The Measurement, Dynamics, and Interpretation of Biological Diversity

by

Sonia Natalie Mitchell

A thesis submitted for the degree of

Doctor of Philosophy (Ph.D.)

Institute of Biodiversity, Animal Health and Comparative Medicine

College of Medical, Veterinary and Life Sciences

University of Glasgow

April 8, 2019

Abstract

The measurement of diversity reflects the variation of *types* (any categorical unit) across one or multiple populations and the distribution of individuals among those types. Clearly useful, a plethora of measures have been developed across widely varying fields ranging from ecology to information theory, economics, and physics. Alarmingly, however, a review of biological literature reveals considerable confusion and disagreement, made worse by the misuse and misinterpretation of measures, conflicting results, and semantic ambiguity. In order to devise a measure of diversity that can be understood across multiple research fields, it must be theoretically well-grounded, powerful, flexible, and robust. Recently, a new framework of diversity measures was developed by Reeve et al. (2016) that satisfies these criteria. This framework of measures is novel in that alpha, beta, and gamma diversity can be assessed at not only the metacommunity level but also the underlying subcommunity level. In addition to this, information on similarity between species can be tailored to suit a particular problem without changing the measures being used. This thesis examines this framework of diversity measures and explores its utility and robustness to many aspects of diversity measurement, in particular beta diversity, the measurement of variation across communities

First, the 'rdiversity' software package was developed in R to calculate these measures (Chapter 2). These measures are then examined in detail, in Chapter 3, by comparing results obtained from these measures to known features in three distinct case studies. The first case study showcases how each measure can be used to extract different signals from a population, by investigating the spatial and temporal biodiversity of the Barro Colorado Island (BCI) Forest dynamics plot. The next two case studies demonstrate the flexibility and utility of these measures by applying diversity-based solutions to more unusual applications, investigating: the demographic diversity of the 2001 population census of England and Wales; and the transmission of antimicrobial resistance between sympatric human and animal host populations.

Chapter 4 extends on the framework to develop new methods to analyse phylogenetic beta diversity. These methods are compared to traditional measures of phylogenetic beta diversity using detailed simulations. Experiments were designed to explore how well each measure was able to detect phylogenetic signals

in community structure (varying the number of tips in the phylogeny, the number of subcommunities, evolutionary rate, whether a phylogeny is ultrametric or non-ultrametric, whether data is incidence-based or abundance-based, nestedness vs. turnover, and so on). Results showed that these measures, particularly the exponentially transformed phylogenetic distance-based beta diversity measures are the most robust across all measures tested, having the greatest power to detect community structure in almost all cases across all measures tested. Following this, a case study highlights the utility of these measures, using phylogenetic data to assess transmission of antimicrobial resistance between human- and animal-origin isolates of Salmonella DT104.

In the final results chapter (Chapter 5), I consider the issue of inaccuracies arising from incomplete sampling, which are ubiquitous in diversity measurement. The robustness of the diversity framework is tested comprehensively under two distinct sampling strategies, reducing the sampling effort per unit area and reducing the area sampled. This is explored using the fully sampled 50 ha BCI Forest dynamics plot dataset, where results show that subcommunity measures are particularly robust to subsampling by subcommunity (reducing the area sampled).

The aim of this thesis is to demonstrate that whilst the concept of diversity continues to be shrouded in ambiguity – where a review of literature reveals as many measures of diversity as possible research questions – there now exists a single framework of flexible and robust measures capable of detailed analyses across different data types, resolutions, and applications.

Dedication

I'd like to thank Gavin Lawson, for his unwavering support[1] and encouragement during the past few years, without which this work surely wouldn't exist.

Richard Reeve, for taking me on as a PhD student, for his extensive supervision and guidance, and for dragging me over the finish line. Even when times were tough, his boundless enthusiasm gave me hope.

Louise Matthews and Christina Cobbold, for giving me confidence in my work, and for showing me the way when I couldn't see the light.

Nardus Mollentze, for being a wonderful bench-mate, a soundboard for ideas, helping me with coding issues, enduring my distractions, and being a good friend. As well as the past and present members of office 308, who are all very lovely.

And last but not least, my parents, Frank and Erliena Mitchell, for making me go to school when I was little, and most importantly, for always believing in me.

[1] Not only patiently listening to my PhD woes, but reassuring me when times were tough, supplying me with endless amounts of coffee, and picking me up from the pub

Contents

List of Tables

List of Figures

Acknowledgements

The work contained within this thesis was funded by BBSRC Doctoral Training Programme grant BB/J013854/1 and was conducted under the supervision of Richard Reeve, Louise Matthews, and Christina Cobbold.

Some of the text describing diversity measures in Chapter 2 was taken from a joint publication with my supervisors (Reeve et al., 2016).

Data was provided from many sources, which are acknowledged where they are used. Specifically, in Chapter 3: the Barro Colorado Island Forest dataset (also used in Chapter 5) was collected with the support of the Center for Tropical Forest Science (CTFS) of the Smithsonian Tropical Research Institute (Condit et al., 2012b, 2017a); the 2001 census of England and Wales was obtained from the Office for National Statistics (2001); and the phenotypic antimicrobial resistance data was provided by Alison Mather (Mather et al., 2012), but originally obtained from the Scottish *Salmonella*, *Shigella* and *Clostridium difficile* Reference Laboratory (SSSCDRL). In Chapter 4, the phylogenetic antimicrobial resistance data analysed in the case study was provided by Alison Mather (Mather et al., 2014), as was the genetic sequence data in Appendix D. The R code used to generate simulated outbreaks in Appendix E was provided by Antonello Di Nardo. The format of the thesis itself was adapted from a latex template developed by Matt Denwood (University of Copenhagen).

Acknowledgement should also be given to my thesis examiners Rebecca Mancy (University of Glasgow) and Hanna Tuomisto (University of Turku), who provided many useful comments and suggestions.

Declaration

I declare that this thesis, and the work contained within it – carried out between September 2014 and April 2019 at the University of Glasgow – is my own, unless otherwise stated or acknowledged, and no part of it has been submitted for any other degree or qualification.

Sonia Natalie Mitchell

Key concepts

Several key terms and concepts are used throughout this text. They are defined here for ease of reference.

Standard mathematical notation is used for any symbolic representation used in mathematical expressions (Table 1).

Symbol	Description
x	Scalars (numbers) are denoted in lowercase, italic, regular font.
$\boldsymbol{x} = (x_1, x_1, \ldots, x_n)$	Vectors are denoted in lowercase, bold, italic font. Elements of a vector are denoted as scalars and enclosed in round brackets.
$x_i \in \{x_1, x_2, \ldots, x_n\}$	Elements of a set are denoted as scalars and enclosed in curly brackets.
$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1N} \\ x_{21} & x_{22} & \ldots & x_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{S1} & x_{S2} & \ldots & x_{SN} \end{bmatrix}$	Matrices are denoted in uppercase, bold, regular font. Elements of a matrix are denoted as scalars, with row-column subscripts.

Table 1: Standard mathematical notation

Weighted power means (also referred to as *weighted generalised means*) are used to calculate many well-known diversity metrics. The r^{th} power mean of $\boldsymbol{v} = (v_1, \ldots, v_n)$, weighted by $\boldsymbol{u} = (u_1, \ldots, u_n)$ is defined as:

$$\mathrm{M}_r(\boldsymbol{u}, \boldsymbol{v}) = \left(\sum_{i:v_i>0}^{n} u_i v_i^r \right)^{1/r} \qquad \text{if } r \neq 0, \pm\infty \tag{1}$$

where r is the *order* of the power mean, and $\sum_i u_i = 1$. Approaching the limits, special cases

are defined as:

$$\mathrm{M}_{-\infty}(\boldsymbol{u},\boldsymbol{v}) = \lim_{r\to-\infty}\mathrm{M}_r = \min_{i:v_i>0}(v_i) \qquad \text{(minimum)} \tag{2}$$

$$\mathrm{M}_0(\boldsymbol{u},\boldsymbol{v}) = \lim_{r\to 0}\mathrm{M}_r = \prod_{i:v_i>0}^{n} v_i^{u_i} \qquad \text{(geometric mean)} \tag{3}$$

$$\mathrm{M}_{\infty}(\boldsymbol{u},\boldsymbol{v}) = \lim_{r\to\infty}\mathrm{M}_r = \max_{i:v_i>0}(v_i) \qquad \text{(maximum)} \tag{4}$$

Standard diversity notation (defined below) is used for consistency and ease of understanding. Throughout this thesis formulae and nomenclature are adapted from the original notation in the literature into a standard form. Following Jost (2007), H is used to represent raw diversity indices that are functions of the basic sum $\sum_{i:p_i>0}^{S} p_i^q$, such as species richness, Shannon entropy, Simpson concentration, the Gini-Simpson index, the Berger-Parker index, and Rényi entropies (Table 2). Any measure of diversity defined in terms of effective numbers of species are represented by the symbol D (Table 3).

The relative abundance of S species in a single community is given by the vector $\boldsymbol{p} = (p_1, \ldots, p_S)$, where p_i is the relative abundance of the i^{th} species. By definition, relative abundances are non-negative ($p_i \geq 0$) and are normalised within a population, such that $\sum_i p_i = 1$.The relative abundance of S species in N subcommunities is given by the $S \times N$ matrix, $\boldsymbol{P}$, where P_{ij} is the relative abundance of species i in subcommunity j and $w_j = \sum_i P_{ij}$ is the relative size of subcommunity j. Relative abundances are normalised within the metacommunity, such that $\sum_i \sum_j P_{ij} = 1$.

Symbol	Equation	Equivalent to	Section
$H_{Richness}(\boldsymbol{p})$	$\mathrm{M}_1\left(\boldsymbol{p}, {}^1\!/\!_{\boldsymbol{p}}\right)$	Species richness[1]	1.2.1
$H_{Shannon}(\boldsymbol{p})$	$\log \mathrm{M}_0\left(\boldsymbol{p}, {}^1\!/\!_{\boldsymbol{p}}\right)$	Shannon entropy	1.2.3
$H_{Simpson}(\boldsymbol{p})$	$\left(\mathrm{M}_{-1}\left(\boldsymbol{p}, {}^1\!/\!_{\boldsymbol{p}}\right)\right)^{-1}$	Simpson concentration	1.2.4
$H_{GiniSimpson}(\boldsymbol{p})$	$1 - \left(\mathrm{M}_{-1}\left(\boldsymbol{p}, {}^1\!/\!_{\boldsymbol{p}}\right)\right)^{-1}$	Gini-Simpson index	1.2.4
$H_{BergerParker}(\boldsymbol{p})$	$\left(\mathrm{M}_{-\infty}\left(\boldsymbol{p}, {}^1\!/\!_{\boldsymbol{p}}\right)\right)^{-1}$	Berger-Parker dominance	1.3
${}^qH(\boldsymbol{p})$	$\log \mathrm{M}_{1-q}\left(\boldsymbol{p}, {}^1\!/\!_{\boldsymbol{p}}\right)$	Rényi entropy	1.3

[1] Although $H_{Richness}$ and species richness have the same numerical value, the former is calculated from proportional abundance data, whereas the latter is not.

Table 2: Standard notation - entropy

Name	Symbol	Equation	Section
Hill numbers	${}^{q}D(\boldsymbol{p})$	$\mathrm{M}_{1-q}\left(\boldsymbol{p}, {}^{1}/_{\boldsymbol{p}}\right)$	1.4
Routledge's alpha	${}^{q}D_{\alpha,R}(\boldsymbol{P})$	$\left(\sum_{j=1}^{N} w_j \sum_{i=1}^{S} p_{ij}^{q}\right)^{1/1-q}$	Section
Jost's alpha	${}^{q}D_{\alpha,J}(\boldsymbol{P})$	$\left(\sum_{j=1}^{N} \frac{w_j^q \sum_{i=1}^{S} p_{ij}^{q}}{\sum w_j^q}\right)^{1/(1-q)}$	1.6.2
Jost's gamma	${}^{q}D_{\gamma,J}(\boldsymbol{P})$	$\sum_{i=1}^{S}\left(p_i^q\right)^{1/(1-q)}$	1.6.2
Jost's beta	${}^{q}D_{\beta,J}(\boldsymbol{P})$	${}^{q}D_{\gamma}/{}^{q}D_{\alpha}$	1.6.2
Leinster & Cobbold's similarity-sensitive diversity	${}^{q}D^{\mathbf{Z}}(\boldsymbol{p})$	$\mathrm{M}_{1-q}\left(\mathbf{Z}\boldsymbol{p}, {}^{1}/_{\mathbf{Z}\boldsymbol{p}}\right)$	1.5.2
Leinster & Cobbold's naïve-type diversity	${}^{q}D^{\mathbf{I}}(\boldsymbol{p})$	$\mathrm{M}_{1-q}\left(\mathbf{I}\boldsymbol{p}, {}^{1}/_{\mathbf{I}\boldsymbol{p}}\right) = {}^{q}D(\boldsymbol{p})$	1.5.2

Table 3: Standard notation - diversity

Reeve et al.'s (2016) similarity-sensitive diversity framework is written using separate notation to avoid confusion. The measures are defined at the species-, subcommunity-, and metacommunity-levels (Species, SC, and MC, respectively). See Table 4 for specific notation and Chapter 2 for explanation.

At the metacommunity level, the general expression for similarity-sensitive diversity of type T is written:

$$ {}^{q}T^{\mathbf{Z}}(\boldsymbol{P}) = \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{q}\boldsymbol{\tau}_j^{\mathbf{Z}}|_{j\in\{1\ldots N\}}\right) \tag{5} $$

which is simply the $1-q$ weighted power mean of the N subcommunity-level diversity values (${}^{q}\boldsymbol{\tau}_j^{\mathbf{Z}}$), weighted by the relative size (w_j) of each subcommunity. Similarly, the general expression for subcommunity-level similarity-sensitive diversity is written:

$$ {}^{q}\tau_j^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}_{\cdot j}, \boldsymbol{t}_{ij}^{\mathbf{Z}}|_{i\in\{1\ldots S\}}\right) \tag{6} $$

Name	Symbol			Section
	Species	SC	MC	
Raw alpha	$a_{i,j}$	${}^{q}\alpha_j^{\mathbf{Z}}$	${}^{q}A^{\mathbf{Z}}$	2.2.4.2
Normalised alpha	$\bar{a}_{i,j}$	${}^{q}\bar{\alpha}_j^{\mathbf{Z}}$	${}^{q}\bar{A}^{\mathbf{Z}}$	2.2.4.1
Raw beta	$b_{i,j}$	${}^{q}\beta_j^{\mathbf{Z}}$	${}^{q}B^{\mathbf{Z}}$	2.2.6
Normalised beta	$\bar{b}_{i,j}$	${}^{q}\bar{\beta}_j^{\mathbf{Z}}$	${}^{q}\bar{B}^{\mathbf{Z}}$	2.2.6.3
Raw rho	$r_{i,j}$	${}^{q}\rho_j^{\mathbf{Z}}$	${}^{q}R^{\mathbf{Z}}$	2.2.6.2
Normalised rho	$\bar{r}_{i,j}$	${}^{q}\bar{\rho}_j^{\mathbf{Z}}$	${}^{q}\bar{R}^{\mathbf{Z}}$	2.2.6.1
Gamma	$g_{i,j}$	${}^{q}\gamma_j^{\mathbf{Z}}$	${}^{q}G^{\mathbf{Z}}$	2.2.5

Table 4: Standard notation - Reeve et al.'s (2016) framework

which is the $1-q$ weighted power mean of the species-level components, $\boldsymbol{t}_{\cdot j}^{\mathbf{Z}} = (t_{1j}^{\mathbf{Z}}, \ldots, t_{Sj}^{\mathbf{Z}})$, where $t_{ij}^{\mathbf{Z}}$ is the species-level component of the i^{th} species in subcommunity j, and $\bar{\boldsymbol{P}}_{\cdot j}$ is the normalised relative abundance of subcommunity j (where $\bar{\boldsymbol{P}}_{\cdot j} = \boldsymbol{P}_{\cdot j}/w_j$).

Throughout this thesis, notation follows that used by Reeve et al. (2016). In some limited special cases, however, some of the measures in Table 4 correspond to existing measures from Table 3:

- ${}^{q}\bar{\alpha}_{j}^{\mathbf{Z}} = {}^{q}D^{\mathbf{Z}}(\bar{\boldsymbol{P}}_{\cdot j})$
- ${}^{q}\bar{G}^{\mathbf{Z}} = {}^{q}D^{\mathbf{Z}}(\boldsymbol{p})$
- ${}^{q}\bar{\alpha}_{j}^{\mathbf{I}} = {}^{q}D(\bar{\boldsymbol{P}}_{\cdot j})$
- ${}^{q}\bar{G}^{\mathbf{I}} = {}^{q}D(\boldsymbol{p}) = {}^{q}D_{\gamma,J}(\boldsymbol{P})$
- ${}^{q}\bar{A}^{\mathbf{I}} = {}^{q}D_{\alpha,R}(\boldsymbol{P})$
- ${}^{q}\bar{A}^{\mathbf{I}} = {}^{q}D_{\alpha,J}(\boldsymbol{P})$ when $q = 1$ or $w_1 = w_2 = \cdots = w_N = \frac{1}{N}$
- ${}^{1}\bar{B}^{\mathbf{I}} = {}^{1}D_{\beta,R}(\boldsymbol{P}) = {}^{1}D_{\beta,J}(\boldsymbol{P})$

Chapter 1

Introduction to diversity

> 'The term 'species diversity' has been defined in such various and disparate ways that it now conveys no information other than "something to do with community structure"; species diversity has become a nonconcept.'
>
> — Hurlbert (1971)

1.1 The problem with diversity

Diversity is a ubiquitous feature of the world around us, with applications ranging from the prioritisation of functional and phenotypic biodiversity in conservation management (Vane-Wright et al., 1991; Forest et al., 2007; Cadotte et al., 2011) to the regulation of genetic diversity in livestock and agricultural ecosystems (Heal et al., 2004; Hajjar et al., 2008; Groeneveld et al., 2010), and the study of antigenic diversity in viral systems (Lipsitch & O'Hagan, 2007; Mumford, 2007; Rambaut et al., 2008). Although the measurement of diversity is often clearly desirable, this somewhat abstract concept has often been difficult to define in practice. Deeply rooted in fields as diverse as ecology, economics, and information theory, diversity can be measured in almost any system[1], requiring only that a community be apportioned categorically. The cross-pollination of measures across such varied fields has resulted in terminological ambiguity, to the extent that even a brief review of literature reveals 'much confusion and pointless argument' (Adams et al., 1997). This is demonstrated by conflicting results, misuse and misinterpretation of measures, and a general disagreement surrounding the most basic diversity-related concepts (see Jurasinski et al., 2009, and associated commentary). Despite (or because of) these acknowledged problems, a solution has yet to be agreed upon.

In recent years, there have been many attempts to unify these measures within a common framework, culminating with the motivation of this thesis, "*How to partition diversity*" (Reeve et al., 2016), which introduces a new framework of similarity-sensitive diversity measures. The purpose of this thesis is to analyse the properties of these measures and investigate their utility, extend these measures for phylogenetic analysis, and develop novel applications in the study of taxonomic diversity, antimicrobial resistance, and viral transmission. To this end, Chapter 1 begins by describing the fundamental concepts underlying the measurement of diversity.

[1] *e.g.* pharmacology (Martin et al., 1995), genetics (Lewontin, 1972), psychology (Junge, 1994), sociology (Lieberson, 1969), archaeology (Leonard & Jones, George, 1989), and so on

1.2 Basic measures of diversity

> 'Expression by means of mathematical symbols frees us from the dangers of the rhetorical inertia of our verbal reasoning and allows us to advance more surely and rapidly in the development of our science.'
>
> — Margalef (1958)

1.2.1 Species richness

At the most basic level, the diversity of an ecological community may be defined as the number of distinct species present, or *species richness* (McIntosh, 1967; Peet, 1974). Simply put, species richness is a count of diversity units – generically denoted here as *types* – be it species or higher taxonomic classifications, functional groups, trait states, lengths of evolutionary history, or any biologically meaningful unit (for simplicity and readability, Chapters 1 & 2 instead refer to *species*). This index of diversity disregards the relative abundance of each species, such that common species are given just as much importance as rarities (and transient species). Species richness is defined, for a community $\boldsymbol{p} = (p_1, \ldots, p_S)$, as:

$$H_{Richness} = \sum_{i=1}^{S} I_{p_i>0} \tag{1.1}$$

where p_i denotes the relative abundance of species i, and the indicator variable:

$$I_{p_i>0} = \begin{cases} 1 & \text{if } p_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1.2}$$

which ensures that only species that are present are counted. To illustrate this, consider the following communities, A and B:

A: ❀ ❖ ♥ ✚ ☆ ☆ ☆ ☆ ☆ ☆

B: ❀ ❀ ❖ ❖ ♥ ♥ ✚ ✚ ☆ ☆

Assuming each distinct symbol represents a distinct species, then a simple count reveals a species richness of 5 in each community. Although the frequency of each species is different, species richness remains the same. The relative abundance of each of these species is not important. Species richness is easy to quantify and interpret, but fails to take into account the relative abundance of species, and therefore considers each of these communities to be equally diverse.

1.2.2 Species evenness

According to Lloyd & Ghelardi (1964), diversity consists of two components, depending not only on species richness, but also on 'equitability' – more commonly referred to as *evenness*. Evenness describes the quantitative equilibrium of the distributional abundance between species, and is typically measured as the ratio of observed diversity to maximum diversity; though many other indices have been proposed (Margalef, 1958; Lloyd & Ghelardi, 1964; Pielou, 1967; McIntosh, 1967; Hill, 1973; Heip, 1974; Peet, 1974, 1975; Alatalo, 1981; Routledge, 1983; Smith & Wilson, 1996; Help et al., 1998; Mouillot & Wilson, 2002; Ricotta, 2004; Tuomisto, 2012). In this context, evenness is low when there is high variation in relative abundance among the species, reaching a maximum in a heterogeneous community when species are uniformly distributed.

1.2.3 Shannon entropy

Abundance weighted indices take into account both richness and the evenness of species. One such example, common in the ecological literature, is Shannon entropy[1]. *Shannon entropy* was originally proposed by information theorist Claude Shannon (1948). It measures the average entropy associated with the outcome of an experiment prior to carrying it out, and can be described as a measure of information content, uncertainty, or 'surprise'. In other words, the uncertainty in predicting the outcome of an experiment is equivalent to the average amount of information gained in observing its result.

In ecological terms, Shannon entropy is the average uncertainty associated with predicting the species-identity of a single individual in a sampling process, and is expressed as,

$$H_{Shannon} = -\sum_{i=1}^{S} p_i \log p_i \tag{1.3}$$

where p_i is the relative abundance of the i^{th} species in a community comprising S species. In this case, entropy reaches a maximum of $\log S$ when all species are equally abundant (all possibilities are equally probable) and a minimum of 0 as the community approaches saturation by a single species (tending towards a single possibility). In other words, the greater the species richness and the more even their distribution (as common species reduce in frequency) the greater the uncertainty in knowing the identity of an individual and the greater the entropy of the system as a whole. Conversely, the lower the species richness and the less even their distribution (as a community becomes dominated by few species) the greater the probability of predicting an observed species and the lower the value of entropy.

[1] Also known as *Shannon information*, the *Shannon-Wiener index*, or in physics, as the *Boltzmann-Gibbs entropy* when $k = 1$ (Boltzmann, 1866)

1.2.4 Simpson indices

Another widely used definition of diversity is the *Simpson concentration index*[1], which is 'a measure of the concentration of the classification [of individuals within a community]' (Simpson, 1949). More specifically, it describes **the probability with which two individuals, selected at random and with replacement, will belong to the same species**. Like Shannon entropy, Simpson's index is sensitive to both species number and evenness, and is expressed:

$$H_{Simpson} = \sum_{i=1}^{S} p_i^2 \tag{1.4}$$

where p_i is the relative abundance of the i^{th} species. Counter-intuitively, Simpson's index reaches a minimum of $1/S$ when species are evenly distributed (though as Simpson's index tends to 0, the diversity of a community tends to infinity), whereas an index of one denotes no diversity at all – *i.e.* only one species present (Figure 1.1a). As a consequence of this, various transforms are regularly used in the literature, most common being the *Gini-Simpson index* and the *inverse Simpson index*.

The Gini-Simpson index[2], $1 - H_{Simpson}$, describes **the probability with which two individuals, selected at random and with replacement, will belong to different species**; where intuitively, zero represents an absence of diversity and 1 represents infinite diversity (Figure 1.1b). The inverse Simpson index[3], $1/H_{Simpson}$ on the other hand, reaches a minimum of one where only one species is present and with a maximum equal to the number of species present, attaining it only when all species are equally abundant (Figure 1.1c).

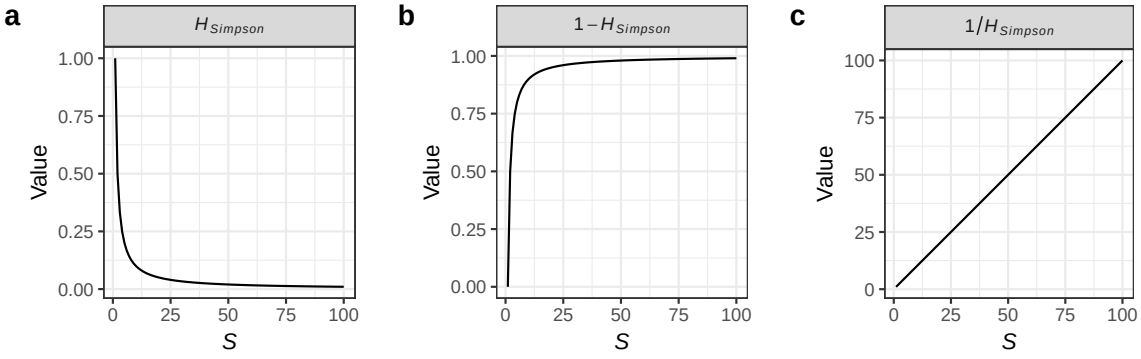


Figure 1.1: Simpson indices measured against species richness: where the x-axis denotes the species richness of evenly distributed communities and the y-axis is the measured value of (a) Simpson concentration index, (b) the Gini-Simpson index, and (c) Simpson diversity.

[1] Also known in economic contexts as the *Herfindahl-Hirschman index* (Hirschman, 1964)
[2] Equivalent to, but independently derived from the *Gini-coefficient* (Gini, 1912)
[3] Also referred to as *Simpson diversity*, or in physics as the *inverse participation ratio*

1.3 Diversity as 'effective numbers'

> '.. the notion of diversity is little more that the notion of the effective number of species present.'
>
> — Hill (1973)

The *effective number of species*[1] is the equivalent number of equally-abundant, virtual species necessary to provide the same diversity as the measured community[2]. This is nicely described by Jost (2006) in that if one were to consider all communities, differing in species richness and evenness, that are equivalent to one another in terms of their diversity index (*e.g.* Simpson's concentration or Shannon entropy), then it is reasonable to suppose that there exists a single community within this set comprising equally-common species, *i.e.* with species of relative abundance $p_i = 1/S$ for all $i = (1, \ldots, S)$. The number of equally-common species in this hypothetical community can be regarded as the effective number of species for any community in this set[3].

The idea of diversity as an effective number of species has been around for some time. Robert MacArthur (1965), citing Shannon entropy as an information-theoretic approach to species diversity, highlights its conversion – for ease of interpretation – to a species richness scale, or an 'equivalent number of equally common species'. Some years later, Mark Hill (1973) derived a class of measures now known as *Hill numbers*, which quantify the effective number of species, whilst varying sensitivity to species rarity. Regrettably, Hill numbers languished in relative obscurity until Lou Jost (2006) reaffirmed their importance, stressing that they 'share a common set of intuitive mathematical properties and behave as one would expect of a diversity, while raw indices do not'. These properties can be illustrated by comparing traditional diversity indices to their effective number equivalents. For example, it is reasonable to expect that a community comprising species of equal abundance should return a value of diversity equivalent to its species richness, and though many traditional indices fail in this regard, their effective number equivalents do not. Consider, the following community:

A: ❀ ❀ ❖ ❖ ♥ ♥ ✚ ✚ ☆ ☆

Shannon entropy highlights the uncertainty in predicting the species-identity of a sample and therefore, the information content of a sampled individual ($H = 1.609$), whilst Simpson's index of concentration calculates the probability of sampling two individuals belonging to the same species ($H = 0.2$). Converting these indices into their effective number equivalents, reveals

[1] Also known in economic contexts as the *numbers equivalent* (Adelman, 1969; Patil & Taillie, 1982) and in physics as the *number of states*

[2] Equivalent to the number of equally-abundant species needed to get the same mean species abundance as observed for the actual (non-equally abundant) species (Tuomisto, 2010a,b, 2011)

[3] When the effective number of species is not an integer, this hypothetical community can never truly exist, since the number of species in a real community must always be an integer

that this community contains *effectively* 5 species (since species richness, the exponential of Shannon entropy, and the inverse of Simpson's index all equal 5). See Table 1.1 for conversions.

The consequences of this are plain when comparing multiple communities. Consider for example, a set of communities of equally-common species, each with a higher species richness than the last. If the diversity of each community were assessed with Shannon entropy, then at higher levels of species richness it becomes increasingly difficult to distinguish between communities (Figure 1.2a). This problem is exaggerated with the Gini-Simpson index, where at much lower values of species richness, communities become almost indistinguishable (flattening in Figure 1.2b). However, when these indices are transformed into their effective number equivalents, the resultant values are directly proportional to the number of species existing within each community (Figures 1.2c & 1.2d). Intuitively, this means that if the number of equally abundant species in a community were to double, its diversity should double too, and vice versa (enabling meaningful comparisons to be made between communities). This fundamental mathematical property is commonly known as Hill's *doubling property*[1] (Hill, 1973).

Table 1.1: Conversion of traditional indices to effective numbers: where p_i is the relative abundance of the i^{th} species in a community, q is the order of diversity, and S is the number of species; adapted from Jost (2006).

Raw Index, H		**Diversity,** D
Measure	**Formula**	
Species richness	$\sum_{i=1}^{S} p_i^0$	H
Shannon entropy	$-\sum_{i=1}^{S} p_i \log p_i$	e^H
Simpson concentration	$\sum_{i=1}^{S} p_i^2$	$\frac{1}{H}$
Gini-Simpson index	$1 - \sum_{i=1}^{S} p_i^2$	$\frac{1}{1-H}$
Rényi entropy	$\frac{1}{q-1} \log \sum_{i=1}^{S} p_i^q$	e^H
Berger-Parker	$\max_{i:p_i>0} p_i$	$\frac{1}{H}$

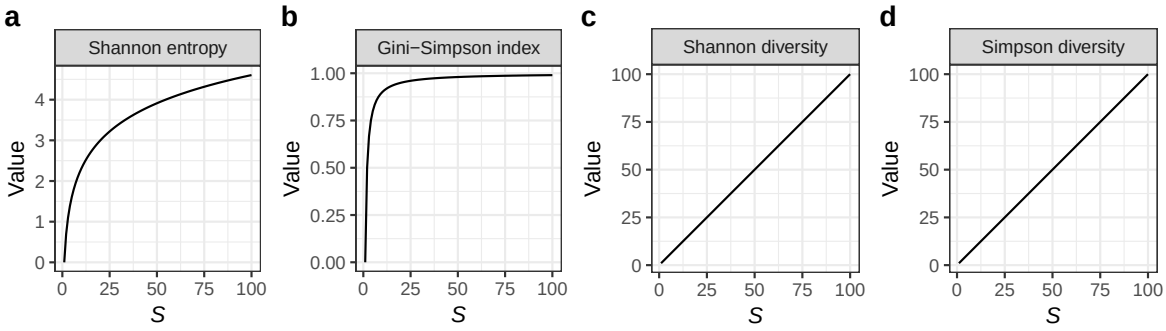


Figure 1.2: Comparing traditional diversity indices on communities of increasing size and species richness: (a) Shannon entropy and (b) Gini-Simpson index versus (c) Shannon diversity and (d) Simpson diversity.

[1] Also referred to as the *replication principle* (Ricotta, 2008), *replication invariance* (economics), and as *Dalton's Principle of Population* (in the context of income inequality, Dalton, 1920)

1.4 Hill numbers

> 'Diversity is not meaningless but has been confounded with the indices used to measure it.'
>
> — Jost (2006)

In 1973, Mark Hill defined a family of diversity measures commonly known as *Hill numbers*[1]. These were developed in response to growing confusion surrounding measures of biodiversity, most notably Hulbert's description of diversity as a 'nonconcept' (Hurlbert, 1971), which arose from a general agreement that whilst many plausible indices were available, none could be proven superior to the rest.

Hill (1973) observed that 'a diversity index is not necessarily itself "diversity"', in that many of the most commonly used diversity indices are special cases of a generalised family of entropies developed by Alfréd Rényi (1961). The Rényi entropy of order q is defined as,

$$ {}^{q}H(\boldsymbol{p}) = \begin{cases} \log\left(\left(\sum_{i:p_i>0} p_i^q\right)^{1/1-q}\right) & \text{if } q \neq 1 \\ -\sum_{i:p_i>0} p_i \log p_i & \text{if } q = 1 \end{cases} \tag{1.5} $$

where $\boldsymbol{p} = (p_1, \ldots, p_S)$ is the relative abundance of S species in a community and q varies the degree of significance attributed to common species relative to those of increasing rarity (described in more detail below). These measures are a generalisation of Shannon entropy, which is ${}^{1}H\,(\boldsymbol{p})$. Hill proposed that these indices could be converted into an 'equivalent number of species', to 'enable us to speak naturally' whilst satisfying the mathematical properties intuitively expected from a measure of diversity. Following (Rényi, 1961), he transformed these expressions such that ${}^{q}H(\boldsymbol{p}) = \log {}^{q}D(\boldsymbol{p})$:

$$ {}^{q}D(\boldsymbol{p}) = \begin{cases} \left(\sum_{i:p_i>0} p_i^q\right)^{1/1-q} & \text{if } q \notin \{1, \infty\} \\ \prod_{i:p_i>0} {p_i}^{-p_i} & \text{if } q = 1 \\ \min_{i:p_i>0} p_i^{-1} & \text{if } q = \infty \end{cases} \tag{1.6} $$

which can also be written as a generalised power mean, ${}^{q}D(\boldsymbol{p}) = \mathrm{M}_{1-q}(\boldsymbol{p}, {}^{1}/_{\boldsymbol{p}})$, where different means are equivalent to commonly used measures of diversity (described in Key Concepts). When we consider the simple case where all species are evenly distributed (so $p_i = {}^{1}/_{S}$ for all i), then ${}^{q}D(\boldsymbol{p}) = S$ for all q. This is the maximum value diversity can take for S species. As the distribution becomes less uniform the diversity of the community drops and ${}^{q}D(\boldsymbol{p})$ decreases as q increases. Analogous to Rényi entropy, these measures – *Hill numbers* – are a generalisation of Shannon diversity, which is ${}^{1}D\,(\boldsymbol{p})$, the exponential of Shannon entropy.

[1] Chao et al. (2014) provide Hill numbers for incidence data based on the Bernoulli product model

According to Hill (1973), the 'diversity number is figuratively a measure of how many species are present if we examine the sample down to a certain depth among its rarities'. This is varied by the parameter q, the *order of diversity*[1], which determines the relative importance attributed to species of differing rarity (for $0 \leq q \leq \infty$). At $q = 1$, each species is weighted exactly by its own abundance. As q decreases, rare species are weighted more strongly and a smaller fraction of the 'effective number of species' is contributed by common species. For example, at $q = 0$, *species richness* describes precisely the number of distinct species in a community by weighting rare and common species equally:

$$ {}^{0}D(\boldsymbol{p}) = \sum_{i:p_i>0} p_i^0 \tag{1.7} $$

As q increases, common species are weighted much more strongly and rare species contribute a much smaller fraction of the 'effective number of species'. For example, at $q = 2$, the *inverse of Simpson's index* is weighted more strongly by species of greater abundance and is therefore less sensitive to rare species:

$$ {}^{2}D(\boldsymbol{p}) = \frac{1}{\sum_{i:p_i>0} p_i^2} \tag{1.8} $$

In the most extreme case, when $q = \infty$, diversity depends only on the most abundant species and disregards rare species completely. This is simply the *inverse Berger-Parker index* (Berger & Parker, 1970), which is written as,

$$ {}^{\infty}D(\boldsymbol{p}) = \frac{1}{\max\limits_{i:p_i>0} p_i} = \min_{i:p_i>0} p_i^{-1} \tag{1.9} $$

where $\max\limits_{i:p_i>0} p_i$ is the relative abundance of the most common species.

In contrast to their associated entropies, Hill numbers are expressed as effective numbers (Section 1.3). This is important as it enables different communities to be easily compared at each value of q. Additionally, a single community may be observed across multiple values of q, combining many well-known indices along a continuum of values known as a *diversity profile* (as in Hill, 1973; Tóthmérész, 1995; Leinster & Cobbold, 2012). This allows an observer to discern a richer and more comprehensive measure of diversity. An example of such a profile is given in Figure 1.3, which is illustrated subsequently. Consider the following communities:

A: ❀ ❖ ♥ ✚ ☆ ☆ ☆ ☆ ☆ ☆

B: ❀ ❀ ❖ ❖ ♥ ♥ ✚ ✚ ☆ ☆

Any individual diversity measure is a single value that represents a single perspective. For example, species richness tells us that these communities have the same number of species. Clearly, these communities have quite different distributions that species richness fails to capture. The differences in community structure are revealed by comparing diversity profiles

[1] Also termed the *sensitivity parameter* (Leinster & Cobbold, 2012) or the *viewpoint parameter* (Reeve et al., 2016)

(Figure 1.3). The slope of each diversity profile reflects the evenness of the distribution of species across each community from a range of different perspectives (different values of q). Specifically, we find that community A (blue) is dominated by a single species, reflected by the steep initial drop in diversity with increasing q. Whereas, the species in community B (yellow) are evenly distributed, reflected in the constant value of diversity for all values of q. If a third community is introduced:

C: ❀ ✿ ❖ ✦ ♥ ♣ ✚ ✘ ☆ ☆

The distribution of species in community C (red) is more evenly distributed than those in community A (blue), reflected in the slower decline in diversity at lower values of q. At $q = \infty$ the effective number of species in communities C and B is equal, since only the most dominant species is considered.

Are Hill numbers a good measure of diversity? Provided that all species are completely distinct from one another, then yes. They are able to account for the number of species as well as their relative abundances, they satisfy a number of reasonable properties, and can be compared easily and intuitively between different communities (Routledge, 1979; Jost, 2006). In reality however, there are often instances where it is important to consider one species to be similar to another, and in these instances it may be useful to consider indices of similarity-sensitive diversity.

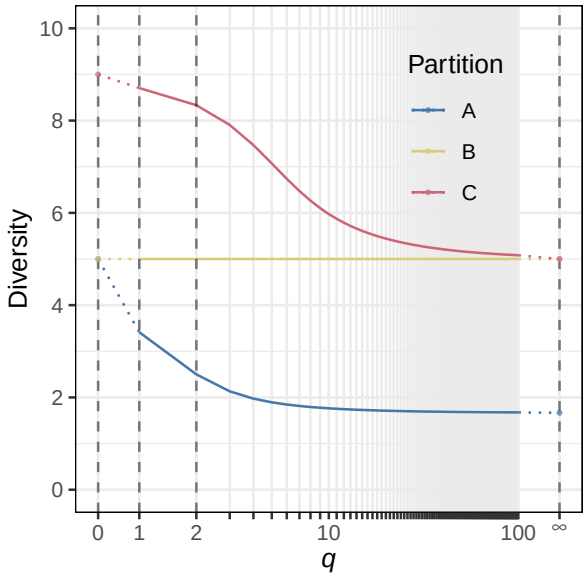


Figure 1.3: Diversity profiles: Contrasting the diversity of three simple communities.

1.5 Similarity-sensitive diversity

> 'Realistic measures of biodiversity should reflect not only the relative abundance of species, but also the differences between them.'
>
> — Leinster & Cobbold (2012)

By the early 1980s, measures of diversity were everywhere, with applications ranging from anthropology (Greenberg, 1956), to genetics (Nei, 1973; Duvick, 1984), economics (Gini, 1912; Hart, 1971), sociology (Horowitz, 1970; Sen, 1973), and ecology (Pielou, 1967; Wilson & Shmida, 1984). At this time, most of these indices considered species (or any meaningful categorical unit) to be completely distinct. The problem with this becomes clear in the following example.

Consider a community comprising two ducks, a cat, a horse, and a donkey. Most traditional measures of diversity regard each species as being identical to itself but completely dissimilar to all other species (A). In reality however, species might share multiple similarities – functional, phenotypic, genetic, *etc.* In this case we might consider that horses and donkeys are 90% similar (A′). If one were to calculate the diversity of this community using traditional measures, then we might consider that to give an overestimation of the true diversity (Figure 1.4), since the degree of differentiation between horses and donkeys in A′ (when similarity is included) is much lower than in A (when similarity is ignored). Capturing this information and accurately quantifying the biodiversity of this community necessitates incorporating some index of similarity into the measurement of diversity.

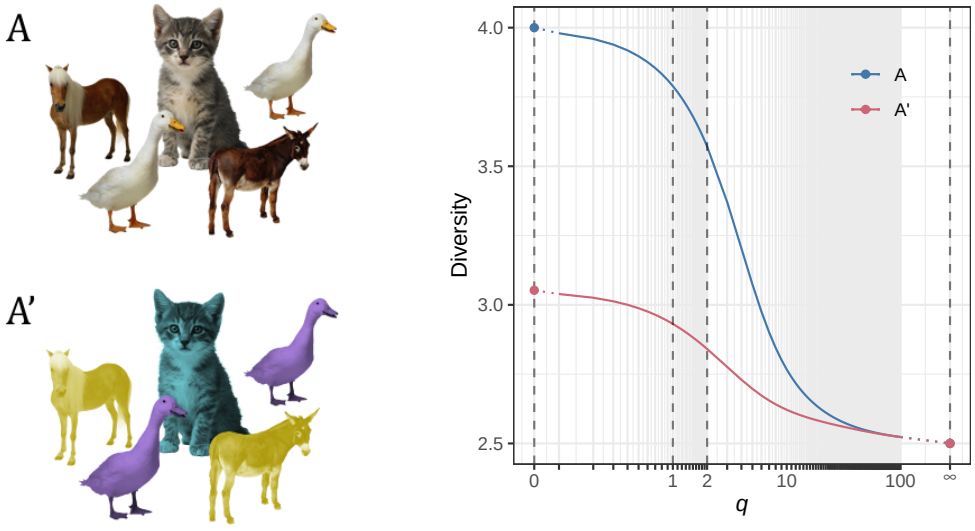


Figure 1.4: Similarity-sensitive diversity profiles: where community A (blue) considers species to be distinct (Hill numbers, Section 1.4); and community A' (red) incorporates species similarity (Leinster & Cobbold's similarity-sensitive diversity, Section 1.5.2).

1.5.1 Rao's quadratic entropy

The most well-known index of similarity-sensitive diversity was developed by statistician, Calyampudi R. Rao (1982) and is commonly known as Rao's *quadratic entropy*[1]. This metric describes **the expected dissimilarity between two individuals sampled at random and with replacement from the same community**, and is written,

$$Q(\boldsymbol{p}) = \sum_{i=1}^{S} \sum_{i'=1}^{S} \Delta_{ii'} p_i p_{i'} \tag{1.10}$$

where $\Delta_{ii'}$ is the distance between species i and i'. In this way, Rao's quadratic entropy combines both the relative abundance of species and the pairwise distance between them. Inter-species distance is zero when both individuals originate from the same species ($\Delta_{ii} = 0$ for all i) and unity when two completely distinct species are sampled ($\Delta_{ii'} = 1$ for all $i \neq i'$). In the special case, when species are completely distinct, Rao's quadratic entropy is essentially a generalisation of Simpson's index of concentration (Simpson, 1949).

1.5.2 Leinster & Cobbold's similarity-sensitive diversity

Approaching this problem from the field of mathematics, Leinster & Cobbold (2012) proposed a *similarity-sensitive* measure of diversity. This framework generalises many of the most popular measures of diversity: species richness, Shannon (1948) entropy, the Gini-Simpson index (Gini, 1912; Simpson, 1949), the Berger-Parker index (Berger & Parker, 1970), Hill (1973) numbers, the Patil-Taille-Tsallis entropies (Patil & Taillie, 1982), and the indices of Hurlbert (1971), and Smith & Grassle (1977). Many similarity-based measures are also closely connected, including: Rao's (1982) quadratic entropy, the entropies of Ricotta & Szeidl (2006), and the phylogenetic indices of Faith (1992), Allen et al. (2009), and Chao et al. (2010), the latter being based on Hill numbers.

Consider a community of S species with relative abundance $\boldsymbol{p} = (p_1, \ldots, p_S)$, where inter-species similarity is encoded in an $S \times S$ *similarity matrix*, $\mathbf{Z}$, which is equivalent to *Rao's dissimilarity* such that $\mathbf{Z} \equiv 1 - \Delta$. The elements of this matrix ($Z_{ii'}$) reflect the similarity between species i and i' on a scale of 0 to 1, such that species are identical when $Z_{ii'} = 1$ and completely dissimilar when $Z_{ii'} = 0$. It follows then, that $(\mathbf{Z}\boldsymbol{p})_i = \sum_{i'} Z_{ii'} p_{i'}$ describes the relative abundance of species similar to species i, defined as the *ordinariness* of the i^{th} species, where similarity can be defined ally, phenotypically, phylogenetically, or in some specific manner determined by the research question under study. If distances ($d_{ii'}$) are available, they must first be translated into similarities such that $0 \leq Z_{ii'} \leq 1$ and $Z_{ii} = 1$. This can be done in a number of ways, for example: by defining $Z_{ii'} = 1 - (\frac{d_{ii'}}{kd_{\max}})$ when $d_{ii'} \leq kd_{\max}$ and $Z_{ii'} = 0$ when $d_{ii'} > kd_{\max}$; or defining $Z_{ii'} = e^{-kd_{ii'}}$ (Nei, 1972) where k is a scaling factor.

[1]Known in interdisciplinary studies as the *Rao-Stirling index* (Stirling, 2007), and independently rediscovered by Ganeshaiah et al. (1997) as the *Avalanche index*

Leinster & Cobbold (2012) derived a general expression for similarity-sensitive diversity, which like Hill (1973) numbers, is based on Rényi's (1961) generalised entropies, and expressed as effective numbers of species. This framework reveals a coherent family of diversity measures, derived as the reciprocal of the "average" ordinariness of a community:

$$ {}^{q}D^{\mathbf{Z}}(\boldsymbol{p}) = \begin{cases} \left(\sum_{i:p_i>0} p_i \left(\mathbf{Z}\boldsymbol{p}\right)_i^{q-1}\right)^{1/1-q} & \text{if } q \notin \{1,\infty\} \\ \prod_{i:p_i>0} (\mathbf{Z}\boldsymbol{p})_i^{-p_i} & \text{if } q = 1 \\ \left(\max\limits_{i:p_i>0} \left(\mathbf{Z}\boldsymbol{p}\right)_i\right)^{-1} & \text{if } q = \infty \end{cases} \tag{1.11} $$

As with Hill numbers, this can be expressed as a weighted power mean,

$$ {}^{q}D^{\mathbf{Z}}(\boldsymbol{p}) = \mathrm{M}_{1-q}(\boldsymbol{p}, 1/\mathbf{Z}\boldsymbol{p}) \tag{1.12} $$

where q defines the type of average being applied. And from here, it is a simple matter to derive Hill numbers, where similarity between species is ignored. To do this, similarity coefficients are defined, $Z_{ii} = 1$ for all i, and $Z_{ii'} = 0$ for all $i \neq i'$. In other words, $\mathbf{Z} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. A simple substitution reveals that $(\mathbf{Z}\boldsymbol{p})_i = (\mathbf{I}\boldsymbol{p})_i = \boldsymbol{p}_i$ and therefore, ${}^{q}D^{\mathbf{I}}(\boldsymbol{p}) = {}^{q}D(\boldsymbol{p})$, where ${}^{q}D(\boldsymbol{p})$ is the Hill number of order q. This can be written as,

$$ {}^{q}D^{\mathbf{I}}(\boldsymbol{p}) = \begin{cases} \left(\sum_{i:p_i>0} p_i \left(\mathbf{I}\boldsymbol{p}\right)_i^{q-1}\right)^{1/1-q} & \text{if } q \notin \{1,\infty\} \\ \prod_{i:p_i>0} (\mathbf{I}\boldsymbol{p})_i^{-p_i} & \text{if } q = 1 \\ \left(\max\limits_{i:p_i>0} \left(\mathbf{I}\boldsymbol{p}\right)_i\right)^{-1} & \text{if } q = \infty \end{cases} \tag{1.13} $$

which, because $\mathbf{I}\boldsymbol{p} = \boldsymbol{p}$ and so $(\mathbf{I}\boldsymbol{p})_i = p_i$, is equal to

$$ {}^{q}D(\boldsymbol{p}) = \begin{cases} \left(\sum_{i:p_i>0} p_i^q\right)^{1/1-q} & \text{if } q \notin \{1,\infty\} \\ \prod_{i:p_i>0} {p_i}^{-p_i} & \text{if } q = 1 \\ \left(\max_{i:p_i>0} p_i\right)^{-1} & \text{if } q = \infty \end{cases} \tag{1.14} $$

This can be written in its general form as,

$$ {}^{q}D^{\mathbf{I}}(\boldsymbol{p}) = \mathrm{M}_{1-q}(\boldsymbol{p}, 1/\mathbf{I}\boldsymbol{p}) = \mathrm{M}_{1-q}(\boldsymbol{p}, 1/\boldsymbol{p}) = {}^{q}D(\boldsymbol{p}) \tag{1.15} $$

which clearly shows how closely related these measures are.

As Leinster & Cobbold (2012) explain, the parameter q allows greater control of the relative emphasis placed on species of differing ordinariness (where $0 \leq q \leq \infty$). Specifically, as q increases, diversity is increasingly weighted by species of high ordinariness, that is, species of high abundance and/or high similarity to other species. For example, similarity-sensitive species richness ($q = 0$), just like its counterpart, equals the species richness of a community irrespective of its relative abundance. However, unlike Hill's formulation, it takes into account the similarity of each species-pair within a community, and so, avoids the inflation in species

richness caused by species of high similarity. Conversely, at $q = \infty$ exists a similarity-sensitive version of the reciprocal of the Berger-Parker index. Whereas Hill's version of $q = \infty$ reflects the dominance of a community with respect to the species of greatest relative abundance, the similarity-sensitive measure takes into account the abundance of similar species, and therefore describes a measure of dominance that considers the most ordinary species within a community. Thus, it is possible that the ordinariness $(\mathbf{Z}\boldsymbol{p})_i$ may be largest for species i, even though the species itself is relatively rare, because it is similar to highly abundant species.

In a single community, when the similarity between species can be quantified, Leinster & Cobbold's (2012) similarity-sensitive diversity measures retain more information than Hill (1973) numbers (that consider species to be completely distinct). However, these measures focus on the underlying variability within a *single community*. The next section looks at beta diversity, which attempts to address the issues of looking at multiple subcommunities, where the way in which diversity is partitioned must be carefully considered.

1.6 Beta diversity

> 'The original definition of "beta diversity" by Whittaker (1960) was already very broad, and since the coining of the concept, it has been tremendously stretched to cover the most varied phenomena'
>
> — Tuomisto (2010b)

In his seminal paper, Robert Whittaker (1960) proposed a new approach for describing the relationship between the heterogeneity of subcommunities and the diversity of species across a broader community. Whittaker developed new terminologies – *alpha*, *beta* and *gamma* – with which to distinguish between these differing aspects of diversity.

(1) *Alpha* diversity, inspired by *Fishers alpha* is described as 'the [average] richness in species of a particular stand or community[1].

(2) *Beta* diversity is somewhat ambiguously described as $beta = {}^{gamma}/_{alpha}$, or 'the extent of change of community composition, or degree of community differentiation, in relation to a complex gradient of environment, or a pattern of environments', which might be measured using indices of compositional similarity such as the *Jaccard*, *Sørensen*, and *Bray-Curtis* indices[2] (Whittaker, 1960, 1972).

(3) *Gamma* diversity is described as the species diversity across all samples in a community, where 'diversity value is a resultant of both [the] alpha and beta diversities of these samples', commonly referred to as *Whittaker's multiplicative law* (Section 1.6.2).

[1] A log series index used to model species diversity at the local scale (Fisher et al., 1943)
[2] Jaccard (1901, 1912); Sørensen (1948); Bray & Curtis (1957)

Though ecologically useful, the general confusion surrounding the concept of *beta diversity* has resulted in a fundamental disagreement regarding its usefulness. Commonly used metrics are often described as 'beta diversities', but are actually measures of compositional heterogeneity, differentiation between subcommunities, relative change in species composition between communities, rate of compositional turnover, distinctness, or nestedness (reviewed in Vellend, 2001; Jurasinski et al., 2009; Anderson et al., 2011). In this chapter, no attempt is made to clarify the conceptual and semantic ambiguity of 'beta diversity'. The reader is instead referred to an extensive two-part review by Hanna Tuomisto, which highlights the range of methods used to develop definitions of beta diversity, and examines the properties of these measures whilst structuring these concepts within a common framework (Tuomisto, 2010a,b). With this in mind and to support further discussion, the following section identifies some of the most commonly used beta-related concepts.

1.6.1 Indices of compositional similarity

Measures of compositional similarity, differentiation, or species overlap, between two samples are numerous in literature (Legendre et al., 2005). Necessarily, studies have been directed at examining the performance of these indices; Wilson & Shmida (1984) and Koleff et al. (2003) review incidence-based measures, whilst Barwell et al. (2015) conducts a similar review of abundance-based measures. Two of the most well-known incidence-based measures are the closely related Jaccard and Sørensen indices, which can be used to quantify variation between communities from presence/absence data. Consider, for example, two communities represented by sets $\boldsymbol{j} = \{S_1, \ldots, S_n\}$ and $\boldsymbol{k} = \{S'_1, \ldots, S'_m\}$, where S_i and S'_i are the species present in communities j and k, respectively.

The *Jaccard index*[1] measures the similarity between two samples and is calculated by measuring the ratio between the intersection and the union of both communities (Jaccard, 1901, 1912). In the language of Jaccard:

$$\beta_{jac} = \frac{|\boldsymbol{j} \cap \boldsymbol{k}|}{|\boldsymbol{j} \cup \boldsymbol{k}|} \tag{1.16}$$

That is, the ratio of the number of species shared by communities j and k relative to the total number of species in the two communities combined. This measure reaches a maximum of 1 when all species are shared, and a minimum of zero when the species in each community are distinct. Conversely, the Jaccard distance[2] can be used to determine compositional dissimilarity between two samples, such that:

$$\beta_{jac'} = 1 - \beta_{jac} = \frac{|\boldsymbol{j} \Delta \boldsymbol{k}|}{|\boldsymbol{j} \cup \boldsymbol{k}|} = \frac{|\boldsymbol{j} \cup \boldsymbol{k}| - |\boldsymbol{j} \cap \boldsymbol{k}|}{|\boldsymbol{j} \cup \boldsymbol{k}|} \tag{1.17}$$

[1] Also referred to as *coefficient of community, or coefficient de communauté*

[2] Also referred to as the *Marczewski-Steinhaus distance* (Marczewski & Steinhaus, 1958), *Tanimoto distance*, or *biotope distance*

Likewise, the *Sørensen-Dice coefficient*[1], independently derived by Thorvald Sørensen (1948) and Lee Dice (1945), is written:

$$\beta_{sor} = \frac{2|\boldsymbol{j} \cap \boldsymbol{k}|}{|\boldsymbol{j}| + |\boldsymbol{k}|} = \frac{2|\boldsymbol{j} \cap \boldsymbol{k}|}{|\boldsymbol{j} \cup \boldsymbol{k}| + |\boldsymbol{j} \cap \boldsymbol{k}|} \tag{1.18}$$

which is related to the Jaccard index by $\beta_{sor} \equiv \frac{2\beta_{jac}}{1+\beta_{jac}}$. Each of these indices compares the distribution of species between two communities, but differ in their perspective. The Jaccard index compares the number of shared species to the cumulative number of species across both communities, whereas the Sørensen index places more importance on species shared by each community, and consequentially compares the number of shared species to the mean number of species within a single assemblage. Both the Jaccard and Sørensen coefficients are commonly used throughout literature, but are limited by their inability to cope with species abundance data.

The Bray-Curtis similarity coefficient[2] – a quantitative analogue of Sørensen's index – is an abundance-based measure used to quantify the dissimilarity between communities $\boldsymbol{j}$ and $\boldsymbol{k}$ (Bray & Curtis, 1957). It is written:

$$\beta_{bc} = \frac{2\sum_{i=1}^{S_j \cap S_k} \min(N_i, M_i)}{\sum_{i=1}^{S_j} N_i + \sum_{i=1}^{S_k} M_i} \tag{1.19}$$

where N_i and M_i denote the total number of individuals of species i in communities j and k, respectively.

1.6.2 Defining the relationship between alpha, beta, and gamma

Multiplicative beta, also known as *Whittaker's multiplicative law*, assumes that the relationship between alpha, beta and gamma diversities is written as:

$$gamma = alpha \times beta. \tag{1.20}$$

Originally defined in terms of Fisher's alpha (Whittaker, 1960), the multiplicative partitioning of diversity was later applied to Shannon diversity (MacArthur, 1965), species richness

[1] Also referred to as *Sørensen binary*, the *Sørensen similarity index*, the *coincidence index*, the *Dice coefficient*, or *binary Sørensen-Dice*

[2] Also referred to as the *Bray-Curtis distance*, the *quantitative Sørensen-Dice index*, the *Sørensen abundance index*, the *proportional similarity index*, or the Odum coefficient, and sometimes incorrectly referred to as Czekanowski's distance (Legendre et al., 2005)

(Whittaker, 1972)[1], and then generalised for all Hill numbers (Routledge, 1979), such that:

$$ {}^{q}D_{\alpha,R} = \begin{cases} \left(\sum_{i:p_i>0} \sum_{j=1}^{N} w_j \bar{P}_{ij}^{q} \right)^{1/(1-q)} & \text{if } q \neq 1 \\ \exp\left(\sum_{j=1}^{N} -w_j \sum_{i:p_i>0} \bar{P}_{ij} \log \bar{P}_{ij} \right) & \text{if } q = 1 \end{cases} \tag{1.21} $$

$$ {}^{q}D_{\gamma} = \begin{cases} \left(\sum_{i:p_i>0} p_i^{q} \right)^{1/1-q} & \text{if } q \notin \{1, \infty\} \\ \prod_{i:p_i>0}^{S} {p_i}^{-p_i} & \text{if } q = 1 \\ \min_{i:p_i>0} p_i^{-1} & \text{if } q = \infty \end{cases} \tag{1.22} $$

$$ {}^{q}D_{\beta,R} = \frac{{}^{q}D_{\gamma}}{{}^{q}D_{\alpha,R}} \tag{1.23} $$

where N is the number of subcommunities and $w_j = \sum_{i=1}^{S} p_{ij}$ is the relative size of subcommunity j. If one were to think of a *metacommunity* as comprising a number of *subcommunities* linked by dispersal (Leibold et al., 2004), then diversity can be partitioned in terms of: Routledge's alpha, ${}^{q}D_{\alpha,R}$, the average effective number of species per subcommunity weighted by the size of the subcommunity; ${}^{q}D_{\gamma}$, the effective number of species in metacommunity as a whole (which is equal to ${}^{q}D$); and ${}^{q}D_{\beta,R}$, the amount of species turnover between subcommunities.

A similar framework, proposed by Jost (2007), uses a different alpha (and therefore beta) component:

$$ {}^{q}D_{\alpha,J} = \begin{cases} \left(\sum_{i:p_i>0} \sum_{j=1}^{N} \frac{w_j^{q}}{w_1^{q}+\cdots+w_N^{q}\bar{P}_{ij}^{q}} \right)^{1/(1-q)} & \text{if } q \neq 1 \\ \exp\left(\sum_{j=1}^{N} -w_j \sum_{i:p_i>0} \bar{P}_{ij} \log \bar{P}_{ij} \right) & \text{if } q = 1 \end{cases} \tag{1.24} $$

$$ {}^{q}D_{\beta,J} = \frac{{}^{q}D_{\gamma}}{{}^{q}D_{\alpha,J}} \tag{1.25} $$

Jost calls these expressions *true alpha diversity*, *true gamma diversity*, and *true beta diversity*, respectively. That is: Jost's alpha, ${}^{q}D_{\alpha,J}$, the average effective number of species per subcommunity weighted by the q^{th} power of the size of the subcommunity; and ${}^{q}D_{\beta,J}$, the number of subcommunities that have no species in common, or the equivalent effective number of compositional units in a community. Routledge's (1979) alpha (advocated by Tuomisto, 2010a) is equivalent to Jost's (2007) alpha (advocated by Chao et al., 2012) when subcommunity weights are equal or at $q = 1$. Since, Routledge and Jost use a multiplicative beta with the same gamma diversity measure, their beta diversity measures are the same under the same conditions (when subcommunity weights are equal or at $q = 1$).

An alternative approach, *additive beta*, demonstrates that traditional indices of 'diversity' (Section 1.3) such as species richness, the Gini-Simpson index, and Shannon entropy, can be

[1]In the 1972 paper, Whittaker also commented on MacArthur's use of the exponential of the Shannon entropy and noted that the raw entropies should not be used.

partitioned additively (MacArthur, 1965; Levins, 1968; Lewontin, 1972; Nei, 1973; Allan, 1975; Patil & Taillie, 1982; Veech et al., 2002; Veech & Crist, 2010; Marcon et al., 2014)[1]. This additive partitioning of 'diversity' was not well known or utilised until its rediscovery by Lande (1996)[2], who redefined the approach using Whittaker's alpha-, beta-, and gamma-notation, such that:

$$gamma = alpha + beta \tag{1.26}$$

where beta diversity is now redefined as 'the average amount of diversity that is *not* found in a single, randomly chosen sample' (Jurasinski et al., 2009). The benefits of each approach have since been debated extensively (Ricotta, 2005; Jost, 2006, 2007; De Bello et al., 2010; Ricotta, 2010; Tuomisto, 2010a; Veech & Crist, 2010; Chao et al., 2012; Marcon et al., 2014).

Additive and multiplicative beta diversities are able to measure the variability across subcommunities, or compare many subcommunities simultaneously. But they are not able to identify specific subcommunities as being different or typical. As discussed in the following chapter, Reeve et al.'s (2016) framework of measures addresses these issues by developing a new formulation for beta diversity with stronger mathematical foundations, but which is neither additive nor multiplicative. This approach also satisfies a list of properties of beta diversity measures proposed by previous authors, which aren't satisfied by any existing measure.

[1]Though MacArthur (1965) was first to partition diversity data additively, using Shannon entropy, he transformed them into their numbers equivalents before interpreting them.

[2]Lande's (1996) additive partitioning of Shannon entropy is related to its multiplicative counterpart through a simple logarithmic transformation

1.7 Thesis outline

The aim of this thesis is to investigate the properties of Reeve et al.'s (2016) framework, to apply these measures to practical applications, and to determine whether different or typical subcommunities can be identified from real data. The framework incorporates a measure of similarity, which can be utilised in the study of phylogenies, for which many customised measures of diversity have already been created. I extend the scope of the framework to handle phylogenetic similarity, and benchmark the resultant measures against existing measures of phylogenetic diversity. Finally, the framework assumes a completely censused metacommunity, which is rarely the case in real systems, so I also investigate how these measures perform with more realistic, incomplete information.

This is done in the following manner:

Chapter 2: Methodological development - the framework of diversity measures proposed in Reeve et al. (2016) is described. Simple examples are used to demonstrate the properties of alpha-, beta-, and gamma-diversity at both the metacommunity and the subcommunity levels. An R package – `rdiversity` – is developed to calculate these measures. The functionality of this package is demonstrated here, whilst also simulating simple community structures to highlight the properties of the diversity framework.

Chapter 3: Validation study - three distinct case studies are highlighted. These are used to validate the framework, demonstrating how these methods can be used to detect signals in real datasets. Diversity-based methods are used to recreate known quantities: (1) the demographic diversity of the human population of England and Wales; (2) the spatial and temporal biodiversity of a well-known forest plot; and (3) the flow of antimicrobial resistance of the Salmonella bacterium between human and animal communities using phenotypic, genetic, and phylogenetic datasets.

Chapter 4: Phylogenetic diversity - Phylogenetic diversity measures provide a means of understanding how evolutionary factors influence patterns of diversity on a temporal scale. In this chapter, traditional measures of phylogenetic diversity are described alongside those derived from the framework. All methods are statistically compared. The chapter concludes with a case study describing the flow of antimicrobial resistance of DT104 using phylogenetic data.

Chapter 5: Subsampling properties - In this chapter, I investigate how well the measures perform under a range of sampling pressures, using data from the Barro Colorado Island (BCI) Forest dynamics plot.

Chapter 6: Discussion.

Chapter 2

Methodological development

> 'The lack of a common framework means that diversity measures from different fields have conflicting fundamental properties, allowing conclusions reached to depend on the measure chosen.'
>
> — Reeve et al. (2016)

2.1 Abstract

The underlying structure of a community is both complex and dynamic, comprising as many measurable variables as sensible avenues of research. Although attempts have been made to explain this complexity in a meaningful way, a single coherent framework for its measurement, analysis, and partitioning has been conspicuously lacking.

In response to this, Reeve et al. (2016) devise an intuitive family of diversity measures, derived from Rényi's (1961) generalisations of Shannon (1948) entropy and Leinster & Cobbold's (2012) expression of similarity-sensitive diversity. These measures expose compositional variability (in terms of taxonomic, phylogenetic, phenotypic, genetic, or functional interest) as alpha, beta and gamma diversities, of both the whole community and its lower, underlying levels.

This chapter considers the following questions:

1. *What are these measures and how do they work?* The properties of each measure are described, alongside a number of worked examples to illustrate their utility.

2. *How are these measures calculated?* A software package – `rdiversity` – was developed in R and published on CRAN to calculate these measures. The functionality of this package is demonstrated with simple examples.

2.2 How to partition diversity

This section follows Reeve et al. (2016), beginning with a description of the notation used in this (and subsequent) chapters. The diversity framework is introduced, after which, simple examples are used to illustrate the utility of each measure.

2.2.1 Introduction

In 1960, Whittaker recognised that local variation in species composition and diversity often exists within a community (Whittaker, 1960). He quantified this by partitioning diversity or multiplicatively in terms of alpha, beta, and gamma diversity. Though useful, there was no consensus on how to measure diversity, as the term gamma diversity was only introduced in Whittaker's paper. Over a decade later, Hill devised an effective number formulation (see Section 1.4), which combines many of these commonly used measures of diversity (Hill, 1973). However, this family of measures languished in obscurity for several decades, before it was championed by Jost (2006), and is now widely accepted as the basis for measurements of gamma diversity.

In his seminal paper, Jost (2007) argued that for a measure of diversity to satisfy required mathematical properties, the partitioning of diversity must be based on Whittaker's multiplicative law (Equations 1.24-1.25). However, he acknowledged that this only works at $q \in \{0, 1\}$ for unequally weighted subcommunities[1], and that there was perhaps no way of partitioning diversity for any other value of q. In addition to this, his methods did not address the similarity between species. Meanwhile Rao (1982) addressed the issue of similarity between species, but only at $q = 2$. In response to this, Leinster & Cobbold (2012) formulated a similarity-sensitive measure of diversity to generalise Rao's approach for all values of q, based on Rényi's (1961) generalised entropies. But three problems remained:

- Could Jost's (2007) partitioning of diversity be extended beyond $q \in \{0, 1\}$?
- Could Leinster & Cobbold's (2012) similarity sensitive diversity be extended to cover the partitioning of gamma diversity into alpha and beta components?
- Could important subcommunities be identified within the community as a whole? (as well as quantifying the variability across them?)

Given the power and generality of Rényi's (1961) entropy-based approach, Reeve et al. (2016) use Rényi's notion of generalised *relative* entropy (the Rényi divergence of order q) to derive measures of the diversity of a subcommunity relative to the metacommunity as a whole (see

[1]For equally weighted subcommunities these measures work for all values of q

Appendix A for an example of this derivation). Rényi's relative entropy (which had previously been ignored in favour of simpler mathematical formulations) is a measure of divergence[1], or a kind of 'distance' between two probability distributions. In terms of diversity, this is the same as asking how close two species distributions are to one another, and in this context compares the distribution of a subcommunity to the distribution of the metacommunity as a whole. This naturally defines subcommunity values (of beta diversity) that provide insight into how different or typical subcommunities are in the context of the metacommunity. As explained in Reeve et al. (2016), these subcommunity measures can be used to construct more traditional metacommunity measures such as gamma and beta diversity. Of these, one resultant beta diversity measure is identical to Jost's beta diversity measures for distinct species at $q \in \{0, 1\}$, but has the advantage of satisfying all of Jost's properties for all values of q. Furthermore, it generalises Leinster and Cobbold's measure of gamma diversity, to give a new similarity-sensitive beta diversity measure. As such, it addresses all three problems introduced above, and in addition to this, it satisfies three new properties (as described in Reeve et al., 2016), which state that:

- Normalised metacommunity (alpha, beta, and gamma) diversities should be *invariant under shattering* of their constituent subcommunities[2];
- All subcommunity (alpha, beta, and gamma) diversities are conditionally independent and therefore invariant to differences in partitioning of the rest of the metacommunity; and
- All subcommunity (alpha, beta, and gamma) diversities should therefore be directly compared within a metacommunity to determine the relative merits of different subcommunities

As already discussed, Reeve's framework breaks down metacommunity-level measures of diversity into subcommunity-level components, which are broken down further into species-level components within each subcommunity. Subcommunity-level diversities are calculated as some form of average (a weighted power mean) of the species-level components that they contain, and metacommunity-level diversities are in turn calculated as an average of their constituent subcommunities. Throughout this chapter I will refer to weighted power mean measures informally as averages for simplicity – so generally, averages are weighted power means not conventional averages, which are specifically referred to as arithmetic means.

In this chapter, I describe each diversity measure in the context of a simple caricature metacommunity (described in the next section). In total there are seven diversity measures:

[1] A generalisation of Kullback-Leibler divergence (Kullback & Leibler, 1951) when $q = 1$

[2] Invariance under shattering refers to the fact that the normalised metacommunity diversity of N subcommunities should not be affected by the incorrect, continued *shattering* of subcommunities, where each subcommunity is assumed to be well-mixed and the breaking subcommunities into subdivisions results in the a relative abundance of species identical to that of the parent subcommunity

two alpha diversity measures, one of which corresponds to traditional notions of alpha diversity; four beta diversity measures, one of which – at the metacommunity level and when $q \in \{0, 1\}$ – corresponds to Jost's beta; and a measure of gamma diversity. Each of these can be measured at the metacommunity level, and broken down into subcommunity and species-level components. One reason for the multiplicity of measures relates to the option to control for (or not control for) the potentially differing sizes of the subcommunities in half of the alpha and beta diversity measures (described as normalised and raw diversities, respectively, below).

2.2.2 General notation

Consider a *metacommunity* comprising S species of relative abundance $\boldsymbol{p} = (p_1, \ldots, p_S)$, partitioned into N distinct *subcommunities* (*e.g.* Figure 2.1, middle image). Then $\mathbf{P}$ is an $S \times N$ matrix, where element P_{ij} denotes the relative abundance of species i in subcommunity j. We can highlight a particular subcommunity, $\boldsymbol{P}_{.j} = (P_{1j}, \ldots, P_{Sj})$, where each element describes the relative abundance of each species in subcommunity j. We assume that the metacommunity is complete, so $\sum_i \sum_j P_{ij} = 1$, we can say that $\sum_i P_{ij} = w_j$ describes the weight of subcommunity j relative to the metacommunity as a whole and $\sum_j w_j = 1$. In other words, w_j is the proportion of the metacommunity that is contained within subcommunity j. Also, $\bar{\boldsymbol{P}}_{.j} = \boldsymbol{P}_{.j}/w_j$ is the normalised relative abundance of species in subcommunity j in isolation. Collapsing $\mathbf{P}$ across all subcommunities yields $\boldsymbol{p} = \sum_j \boldsymbol{P}_{.j}$, the relative abundance of species in the metacommunity as a whole. The matrix $\mathbf{Z}$ is an $S \times S$ matrix, where element $Z_{ii'}$ defines the pairwise similarity between species i and i', where $Z_{ii'} = 1$ denotes complete similarity and $Z_{ii'} = 0$ when species are completely distinct (as in Section 1.5.2). Therefore, $(\mathbf{Z}\boldsymbol{p})_i$ is the '*relative abundance of species similar to the* i^{th} ', termed the *ordinariness* of species i (Leinster & Cobbold, 2012). We then define $(\mathbf{Z}\boldsymbol{p})_i$ as metacommunity ordinariness of species i, $(\mathbf{Z}\bar{\boldsymbol{P}}_{.j})_i$ as the ordinariness of species i in subcommunity j, and $(\mathbf{Z}\bar{\boldsymbol{P}}_{ij})_i/(\mathbf{Z}\boldsymbol{p})_i$ as relative ordinariness of species i in subcommunity j relative to the metacommunity.

Reeve et al. (2016) also introduce new terminology to represent the simplest or most extreme cases of community diversity. The first is the *naïve-type* case, wherein different species have zero similarity ($\mathbf{Z} = \mathbf{I}$). The second is the *naïve-community* case, wherein there are no shared species between subcommunities and species from different subcommunities have zero similarity (although similarities within a subcommunity may be non-zero). Finally, new notation is defined for each of the diversity measures: Metacommunity diversities are written in romanised upper-case (*e.g.* ${}^{q}T^{\boldsymbol{Z}}$); subcommunity diversities are written in Greek lower-case (*e.g.* ${}^{q}\tau_j^{\boldsymbol{Z}}$); and species-level components are written in romanised lower-case (*e.g.* $t_{ij}^{\boldsymbol{Z}}$). Metacommunity level diversities are calculated as weighted power means (described in Key Concepts) of their subcommunity-level components, which are in turn calculated as weighted power means of their species-level subcomponents.

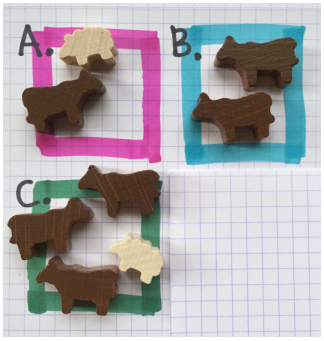
2.2.3 Methods

The example used throughout this chapter is a simple metacommunity, comprising three distinct subcommunities with different species distributions (shown in Figure 2.1). As a naïve observer, it is clear that each subcommunity contains a different distribution: *subcommunity A* (pink) contains an even distribution of species, *subcommunity B* (blue) contains only one species, and *subcommunity C* (green) has the same distribution of species as the metacommunity as a whole, but only half the total abundance of each species.

The aim here is to identify which of Reeve et al.'s (2016) measures are best able to pick out the patterns that visually distinguish these subcommunities. For simplicity, diversity is calculated in the naïve-type case, such that each species is considered completely distinct ($\mathbf{Z} \equiv \mathbf{I}$). I will also show the explicit mathematical calculation at $q = 0$, whilst nevertheless plotting the diversities for all values of q. In each case, I give a general overview of the measure – where text is taken and adapted from Reeve et al. (2016). After which, I examine the variation in species composition within subcommunities, across subcommunities, and at the metacommunity level. Species-level components are used to inform the way in which subcommunity-level measures are constructed, but are are not discussed in any great detail, beyond a simple mathematical description. I start with the alpha and gamma diversities, since they are closely related and easier to understand, before moving on to the beta diversities.

$$\mathbf{ZP} = \mathbf{P} = \begin{array}{c} \begin{matrix} A & B & C \end{matrix} \\ \begin{bmatrix} 1/8 & 2/8 & 3/8 \\ 1/8 & 0 & 1/8 \end{bmatrix} \end{array} \begin{matrix} \\ cows \\ sheep \end{matrix}$$

$$\bar{\mathbf{P}} = \begin{array}{c} \begin{matrix} A & B & C \end{matrix} \\ \begin{bmatrix} 1/2 & 1/2 & 3/2 \\ 1/2 & 0 & 1/4 \end{bmatrix} \end{array} \begin{matrix} \\ cows \\ sheep \end{matrix}$$



$$\boldsymbol{p} = \begin{array}{c} \begin{matrix} cows & sheep \end{matrix} \\ \begin{pmatrix} 6/8, & 2/8 \end{pmatrix} \end{array}$$

$$\boldsymbol{w} = \begin{array}{c} \begin{matrix} A & B & C \end{matrix} \\ \begin{pmatrix} 1/4, & 1/4, & 1/2 \end{pmatrix} \end{array}$$

$$\mathbf{Z} = \begin{array}{c} \begin{matrix} cows & sheep \end{matrix} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \begin{matrix} \\ cows \\ sheep \end{matrix}$$

Figure 2.1: Illustrative example of a simple metacommunity: comprising subcommunities A (pink), B (blue), and C (green). This metacommunity is represented numerically by a matrix of relative abundances, $\mathbf{P} = [P_{ij}]$; a matrix of normalised relative abundances, $\bar{\mathbf{P}} = [P_{ij}]$; a vector of relative abundances collapsed across subcommunities, $\boldsymbol{p} = (p_1, \ldots, p_S)$; a vector of subcommunity weights, $\boldsymbol{w} = (w_1, \ldots, w_N)$; a matrix of pairwise species similarities, $\mathbf{Z} = [Z_{ii'}]$; and a matrix of the ordinariness of species, $\mathbf{ZP} = [(ZP)_{ij}]$ (in the metacommunity).

2.2.4 Alpha diversities ($^{q}\bar{\alpha}_{j}^{\mathbf{Z}}$, $^{q}\alpha_{j}^{\mathbf{Z}}$, $^{q}\bar{A}^{\mathbf{Z}}$, and $^{q}A^{\mathbf{Z}}$)

Two kinds of alpha diversity are included in this framework. These are normalised and raw alpha diversities. All of the alpha diversity measures are listed in Table 2.1.

Table 2.1: Alpha diversities: Unified mathematical framework describing measures of similarity-sensitive subcommunity- and metacommnuity-level diversities, alongside their species-level components. Each subcommunity-level measure is calculated as a weighted power mean (described in Key Concepts) of the species-level components and is quantified relative to the metacommunity as a whole. Each metacommunity-level measure is calculated as a weighted power mean of the subcommunity-level values.

Formula	Description
Species-level	
$a_{i,j}^{\mathbf{Z}} = \left(\mathbf{Z}\boldsymbol{P}_{.j}\right)_{i}^{-1}$	**Raw alpha**: species-level component of $^{q}\alpha_{j}^{\mathbf{Z}}$
$\bar{a}_{i,j}^{\mathbf{Z}} = \left(\mathbf{Z}\bar{\boldsymbol{P}}_{.j}\right)_{i}^{-1}$	**Normalised alpha**: species-level component of $^{q}\bar{\alpha}_{j}^{\mathbf{Z}}$
Subcommunity-level	
$^{q}\alpha_{j}^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}_{.j}, \left(\mathbf{Z}\boldsymbol{P}_{.j}\right)_{i}^{-1}\right)$	**Raw alpha**: estimate of naïve-community metacommunity diversity
$^{q}\bar{\alpha}_{j}^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}_{.j}, \left(\mathbf{Z}\bar{\boldsymbol{P}}_{.j}\right)_{i}^{-1}\right)$	**Normalised alpha**: similarity-sensitive diversity of subcommunity j in isolation
Metacommunity-level	
$^{q}A^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{q}\alpha_{j}^{\mathbf{Z}}\right)$	**Raw alpha**: naïve-community metacommunity diversity
$^{q}\bar{A}^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{q}\bar{\alpha}_{j}^{\mathbf{Z}}\right)$	**Normalised alpha**: average similarity-sensitive diversity of subcommunities

2.2.4.1 Normalised alpha diversity ($^{q}\bar{\alpha}_{j}^{\mathbf{Z}}$ and $^{q}\bar{A}^{\mathbf{Z}}$)

Normalised subcommunity alpha diversity is *the similarity-sensitive diversity of subcommunity j in isolation* ($^{q}\bar{\alpha}_{j}^{\mathbf{Z}}$, Table 2.1), and is equivalent to Leinster & Cobbold's (2012) similarity-sensitive diversity of a single community (Equation 1.11). As with all of the subcommunity measures that follow, it does not depend on how the rest of the metacommunity is partitioned into subcommunities. Normalised metacommunity alpha diversity is *the average*[1] *similarity-sensitive diversity of subcommunities in the metacommunity* ($^{q}\bar{A}^{\mathbf{Z}}$, Table 2.1) and is simply the average of the normalised subcommunity alpha diversities. So both subcommunity and metacommunity alpha are effective numbers and take values between 1 and S, where S is the total number of species.

[1]Recalling that throughout this chapter I refer to weighted power mean measures informally as averages, for simplicity

Example

Figure 2.2a shows the metacommunity broken down into subcommunities, alongside the species-level (Figure 2.2b), subcommunity-level (Figure 2.2c), and metacommunity-level (Figure 2.2d) measures of diversity. Figures 2.2b-d can be read across to see how the lower level components combine to form their higher level averages. This format is repeated across all of the examples in this chapter.

The species-level components of the effective number of species in subcommunity A are calculated as the inverse of the ordinariness (or "specialness") of the species in that subcommunity (Figure 2.2b)[1]. In the naïve-type case, this is simply the inverse of the relative abundance of each species in the subcommunity:

$$\begin{aligned}
\bar{a}^{\mathbf{I}}_{cows,A} &= \left(\mathbf{I}\bar{\boldsymbol{P}}_{\cdot A}\right)^{-1}_{cows} = \left(\bar{\boldsymbol{P}}_{\cdot A}\right)^{-1}_{cows} = \bar{P}^{-1}_{cows,A} = \left(\frac{P_{cows,A}}{w_A}\right)^{-1} \\
&= \left(\frac{1/8}{1/4}\right)^{-1} = \left(\frac{1}{2}\right)^{-1} = 2 \\
\bar{a}^{\mathbf{I}}_{sheep,A} &= \left(\mathbf{I}\bar{\boldsymbol{P}}_{\cdot A}\right)^{-1}_{sheep} = \left(\bar{\boldsymbol{P}}_{\cdot A}\right)^{-1}_{sheep} = \bar{P}^{-1}_{sheep,A} = \left(\frac{P_{sheep,A}}{w_A}\right)^{-1} \\
&= \left(\frac{1/8}{1/4}\right)^{-1} = \left(\frac{1}{2}\right)^{-1} = 2
\end{aligned}$$

which tells us that the cows and sheep in subcommunity A each have a specialness of 2, which is a component of diversity.

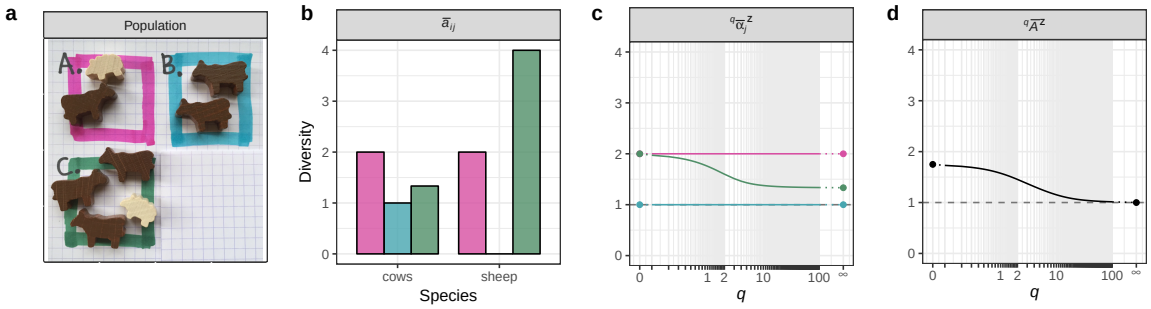


Figure 2.2: Partitioning normalised alpha diversity at hierarchical levels of community structure: (a) metacommunity composition, and alpha diversity calculated at the (b) species, (c) subcommunity, and (d) metacommunity levels. As in (a), subcommunities A, B, and C are coloured pink, blue, and green, respectively. The dashed line highlights the minimum level of diversity.

[1]Following Leinster & Cobbold (2012), the average ordinariness ($\sum_{i=1}^{S} p_i(\mathbf{Z}\boldsymbol{p})_i$) is a measure of concentration. That is, when the average ordinariness of a subcommunity is high, the subcommunity is concentrated into a small number of similar species. Since $\sum_{i=1}^{S} p_i(\mathbf{Z}\boldsymbol{p})_i$ is simply the similarity-sensitive equivalent of Simpson's index of concentration, then the diversity of a subcommunity (at $q = 2$) can be calculated as the inverse of the weighted average ordinariness. This can be extended to all values of q.

The arithmetic mean of $\bar{a}^{\mathbf{I}}_{cows,A}$ and $\bar{a}^{\mathbf{I}}_{sheep,A}$ gives the effective number of species in subcommunity A at $q = 0$, which is the species richness[1]:

$$\begin{aligned}
{}^{0}\bar{\alpha}^{\mathbf{I}}_{A} &= \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}, \bar{a}^{\mathbf{I}}_{i,A}\right) \\
&= \left(\bar{P}_{cows,A} \times \bar{a}^{\mathbf{I}}_{cows,A}\right) + \left(\bar{P}_{sheep,A} \times \bar{a}^{\mathbf{I}}_{sheep,A}\right) \\
&= \left(\tfrac{1}{2} \times 2\right) + \left(\tfrac{1}{2} \times 2\right) = 2
\end{aligned}$$

For subcommunity A, ${}^{q}\bar{\alpha}^{\mathbf{I}}_{A}$ remains constant for all values of q, since species are evenly distributed (Figure 2.2c). Likewise, subcommunity B has an effective number of species equal to 1 for all values of q, so all power means are the same, though here this is trivial as it comprises only a single species. In subcommunity C, however, as q increases – and ${}^{q}\bar{\alpha}^{\mathbf{I}}_{j}$ becomes more conservative, in the sense that calculated diversity drops (in this case towards the specialness of the least special species in the subcommunity) – the effective number of species decreases until at $q = \infty$ only the least special species (in this case cows) is considered, *i.e.* ${}^{\infty}\bar{\alpha}^{\mathbf{I}}_{C} = \bar{a}^{\mathbf{I}}_{cows,C} = \left(\frac{3/8}{1/2}\right)^{-1} = 1\frac{1}{3}$.

Metacommunity values are calculated as a power mean of the subcommunity values. So at $q = 0$, the arithmetic mean of species richness across all subcommunities is calculated as:

$$\begin{aligned}
{}^{0}\bar{A}^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{0}\bar{\alpha}^{\mathbf{I}}_{A}\right) \\
&= \left(p_A \times {}^{0}\bar{\alpha}^{\mathbf{I}}_{A}\right) + \left(p_B \times {}^{0}\bar{\alpha}^{\mathbf{I}}_{B}\right) + \left(p_C \times {}^{0}\bar{\alpha}^{\mathbf{I}}_{C}\right) \\
&= \left(\tfrac{1}{4} \times 2\right) + \left(\tfrac{1}{4} \times 1\right) + \left(\tfrac{1}{2} \times 2\right) = 1\tfrac{3}{4}
\end{aligned}$$

which tells us that the subcommunities in the metacommunity have on average $1\frac{3}{4}$ species. This is simply Whittaker's *alpha diversity* of the metacommunity. Similarly, ${}^{\infty}\bar{A}^{\mathbf{I}}$ is calculated as the minimum of the set $\left\{{}^{\infty}\bar{\alpha}^{\mathbf{I}}_{A}, {}^{\infty}\bar{\alpha}^{\mathbf{I}}_{B}, {}^{\infty}\bar{\alpha}^{\mathbf{I}}_{C}\right\}$ which is the effective number of species (just the cows!) in the least diverse subcommunity B, so ${}^{\infty}\bar{A}^{\mathbf{I}} = {}^{\infty}\bar{\alpha}^{\mathbf{I}}_{B} = \bar{a}^{\mathbf{I}}_{cows,B} = 1$ (Figure 2.2d).

2.2.4.2 Raw alpha (${}^{q}\alpha^{\mathbf{Z}}_{j}$ and ${}^{q}A^{\mathbf{Z}}$)

The raw subcommunity alpha diversity is *a subcommunity-level estimate of naïve-community metacommunity diversity* (${}^{q}\alpha^{\mathbf{Z}}_{j}$, Table 2.1), which is related to the normalised subcommunity alpha diversity via a rescaling by the size of the subcommunity (${}^{q}\bar{\alpha}^{\mathbf{Z}}_{j}/w_j = {}^{q}\alpha^{\mathbf{Z}}_{j}$). In the absence of other information, ${}^{q}\alpha^{\mathbf{Z}}_{j}$ can be used to estimate the metacommunity (gamma) diversity, ${}^{q}G^{\mathbf{Z}}$. This estimate will be exact in the naïve-community case if every other subcommunity takes the same value of ${}^{q}\alpha^{\mathbf{Z}}_{j}$, but in general will tend to be an overestimate due to commonalities between species in different subcommunities. Raw metacommunity alpha diversity is *the*

[1]For other values of q, other means are used – a weighted geometric mean at $q = 1$, a harmonic mean at $q = 2$, and so on.

average of the raw subcommunity alpha diversities (${}^{q}A^{\mathbf{Z}}$, Table 2.1), and is an upper bound on the metacommunity gamma diversity ${}^{q}G^{\mathbf{Z}}$. This allows ${}^{q}G^{\mathbf{Z}}$ to be constrained without any knowledge of the relationships between species across the subcommunities.

Example

The species-level components of ${}^{q}\alpha_{A}^{\mathbf{I}}$ are calculated as the scaled specialness of the species in subcommunity A, which in the naïve-type case is the relative abundance of each species in the subcommunity as a fraction of the metacommunity:

$$a_{cows,A}^{\mathbf{I}} = (\mathbf{I}\boldsymbol{P}_{\cdot A})_{cows}^{-1} = P_{cows,A}^{-1} = \left(\tfrac{1}{8}\right)^{-1} = 8$$
$$a_{sheep,A}^{\mathbf{I}} = (\mathbf{I}\boldsymbol{P}_{\cdot A})_{sheep}^{-1} = P_{sheep,A}^{-1} = \left(\tfrac{1}{8}\right)^{-1} = 8$$

At $q = 0$, subcommunity ${}^{q}\alpha_{A}^{\mathbf{I}}$ is calculated as the weighted mean of $a_{cows,A}^{\mathbf{I}}$ and $a_{sheep,A}^{\mathbf{I}}$:

$$\begin{aligned} {}^{0}\alpha_{A}^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}, a_{i,A}^{\mathbf{I}}\right) \\ &= \left(\bar{P}_{cows,A} \times a_{cows,A}^{\mathbf{I}}\right) + \left(\bar{P}_{sheep,A} \times a_{sheep,A}^{\mathbf{I}}\right) \\ &= \left(\tfrac{1}{2} \times 8\right) + \left(\tfrac{1}{2} \times 8\right) = 8 \end{aligned}$$

which remains constant as q increases, since the species in subcommunity A are evenly distributed, and therefore $a_{cows,A}^{\mathbf{I}}$ and $a_{sheep,A}^{\mathbf{I}}$ are equal in value (see Figure 2.3b-c), as for ${}^{q}\bar{\alpha}_{j}$. Raw subcommunity alpha is a novel measure of diversity, whose value is high when a subcommunity contains an even distribution of many species and when the subcommunity is small relative to the metacommunity. At $q = 0$, ${}^{0}\alpha_{B}^{\mathbf{I}}$ is lower than ${}^{0}\alpha_{A}^{\mathbf{I}}$ because subcommunity B is missing sheep, though both subcommunities are the same weight. Conversely, ${}^{0}\alpha_{C}^{\mathbf{I}}$ is lower than ${}^{0}\alpha_{A}^{\mathbf{I}}$ because subcommunity C is larger in size than A, though it contains the same number of species. At $q = \infty$, subcommunity C considers only the most conservative estimate of ${}^{q}\alpha_{C}^{\mathbf{I}}$, that of the highly abundant cows (*i.e.* ${}^{\infty}\alpha_{C}^{\mathbf{I}} = a_{cows,C}^{\mathbf{I}} = \left(\tfrac{3}{8}\right)^{-1} = 2\tfrac{2}{3}$).

The arithmetic mean of $\{{}^{0}\alpha_{A}^{\mathbf{I}}, {}^{0}\alpha_{B}^{\mathbf{I}}, {}^{0}\alpha_{C}^{\mathbf{I}}\}$ gives ${}^{0}A^{\mathbf{I}}$, the naïve-community metacommunity

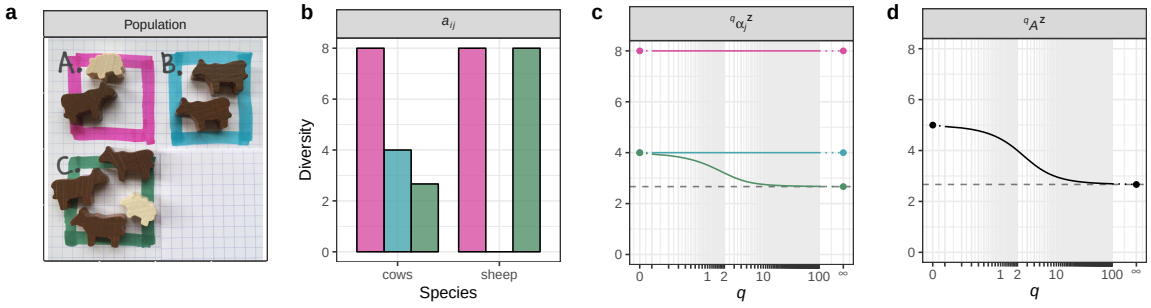


Figure 2.3: Partitioning raw alpha diversity at hierarchical levels of community structure: (a) metacommunity composition, and raw alpha diversity calculated at the (b) species, (c) subcommunity, and (d) metacommunity levels. As in (a), subcommunities A, B, and C are coloured pink, blue, and green, respectively. The dashed line highlights the minimum level of diversity.

diversity at $q = 0$:

$$\begin{aligned} {}^{0}A^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{0}\alpha_{A}^{\mathbf{I}}\right) \\ &= \left(p_A \times {}^{0}\alpha_{A}^{\mathbf{I}}\right) + \left(p_B \times {}^{0}\alpha_{B}^{\mathbf{I}}\right) + \left(p_C \times {}^{0}\alpha_{C}^{\mathbf{I}}\right) \\ &= \left(\tfrac{1}{4} \times 8\right) + \left(\tfrac{1}{4} \times 4\right) + \left(\tfrac{1}{2} \times 4\right) = 5 \end{aligned}$$

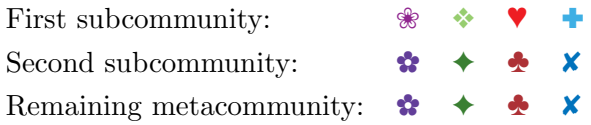
which tells us that metacommunity ${}^{0}G^{\mathbf{I}}$ is less than 5. This value decreases as q increases until the most conservative value is reached at $q = \infty$, such that ${}^{\infty}A^{\mathbf{I}} = {}^{\infty}\alpha_{C}^{\mathbf{I}} = a_{cows,C}^{\mathbf{I}} = 2\frac{2}{3}$ (Figure 2.3d).

2.2.5 Gamma diversity (${}^{q}\gamma_{j}^{\mathbf{Z}}$ and ${}^{q}G^{\mathbf{Z}}$)

Metacommunity gamma diversity is *the similarity-sensitive diversity of the unpartitioned metacommunity* (${}^{q}G^{\mathbf{Z}}$, Table 2.2). This is calculated as an average of the new 'subcommunity gamma diversities', which measure *the contribution per individual in the subcommunity to the diversity of the metacommunity as a whole* (${}^{q}\gamma_{j}^{\mathbf{Z}}$, Table 2.2). It is a new kind of diversity measure, which is able to identify new patterns that are harder to observe using traditional alpha and beta diversity measures.

Since ${}^{q}\gamma_{j}^{\mathbf{Z}}$ measures the contribution to diversity per individual, an increase in ${}^{q}\bar{\alpha}_{j}^{\mathbf{Z}}$ does not necessarily cause a change in ${}^{q}\gamma_{j}^{\mathbf{Z}}$. Consider, for example, a naïve-type case with all species equally abundant in the metacommunity. If two subcommunities have evenly distributed species, but the first subcommunity contains k times as many species as the second, then $\bar{\alpha}_1 = k\bar{\alpha}_2$; but $\gamma_1 = \gamma_2$, because all the individuals in both subcommunities are members of equally 'valuable' species and so contribute equally to metacommunity diversity, being of species that are equally abundant in the metacommunity. For example:

then $\bar{\alpha}_1 = 8$ and $\bar{\alpha}_2 = 4$, but $\gamma_1 = \gamma_2 = 8$. On the other hand, if two equally sized subcommunities have different constituent species, but with the same relative abundances, in such a way that all of the species in the first subcommunity are k times rarer in the metacommunity than the species in second, then $\bar{\alpha}_1 = \bar{\alpha}_2$ but $\gamma_1 = k\gamma_2$. This reflects the higher contribution of subcommunity A to metacommunity diversity, through its rarer and therefore more special species. For example:



then $\bar{\alpha}_1 = \bar{\alpha}_2 = 4$, but $\gamma_1 = 12$ and $\gamma_2 = 6$.

Table 2.2: Gamma diversities: Unified mathematical framework describing measures of similarity-sensitive subcommunity- and metacommnuity-level diversities, alongside their species-level components. Each subcommunity-level measure is calculated as a weighted power mean (described in Key Concepts) of the species-level components and is quantified relative to the metacommunity as a whole. Each metacommunity-level measure is calculated as a weighted power mean of the subcommunity-level values.

Formula	Description
Species-level	
$g_{i,j}^{\mathbf{Z}} = \mathbf{Z}\boldsymbol{p}_{.j}^{-1}$	**Gamma**: species-level component of ${}^{q}\gamma_{j}^{\mathbf{Z}}$
Subcommunity-level	
${}^{q}\gamma_{j}^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}_{.j}, (\mathbf{Z}\boldsymbol{p})^{-1}\right)$	**Gamma**: contribution per individual toward metacommunity diversity
Metacommunity-level	
${}^{q}G_{j}^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{q}\gamma_{j}^{\mathbf{Z}}\right)$	**Gamma**: metacommunity similarity-sensitive diversity

Example

The species-level components of the contribution per individual toward metacommunity diversity are calculated (in the naïve-type case) as the inverse of the metacommunity-abundance of each species in subcommunity j, which for subcommunity A is written:

$$g_{cows,A}^{\mathbf{I}} = (\mathbf{I}\boldsymbol{p})_{cows}^{-1} = p_{cows}^{-1} = \frac{1}{6/8} = 1\tfrac{1}{3}$$

$$g_{sheep,A}^{\mathbf{I}} = (\mathbf{I}\boldsymbol{p})_{sheep}^{-1} = p_{sheep}^{-1} = \frac{1}{2/8} = 4$$

The arithmetic mean of $g_{cows,A}^{\mathbf{I}}$ and $g_{sheep,A}^{\mathbf{I}}$ yields the contribution per individual toward metacommunity diversity,

$$\begin{aligned} {}^{0}\gamma_{A}^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}, g_{i,A}^{\mathbf{I}}\right) \\ &= \left(\bar{P}_{cows,A} \times g_{cows,A}^{\mathbf{I}}\right) + \left(\bar{P}_{sheep,A} \times g_{sheep,A}^{\mathbf{I}}\right) \\ &= \left(\tfrac{1}{2} \times 1\tfrac{1}{3}\right) + \left(\tfrac{1}{2} \times 4\right) = 2\tfrac{2}{3} \end{aligned}$$

which is the highest of all the subcommunity values because subcommunity A contains a higher proportion of rare species (in this case sheep). As q increases, ${}^{q}\gamma_{A}^{\mathbf{I}}$ decreases, until at $q = \infty$, only the most conservative estimate of the contribution per individual toward metacommunity diversity is considered, *i.e.* ${}^{\infty}\gamma_{A}^{\mathbf{I}} = g_{cows,A}^{\mathbf{I}} = 1\frac{1}{3}$.

Naïve-type metacommunity ${}^{q}G^{\mathbf{Z}}$ at $q = 0$ is equivalent to Whittaker's total diversity (*gamma diversity*), calculated as the arithmetic mean across all subcommunities:

$$\begin{aligned} {}^{0}G^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{0}\gamma_{A}^{\mathbf{I}}\right) \\ &= \left(p_A \times {}^{0}\gamma_{A}^{\mathbf{I}}\right) + \left(p_B \times {}^{0}\gamma_{B}^{\mathbf{I}}\right) + \left(p_C \times {}^{0}\gamma_{C}^{\mathbf{I}}\right) \\ &= \left(\tfrac{1}{4} \times 2\tfrac{2}{3}\right) + \left(\tfrac{1}{4} \times 1\tfrac{1}{3}\right) + \left(\tfrac{1}{2} \times 2\right) = 2 \end{aligned}$$

which is the number of species present in the metacommunity, as expected. This value decreases with increasing q, until ${}^{\infty}G^{\mathbf{I}} = {}^{\infty}\gamma^{\mathbf{I}}_{\{A,B,C\}} = g^{\mathbf{I}}_{cows,\{A,B,C\}}$.

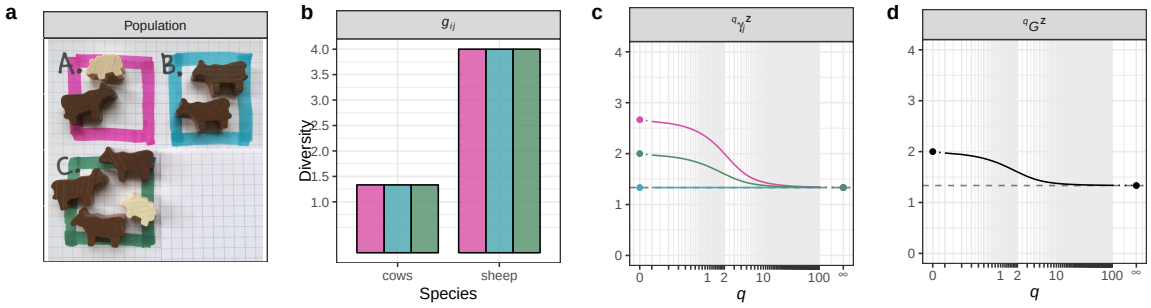


Figure 2.4: Partitioning gamma diversity at hierarchical levels of community structure: (a) metacommunity composition, and gamma diversity calculated at the (b) species, (c) subcommunity, and (d) metacommunity levels. As in (a), subcommunities A, B, and C are coloured pink, blue, and green, respectively. The dashed line highlights the minimum level of diversity.

2.2.6 Beta diversities (${}^{q}\bar{\rho}^{\mathbf{Z}}_{j}$, ${}^{q}\rho^{\mathbf{Z}}_{j}$, ${}^{q}\bar{\beta}^{\mathbf{Z}}_{j}$, ${}^{q}\beta^{\mathbf{Z}}_{j}$, ${}^{q}\bar{R}^{\mathbf{Z}}$, ${}^{q}R^{\mathbf{Z}}$, ${}^{q}\bar{B}^{\mathbf{Z}}$, and ${}^{q}B^{\mathbf{Z}}$)

Four kinds of beta diversity naturally emerge from this framework, reflecting different aspects of the relationship between the metacommunity and its constituent subcommunities. These are representativeness, redundancy, the effective number of subcommunities, and distinctiveness. Representativeness and the effective number of subcommunities are normalised measures, whilst redundancy and distinctiveness are raw. All beta diversity measures are listed in Table 2.3.

2.2.6.1 Representativeness (${}^{q}\bar{\rho}^{\mathbf{Z}}_{j}$ and ${}^{q}\bar{R}^{\mathbf{Z}}$)

The subcommunity ${}^{q}\bar{\rho}^{\mathbf{Z}}_{j}$ diversity measures *how representative, or typical, the subcommunity is of the metacommunity* (Table 2.3). In the naïve-type case, the maximum value of ${}^{q}\bar{\rho}^{\mathbf{I}}_{j}$ is 1, which is attained when all the species in the metacommunity are present in the subcommunity ($q = 0$), or more consequentially at higher q, when the distribution of species in the subcommunity is identical to that of the metacommunity, meaning that the subcommunity represents the metacommunity faithfully. Subcommunity representativeness, ${}^{q}\bar{\rho}^{\mathbf{I}}_{j}$, is minimised (with value w_j) when the species present in subcommunity j are not present anywhere else in the metacommunity: that is, if we lost the subcommunity then no individual in the remaining metacommunity would have any similarity to what was lost. Consequently, a subcommunity is also generally more representative when it constitutes more of the metacommunity. Metacommunity ${}^{q}\bar{R}^{\mathbf{Z}}$ is simply *the average representativeness of subcommunities in the metacommunity* (Table 2.3), which is calculated as the weighted power mean of the subcommunity-level measures.

Table 2.3: Beta diversities: Unified mathematical framework describing measures of similarity-sensitive subcommunity- and metacommnuity-level diversities, alongside their species-level components. Each subcommunity-level measure is calculated as a weighted power mean (described in Key Concepts) of the species-level components and is quantified relative to the metacommunity as a whole. Each metacommunity-level measure is calculated as a weighted power mean of the subcommunity-level values.

Formula	Description
Species-level	
$r^{\mathbf{Z}}_{i,j} = {}^{(\mathbf{Z}\boldsymbol{p})_i}/_{\left(\mathbf{Z}\boldsymbol{P}_{\cdot j}\right)_i}$	**Raw rho**: species-level component of ${}^q\rho^{\mathbf{Z}}_j$
$\bar{r}^{\mathbf{Z}}_{i,j} = {}^{(\mathbf{Z}\boldsymbol{p})_i}/_{\left(\mathbf{Z}\bar{\boldsymbol{P}}_{\cdot j}\right)_i}$	**Normalised rho**: species-level component of ${}^q\bar{\rho}^{\mathbf{Z}}_j$
$b^{\mathbf{Z}}_{i,j} = {}^{\left(\mathbf{Z}\boldsymbol{P}_{\cdot j}\right)_i}/_{(\mathbf{Z}\boldsymbol{p})_i}$	**Raw beta**: species-level component of ${}^q\beta^{\mathbf{Z}}_j$
$\bar{b}^{\mathbf{Z}}_{i,j} = {}^{\left(\mathbf{Z}\bar{\boldsymbol{P}}_{\cdot j}\right)_i}/_{(\mathbf{Z}\boldsymbol{p})_i}$	**Normalised beta**: species-level component of ${}^q\bar{\beta}^{\mathbf{Z}}_j$
Subcommunity-level	
${}^q\rho^{\mathbf{Z}}_j = \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}_{\cdot j}, {}^{(\mathbf{Z}\boldsymbol{p})_i}/_{\left(\mathbf{Z}\boldsymbol{P}_{\cdot j}\right)_i}\right)$	**Raw rho**: redundancy of subcommunity j
${}^q\bar{\rho}^{\mathbf{Z}}_j = \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}_{\cdot j}, {}^{(\mathbf{Z}\boldsymbol{p})_i}/_{\left(\mathbf{Z}\bar{\boldsymbol{P}}_{\cdot j}\right)_i}\right)$	**Normalised rho**: representativeness of subcommunity j
${}^q\beta^{\mathbf{Z}}_j = {}^1/_{{}^q\rho^{\mathbf{Z}}_j}$	**Raw beta**: distinctiveness of subcommunity j
${}^q\bar{\beta}^{\mathbf{Z}}_j = {}^1/_{{}^q\bar{\rho}^{\mathbf{Z}}_j}$	**Normalised beta**: estimate of effective number of distinct subcommunities
Metacommunity-level	
${}^qR^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^q\rho^{\mathbf{Z}}_j\right)$	**Raw rho**: average redundancy of subcommunities
${}^q\bar{R}^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^q\bar{\rho}^{\mathbf{Z}}_j\right)$	**Normalised rho**: average representativeness of subcommunities
${}^qB^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^q\beta^{\mathbf{Z}}_j\right)$	**Raw beta**: average distinctiveness of subcommunities
${}^q\bar{B}^{\mathbf{Z}} = \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^q\bar{\beta}^{\mathbf{Z}}_j\right)$	**Normalised beta**: effective number of distinct subcommunities

For general $\mathbf{Z}$, a low value of ${}^q\bar{\rho}^{\mathbf{Z}}_j$ indicates a subcommunity that has little in common with the metacommunity as a whole, and is in this sense not very representative of the metacommunity. It is often the case that similarities between individuals in the same subcommunity are, on average, greater than similarities between individuals in different subcommunities, and in this scenario ${}^q\bar{\rho}^{\mathbf{Z}}_j$ is less than 1. As an example, consider a naïve-type case with all species equally abundant in each subcommunity in which they are present, and all species equally abundant in the metacommunity as a whole. If every subcommunity contains only a proportion, r, of the total number of species then ${}^q\bar{\rho}^{\mathbf{Z}}_j = {}^q\bar{R}^{\mathbf{Z}} = r$, reflecting the fact that each subcommunity represents a fraction r of the total metacommunity. At $q = 0$, when $\mathbf{Z} = \mathbf{I}$,

$$
\begin{aligned}
{}^0\bar{\rho}^{\mathbf{I}}_j &= \mathrm{M}_1\left(\bar{\boldsymbol{P}}_{\cdot j}, \frac{\boldsymbol{p}}{\bar{\boldsymbol{P}}_{\cdot j}}\right) \\
&= \sum_{i:\bar{P}_{ij}>0} \bar{P}_{ij} \times \frac{p_i}{\bar{P}_{ij}} = \sum_{i:\bar{P}_{ij}>0} p_i
\end{aligned}
$$

which is the relative abundance in the metacommunity of all of the species that are present in subcommunity j. So representativeness in this special case is just the proportion of individuals in the metacommunity that have any representation in the subcommunity.

Example

Species-level estimates of the representativeness of subcommunity A are calculated for each species as the inverse of the relative ordinariness of species i in subcommunity j relative to the metacommunity. For subcommunity A, these are calculated as:

$$\bar{r}^{\mathbf{I}}_{cows,A} = \frac{(\mathbf{I}\boldsymbol{p})_{cows}}{\left(\mathbf{I}\bar{\boldsymbol{P}}_{\cdot A}\right)_{cows}} = \frac{p_{cows}}{\bar{P}_{cows,A}} = \frac{6/8}{1/2} = 1\tfrac{1}{2}$$

$$\bar{r}^{\mathbf{I}}_{sheep,A} = \frac{(\mathbf{I}\boldsymbol{p})_{sheep}}{\left(\mathbf{I}\bar{\boldsymbol{P}}_{\cdot A}\right)_{sheep}} = \frac{p_{sheep}}{\bar{P}_{sheep,A}} = \frac{2/8}{1/2} = \frac{1}{2}$$

At the subcommunity level, the arithmetic mean of $\bar{r}^{\mathbf{I}}_{cows,A}$ and $\bar{r}^{\mathbf{I}}_{sheep,A}$ gives the representativeness of subcommunity A at $q = 0$:

$$\begin{aligned} {}^{0}\bar{\rho}^{\mathbf{I}}_{A} &= \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}, \bar{r}^{\mathbf{I}}_{i,A}\right) \\ &= \left(\bar{P}_{cows,A} \times \bar{r}_{cows,A}\right) + \left(\bar{P}_{sheep,A} \times \bar{r}_{sheep,A}\right) \\ &= \left(\tfrac{1}{2} \times 1\tfrac{1}{2}\right) + \left(\tfrac{1}{2} \times \tfrac{1}{2}\right) = 1 \end{aligned}$$

Since all species (and hence all individuals) have at least 1 species representative in subcommunities A and C, the representativeness of these subcommunities is 1 at $q = 0$ (Figure 2.5c). However, only the cows (which constitute $\frac{3}{4}$ of the individuals) reside in subcommunity B, and so its representativeness is $\frac{3}{4}$.

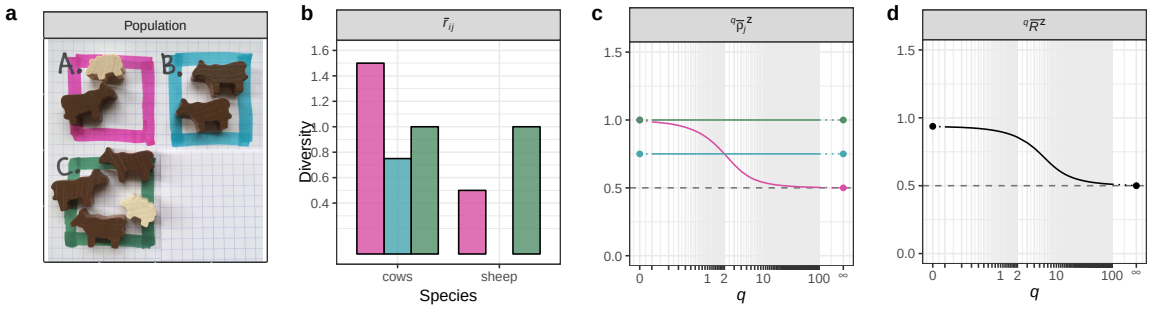


Figure 2.5: Partitioning representativeness at hierarchical levels of community structure: (a) metacommunity composition, and representativeness calculated at the (b) species, (c) subcommunity, and (d) metacommunity levels. As in (a), subcommunities A, B, and C are coloured pink, blue, and green, respectively. The dashed line highlights the minimum level of diversity.

As q increases, ${}^{q}\bar{\rho}^{\mathbf{I}}_{A}$ places more importance on species of low representativeness (species with a high subcommunity-ordinariness, but low ordinariness across the metacommunity as a whole), in this case sheep. Since the distribution of species within subcommunity C is identical to that of the metacommunity, the species in subcommunity C are equally representative of the metacommunity ($\bar{r}^{\mathbf{I}}_{cows,C} = \bar{r}^{\mathbf{I}}_{sheep,C}$) for all values of q. Likewise, subcommunity B is $\frac{3}{4}$ representative of the metacommunity for all values of q, since the representativeness of species in this subcommunity relates to the only species present (cows).

Subcommunity A on the other hand, has different values of representativeness for its two constituent species, and so at $q = \infty$, where only the lower representativeness (that of sheep) is considered, ${}^{\infty}\bar{\rho}^{\mathbf{I}}_{A}$ takes its most conservative value, *i.e.* ${}^{\infty}\bar{\rho}^{\mathbf{I}}_{A} = \bar{r}^{\mathbf{I}}_{sheep,A} = \frac{1}{2}$.

At $q = 0$, metacommunity ${}^{0}\bar{R}^{\mathbf{I}}$ is calculated as the arithmetic mean of ${}^{0}\bar{\rho}^{\mathbf{I}}_{A}$, ${}^{0}\bar{\rho}^{\mathbf{I}}_{B}$, and ${}^{0}\bar{\rho}^{\mathbf{I}}_{C}$ (Figure 2.5d),

$$\begin{aligned} {}^{0}\bar{R}^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{0}\bar{\rho}^{\mathbf{I}}_{A}\right) \\ &= \left(p_A \times {}^{0}\bar{\rho}^{\mathbf{I}}_{A}\right) + \left(p_B \times {}^{0}\bar{\rho}^{\mathbf{I}}_{B}\right) + \left(p_C \times {}^{0}\bar{\rho}^{\mathbf{I}}_{C}\right) \\ &= \left(\tfrac{1}{4} \times 1\right) + \left(\tfrac{1}{4} \times \tfrac{3}{4}\right) + \left(\tfrac{1}{2} \times 1\right) = \tfrac{15}{16} \end{aligned}$$

which tells us that on average, subcommunities are $\frac{15}{16}$ representative of the metacommunity. At $q = \infty$, metacommunity ${}^{q}\bar{R}^{\mathbf{I}}$ is equal to the most conservative species-level component of representativeness (the sheep in subcommunity A), such that ${}^{\infty}\bar{R}^{\mathbf{I}} = {}^{\infty}\bar{\rho}^{\mathbf{I}}_{A} = \bar{r}^{\mathbf{I}}_{sheep,A} = \frac{1}{2}$.

When a fourth subcommunity is added, comprising a new species (a single pig) that is unique in the metacommunity (Figure 2.6a), this subcommunity becomes the least representative subcommunity (Figure 2.6c) since it represents none of the species present elsewhere, and therefore ${}^{q}\bar{\rho}^{\mathbf{I}}_{D} = w_D = \frac{1}{9}$ for all q, and ${}^{q}\bar{R}^{\mathbf{I}}$ drops to ${}^{q}\bar{\rho}^{\mathbf{I}}_{D} = \frac{1}{9}$ as $q \to \infty$.

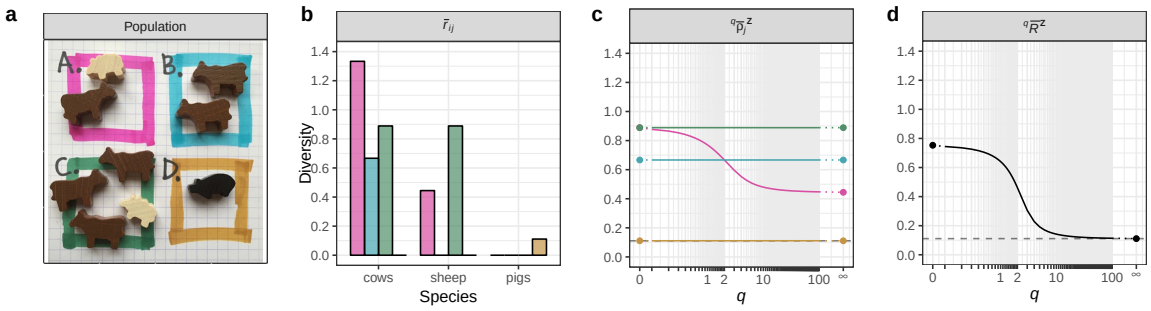


Figure 2.6: Partitioning representativeness at hierarchical levels of community structure, with an additional subcommunity: (a) metacommunity composition, and representativeness calculated at the (b) species, (c) subcommunity, and (d) metacommunity levels. As in (a), subcommunities A, B, C, and D are coloured pink, blue, green, and ochre respectively.

2.2.6.2 Redundancy (${}^{q}\rho_{j}^{\mathbf{Z}}$ and ${}^{q}R^{\mathbf{Z}}$)

Subcommunity ${}^{q}\rho_{j}^{\mathbf{Z}}$ diversity represents *the redundancy of the subcommunity within the metacommunity* (Table 2.3), calculated as the weighted power mean of the redundancy of each individual in the subcommunity. Like α and $\bar{\alpha}$, it is a rescaling of ${}^{q}\bar{\rho}_{j}^{\mathbf{Z}}$ by the size of the subcommunity, w_j, and therefore behaves similarly. In particular, it measures the extent to which the diversity of the metacommunity would be preserved if the subcommunity were to be destroyed. Redundancy is minimised (with value 1) when the species present in subcommunity j are not redundant in the metacommunity: that is, if we lost the subcommunity then no individual in the remaining metacommunity would have any similarity to what was lost.

Metacommunity ${}^{q}R^{\mathbf{Z}}$ is *the average redundancy of subcommunities in the metacommunity* (Table 2.3), which is calculated as the weighted power mean of the ${}^{q}\rho_{j}^{\mathbf{Z}}$s across subcommunities. This measure takes a minimum value of 1 in the naïve-community case (when there is no redundancy), and increases as the subcommunities become more alike in their composition – again, be that through shared species or increased similarity between species. When all N subcommunities in a metacommunity are identical in size and composition, then its redundancy, naturally, is N. This is the maximum value of ${}^{q}R^{\mathbf{Z}}$.

Example

Species-level components of the redundancy of subcommunity A (Figure 2.7b) can be simply calculated, in the naïve-type case, as the abundance of species s in the metacommunity divided by the abundance of species i in the subcommunity. More formally, in terms of relative abundances:

$$r^{\mathbf{I}}_{cows,A} : \frac{(\mathbf{I}\boldsymbol{p})_{cows}}{(\mathbf{I}\boldsymbol{P}_{.A})_{cows}} = \frac{p_{cows}}{P_{cows,A}} = \frac{6/8}{1/8} = 6$$

$$r^{\mathbf{I}}_{sheep,A} : \frac{(\mathbf{I}\boldsymbol{p})_{sheep}}{(\mathbf{I}\boldsymbol{P}_{.A})_{sheep}} = \frac{p_{sheep}}{P_{sheep,A}} = \frac{2/8}{1/8} = 2$$

This means there are 6 times more cows in the metacommunity as a whole, than there are in

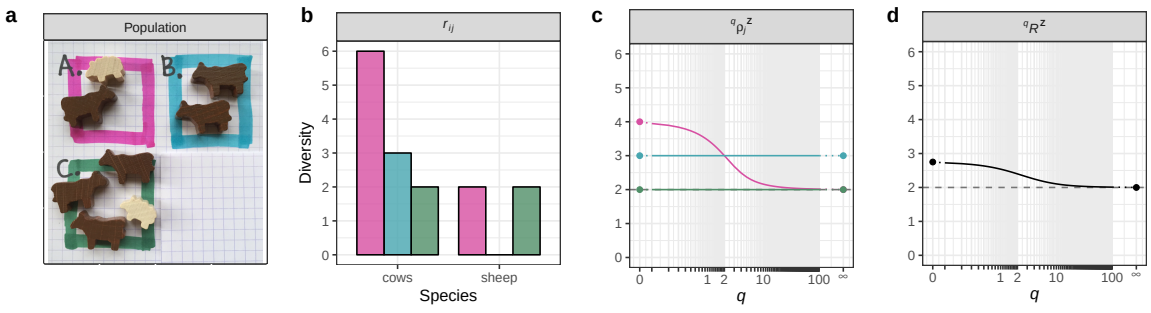


Figure 2.7: Partitioning redundancy at hierarchical levels of community structure: (a) metacommunity composition, and redundancy calculated at the (b) species, (c) subcommunity, and (d) metacommunity levels. As in (a), subcommunities A, B, and C are coloured pink, blue, and green, respectively. The dashed line highlights the minimum level of diversity.

subcommunity A. Therefore, if the subcommunity were destroyed, the metacommunity would still contain 5/6 of the current cow population.

At the subcommunity level (Figure 2.7c), ${}^{q}\rho_{j}^{\mathbf{I}}$ describes how redundant subcommunity A is of the metacommunity, which at $q = 0$ is the arithmetic mean of $r_{cows,A}^{\mathbf{I}}$ and $r_{sheep,A}^{\mathbf{I}}$:

$$\begin{aligned} {}^{0}\rho_{A}^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\bar{\boldsymbol{P}}, r_{i,A}^{\mathbf{I}}\right) \\ &= \left(\bar{P}_{cows,A} \times r_{cows,A}^{\mathbf{I}}\right) + \left(\bar{P}_{sheep,A} \times r_{sheep,A}^{\mathbf{I}}\right) \\ &= \left(\tfrac{1}{2} \times 6\right) + \left(\tfrac{1}{2} \times 2\right) = 4 \end{aligned}$$

This value is high when species that have a low ordinariness in subcommunity j have a high ordinariness across the metacommunity as a whole, and the subcommunity is small. As q increases, ${}^{q}\rho_{j}^{\mathbf{I}}$ becomes increasingly more conservative, until a minimum is reached at $q = \infty$, where only the least redundant species is considered (*i.e.* ${}^{\infty}\rho_{A}^{\mathbf{I}} = r_{sheep,A}^{\mathbf{I}} = 2$).

Since all species in subcommunity C are equally redundant, each comprising half of their total abundance in the metacommunity, the average redundancy of species in subcommunity C remains constant for all values of q (*i.e.* ${}^{0}\rho_{C}^{\mathbf{I}} = r_{cows,C}^{\mathbf{I}} = r_{sheep,C}^{\mathbf{I}} = 2$). Likewise for subcommunity B, since ${}^{q}\rho_{j}^{\mathbf{I}}$ is calculated only with respect to the species present in each subcommunity, all species are equally redundant (there is only one), and therefore ${}^{q}\rho_{j}^{\mathbf{Z}}$ is constant for all values of q (with value 3).

At the metacommunity level (Figure 2.7d), the average redundancy across subcommunities at $q = 0$ is calculated as the arithmetic mean of ${}^{0}\rho_{A}^{\mathbf{I}}$, ${}^{0}\rho_{B}^{\mathbf{I}}$, and ${}^{0}\rho_{C}^{\mathbf{I}}$:

$$\begin{aligned} {}^{0}R^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{0}\rho_{A}^{\mathbf{I}}\right) \\ &= \left(p_A \times {}^{0}\rho_{A}^{\mathbf{I}}\right) + \left(p_B \times {}^{0}\rho_{B}^{\mathbf{I}}\right) + \left(p_C \times {}^{0}\rho_{C}^{\mathbf{I}}\right) \\ &= \left(\tfrac{1}{4} \times 4\right) + \left(\tfrac{1}{4} \times 3\right) + \left(\tfrac{1}{2} \times 2\right) = 2\tfrac{3}{4} \end{aligned}$$

As q increases, ${}^{q}R^{\mathbf{I}}$ decreases, until at $q = \infty$ only the of the least redundant species, in the least redundant subcommunity is observed (*i.e.* ${}^{\infty}R^{\mathbf{I}} = {}^{\infty}\rho_{A}^{\mathbf{I}} = r_{sheep,A}^{\mathbf{I}} = 2$).

2.2.6.3 Effective number of distinct subcommunities (${}^{q}\bar{\beta}_{j}^{\mathbf{Z}}$ and ${}^{q}\bar{B}^{\mathbf{Z}}$)

Subcommunity ${}^{q}\bar{\beta}_{j}^{\mathbf{Z}} = 1/{}^{q}\bar{\rho}_{j}^{\mathbf{Z}}$ is *an estimate of the effective number of subcommunities in the metacommunity.* To understand this, it is useful to consider ${}^{q}\bar{B}^{\mathbf{Z}}$, the average of the ${}^{q}\bar{\beta}_{j}^{\mathbf{Z}}$s, which is *the effective number of completely distinct subcommunities in the metacommunity.* Just as the effective number of distinct species is greatest when distinct species are equally abundant, the effective number of distinct subcommunities is greatest when distinct subcommunities are of equal size. Generally, ${}^{q}\bar{\beta}_{j}^{\mathbf{Z}}$ is an estimate of ${}^{q}\bar{B}^{\mathbf{Z}}$ based on subcommunity j, and is high when

that subcommunity is both distinctive and small. For example, when the subcommunities are completely distinct (in the naïve-community case), ${}^{q}\bar{B}^{\mathbf{Z}}$ is equal to the Hill number ${}^{q}D(\boldsymbol{w})$, and if all of the subcommunities are of equal size then ${}^{q}\bar{B}^{\mathbf{Z}}$ takes its maximum value of N.

Example

Species-level components of ${}^{0}\bar{\beta}^{\mathbf{I}}_{A}$ (Figure 2.8b) – the estimate of the effective number of distinct subcommunities – are calculated in the naïve sense as a ratio of the abundance of species in the subcommunity to the abundance of species in the metacommunity:

$$\bar{b}^{\mathbf{I}}_{cows,A} = \frac{\left(\mathbf{I}\bar{\boldsymbol{P}}_{\cdot A}\right)_{cows}}{\left(\mathbf{I}\boldsymbol{p}\right)_{cows}} = \frac{\bar{P}_{cows,A}}{p_{cows}} = \frac{1/2}{6/8} = \frac{2}{3}$$

$$\bar{b}^{\mathbf{I}}_{sheep,A} = \frac{\left(\mathbf{I}\bar{\boldsymbol{P}}_{\cdot A}\right)_{sheep}}{\left(\mathbf{I}\boldsymbol{p}\right)_{sheep}} = \frac{\bar{P}_{sheep,A}}{p_{sheep}} = \frac{1/2}{2/8} = 2$$

The subcommunity estimate of the effective number of distinct subcommunities in the metacommunity (Figure 2.8c) is calculated as the inverse of ${}^{0}\bar{\rho}^{\mathbf{I}}_{A}$, or:

$$\begin{aligned}
{}^{0}\bar{\beta}^{\mathbf{I}}_{A} &= \mathrm{M}_{q-1}\left(\bar{\boldsymbol{P}}, \bar{b}^{\mathbf{I}}_{i,A}\right) \\
&= \left(\left(\bar{P}_{cows,A} \times (\bar{b}^{\mathbf{I}}_{cows,A})^{-1}\right) + \left(\bar{P}_{sheep,A} \times (\bar{b}^{\mathbf{I}}_{sheep,A})^{-1}\right)\right)^{-1} \\
&= \left(\left(\tfrac{1}{2} \times \tfrac{3}{2}\right) + \left(\tfrac{1}{2} \times \tfrac{1}{2}\right)\right)^{-1} = 1
\end{aligned}$$

In this case, this is simply due to the way the species-level components are averaged ($q = 0$ calculates the arithmetic mean). Since the species-level components are not equal, ${}^{q}\bar{\beta}^{\mathbf{I}}_{j}$ reaches its *least* conservative value at $q = \infty$ (since ${}^{q}\bar{\beta}^{\mathbf{I}}_{j}$ is the inverse of ${}^{q}\bar{\rho}^{\mathbf{I}}_{j}$, which is *most* conservative at $q = \infty$), *i.e.* ${}^{\infty}\bar{\beta}^{\mathbf{I}}_{A} = \max \bar{b}^{\mathbf{I}}_{ij} = \bar{b}^{\mathbf{I}}_{sheep,A} = 2$). On the other hand, subcommunity C estimates that the metacommunity holds exactly one subcommunity because the relative abundance of species in subcommunity C exactly equals the relative abundance of species in the metacommunity (the subcommunity is fully representative of both species in the metacommunity). This value remains constant for all values of q, since $\bar{b}^{\mathbf{I}}_{cows,C} = \bar{b}^{\mathbf{I}}_{sheep,C} = 1$.

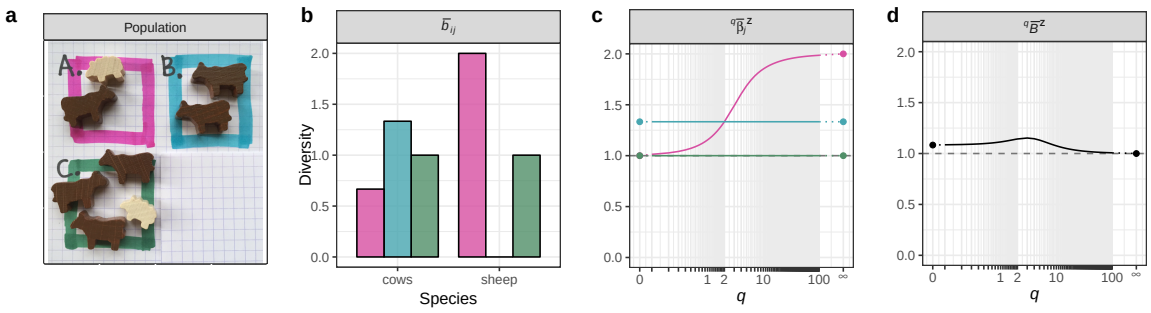


Figure 2.8: Partitioning the effective number of subcommunities at hierarchical levels of community structure: (a) metacommunity composition, and the effective number of subcommunities calculated at the (b) species, (c) subcommunity, and (d) metacommunity levels. As in (a), subcommunities A, B, and C are coloured pink, blue, and green, respectively. The dashed line highlights the minimum level of diversity.

Subcommunity B, on the other hand, has a higher estimate of ${}^0\bar{\beta}^{\mathbf{I}}_B$ because the sheep are missing and therefore more subcommunities are needed to create the metacommunity as a whole. This value remains constant for all values of q, as there is only one species present, so all averages are the same.

At the metacommunity level (Figure 2.8d), the average estimate of the effective number of subcommunities at $q = 0$ is calculated as the arithmetic mean of ${}^0\bar{\beta}^{\mathbf{I}}_j$:

$$\begin{aligned} {}^0\bar{B}^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^0\bar{\beta}^{\mathbf{I}}_A\right) \\ &= \left(p_A \times {}^0\bar{\beta}^{\mathbf{I}}_A\right) + \left(p_B \times {}^0\bar{\beta}^{\mathbf{I}}_B\right) + \left(p_C \times {}^0\bar{\beta}^{\mathbf{I}}_C\right) \\ &= \left(\tfrac{1}{4} \times 1\right) + \left(\tfrac{1}{4} \times 1\tfrac{1}{3}\right) + \left(\tfrac{1}{2} \times 1\right) = 1\tfrac{1}{12} \end{aligned}$$

Both ${}^q\bar{B}^{\mathbf{Z}}$ and ${}^qB^{\mathbf{Z}}$ are different from other measures within the framework in that they are not monotonic, not necessarily consistently increasing or decreasing with q (*e.g.* see Figure 2.8d). At $q = \infty$, ${}^q\bar{B}^{\mathbf{I}}$ is calculated as the most conservative value of ${}^\infty\bar{\beta}^{\mathbf{I}}_j$. That is, ${}^\infty\bar{B}^{\mathbf{I}} = \min_j {}^\infty\bar{\beta}^{\mathbf{I}}_j = {}^\infty\bar{\beta}^{\mathbf{I}}_C = \max_i \bar{b}^{\mathbf{I}}_{i,C} = \bar{b}^{\mathbf{I}}_{sheep,C} = 1$

2.2.6.4 Distinctiveness (${}^q\beta^{\mathbf{Z}}_j$ and ${}^qB^{\mathbf{Z}}$)

Subcommunity ${}^q\beta^{\mathbf{Z}}_j$ diversity measures *the overall distinctiveness of a given subcommunity* – or as the reciprocal of ${}^q\rho^{\mathbf{Z}}_j$, the concentration of species within it – or equivalently *the average distinctiveness of each individual within a given subcommunity relative to the metacommunity* (Table 2.3). It takes its maximum value of 1 when ${}^q\rho^{\mathbf{Z}}_j$ is minimised. This occurs when every individual in the subcommunity is completely dissimilar to every individual outside the subcommunity, so the subcommunity is completely distinct. It is small when the subcommunity has much in common with the rest of the metacommunity, be that through shared species or high similarity between species. It can also be understood as a kind of turnover: not in the traditional sense between adjacent subcommunities, but between subcommunity j and the rest of the metacommunity.

The average of the subcommunity ${}^q\beta^{\mathbf{Z}}_j$ diversities, ${}^qB^{\mathbf{Z}}$, is a measure of *the average distinctiveness of subcommunities in the metacommunity* (and lies between 0 and 1). To see the connection with turnover, consider a naïve-type case with each subcommunity containing the same number of species in equal abundance, and each species present in k subcommunities, with a fraction $1/k$ changing from one subcommunity to the next ordered along a spatial gradient. Thus we have a turnover of $1/k$ along the gradient: and indeed, ${}^q\beta^{\mathbf{Z}}_j = 1/k$ (apart from at the ends), and ${}^qB^{\mathbf{Z}} \to 1/k$ as the number of subcommunities becomes large.

Example

The distinctiveness of cows in subcommunity A, in the naïve sense, is calculated as the number of cows in the subcommunity divided by the number of cows in the metacommunity. Which is written more formally, in terms of relative abundances:

$$b^{\mathbf{I}}_{cows,A} : \frac{(\mathbf{I}\boldsymbol{P}_{\cdot A})_{cows}}{(\mathbf{I}\boldsymbol{p})_{cows}} = \frac{P_{cows,A}}{p_{cows}} = \frac{1/8}{6/8} = \frac{1}{6}$$
$$b^{\mathbf{I}}_{sheep,A} : \frac{(\mathbf{I}\boldsymbol{P}_{\cdot A})_{sheep}}{(\mathbf{I}\boldsymbol{p})_{sheep}} = \frac{P_{sheep,A}}{p_{sheep}} = \frac{1/8}{2/8} = \frac{1}{2}$$

At the subcommunity level, the distinctiveness of species in subcommunity A at $q = 0$ is calculated as the inverse of ${}^{0}\rho^{\mathbf{I}}_{A}$, or:

$$\begin{aligned}
{}^{0}\beta^{\mathbf{I}}_{A} &= \mathrm{M}_{q-1}\left(\bar{\boldsymbol{P}}, b^{\mathbf{I}}_{i,A}\right) \\
&= \left(\left(\bar{P}_{cows,A} \times (b^{\mathbf{I}}_{cows,A})^{-1}\right) + \left(\bar{P}_{sheep,A} \times (b^{\mathbf{I}}_{sheep,A})^{-1}\right)\right)^{-1} \\
&= \left(\left(\tfrac{1}{2} \times 6\right) + \left(\tfrac{1}{2} \times 2\right)\right)^{-1} = \tfrac{1}{4}
\end{aligned}$$

which means that subcommunity A has a distinctiveness of 25%. The distinctiveness of a subcommunity is greatest when species have a high subcommunity ordinariness, but low metacommunity ordinariness, and when the subcommunity itself is small. Since ${}^{q}\beta^{\mathbf{I}}_{j}$ is the inverse of ${}^{q}\rho^{\mathbf{I}}_{j}$, which is most conservative at $q = \infty$, the distinctness of subcommunity A is *least* conservative at $q = \infty$ (*i.e.* ${}^{\infty}\beta^{\mathbf{I}}_{A} = \max_i b^{\mathbf{I}}_{ij} = b^{\mathbf{I}}_{sheep,A} = \frac{1}{2}$).

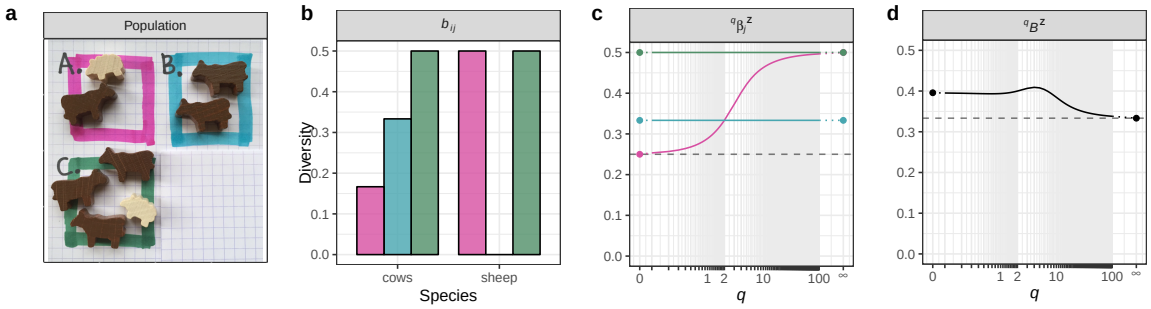


Figure 2.9: Partitioning distinctiveness at hierarchical levels of community structure: (a) metacommunity composition, and distinctiveness calculated at the (b) species, (c) subcommunity, and (d) metacommunity levels. As in (a), subcommunities A, B, and C are coloured pink, blue, and green, respectively. The dashed line highlights the minimum level of diversity.

The arithmetic mean across subcommunities, gives the average distinctiveness of subcommunities at $q = 0$:

$$\begin{aligned} {}^{0}B^{\mathbf{I}} &= \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{0}\bar{\beta}^{\mathbf{I}}_{A}\right) \\ &= \mathrm{M}_{1-q}\left(\boldsymbol{w}, {}^{0}\beta^{\mathbf{I}}_{A}\right) \\ &= \left(p_A \times {}^{0}\beta^{\mathbf{I}}_{A}\right) + \left(p_B \times {}^{0}\beta^{\mathbf{I}}_{B}\right) + \left(p_C \times {}^{0}\beta^{\mathbf{I}}_{C}\right) \\ &= \left(\tfrac{1}{4} \times \tfrac{1}{4}\right) + \left(\tfrac{1}{4} \times \tfrac{1}{3}\right) + \left(\tfrac{1}{2} \times \tfrac{1}{2}\right) = \tfrac{19}{48} \end{aligned}$$

which tells us that, on average, subcommunities are approximately 40% distinct. As with ${}^{q}\bar{B}^{\mathbf{I}}$, ${}^{q}B^{\mathbf{I}}$ is not monotonic, not consistently increasing or decreasing with q. At $q = \infty$, ${}^{q}B^{\mathbf{I}}$ is calculated as the *most* conservative value of ${}^{\infty}\beta^{\mathbf{I}}_{j}$ (*i.e.* ${}^{\infty}B^{\mathbf{I}} = \min_j {}^{\infty}\beta^{\mathbf{I}}_{j} = {}^{\infty}\beta^{\mathbf{I}}_{B} = \max_i b^{\mathbf{I}}_{iB} = b^{\mathbf{I}}_{cows,B} = \frac{1}{3}$).

2.3 The `rdiversity` package `v1.3.0`

All work contained within this thesis was done in R. For this purpose, I developed software – `rdiversity` – to provide the necessary functionality and intuitive access to the framework. This package is published on CRAN. In the following section, a simple example is used to demonstrate the functionality of this software.

First the package must be installed. This is done in the standard way:

```
install.packages("rdiversity")
```

The package can then be loaded into an instance of R and for this example, a simple metacommunity is generated. To calculate the diversity of this metacommunity, a `metacommunity` object must be created. This object contains all the information needed to calculate diversity.

```
library(rdiversity)

# Generate relative abundances from abundance in Chapter 2 examples
dat <- cbind(c(1,1), c(2,0), c(3,1))
dat <- dat/sum(dat)
colnames(dat) <- LETTERS[1:3]
row.names(dat) <- c("cows", "sheep")

# Create metacommunity object
mc <- metacommunity(partition=dat)
```

The `metacommunity()` function takes two arguments, `partition` and `similarity` (assuming naïve-type if similarity is missing), and creates an S4 object of class `metacommunity` with the following slots:

`@type_abundance`	The relative abundance of species within each subcommunities relative to the metacommunity as a whole; a matrix, $\mathbf{P} = [P_{ij}]$
`@subcommunity_weights`	The relative weights of subcommunities in the metacommunity as a whole; a vector, $\boldsymbol{w} = (w_1, ..., w_N)$
`@type_weights`	The relative weights of species in the metacommunity as a whole, a vector $\boldsymbol{p} = (p_1, \ldots, p_S)$
`@dat_ID`	The type of diversity being calculated
`@similarity`	The pair-wise similarity of species within a metacommunity; a matrix, $\mathbf{Z} = [Z_{ii'}]$
`@ordinariness`	The (cached) ordinariness of species within each of the subcommunities; a matrix, $\mathbf{ZP} = [ZP_{ij}]$

If the dataset is small enough, the similarity and ordinariness matrices are pre-computed and stored in the following slots:

`@similarity_components`	The components necessary to calculate similarity (empty when precompute_dist = TRUE)
`@similarity_parameters`	List of parameters associated with converting pairwise distances to similarities
`@similarity_parameters`	List of parameters associated with converting pairwise distances to similarities

If the dataset is too large, then similarity (and ordinariness) are calculated during the diversity calculation, on the fly, with associated data stored in this slot:

`@similarity_components`	The components necessary to calculate similarity (empty when precompute_dist = TRUE)

A metacommunity originating from a phylogeny (see Chapter 4) may contain three additional slots:

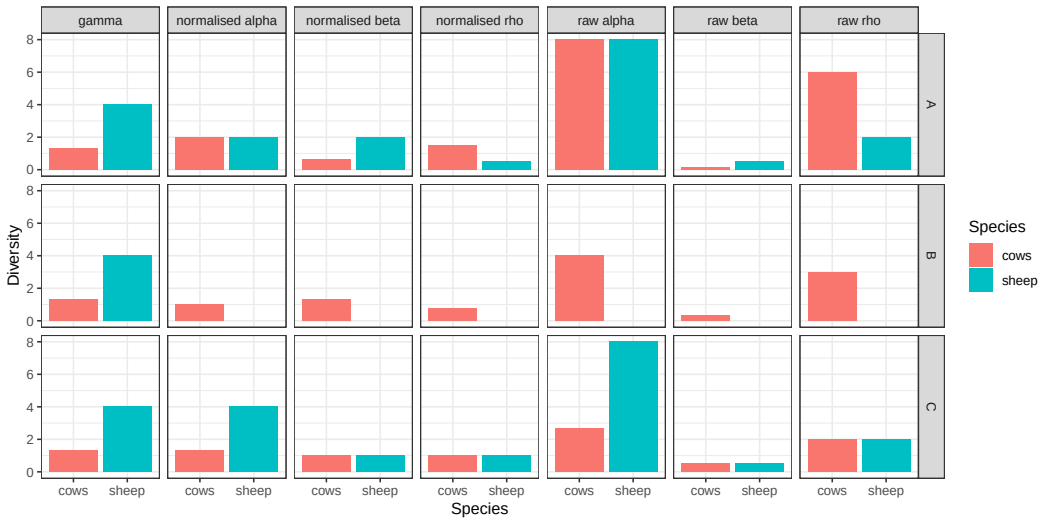
`@raw_abundance`	The relative abundance of present-day species (where `@type_abundance` is then considered to be 'historical species', see Leinster & Cobbold, 2012); a matrix, $\mathbf{P'} = [P'_{ij}]$.
`@raw_structure`	The length of evolutionary history of each 'historical species'; a matrix.
`@parameters`	Parameters associating the 'historical species' to present-day species.

The `inddiv()`, `subdiv()`, and `metadiv()` functions calculate species-level components, and subcommunity- and metacommunity-levels diversities, respectively. Each of these functions take the same arguments:

`data`	This can be input as a metacommunity object (to calculate all measures) or as an individual diversity component (to calculate a single measure).
`qs`	A vector of q-values.

The `inddiv()` function is used to calculate species-level components. Interestingly, in the plot below, the species-level gamma component of each species is the same for each subcommunity. This is because the gamma component of sheep in subcommunity A is calculated as the inverse of the ordinariness of sheep in the metacommunity as a whole, likewise for sheep in subcommunities B and C.

```
library(ggplot2)

# Calculate and plot species-level components at q = 0
sp_res <- rdiversity::inddiv(data=mc, qs=0)

# Plot species-level components
ggplot(sp_res, aes(x=type_name, y=diversity, fill=type_name)) +
  theme_bw() + geom_bar(stat="identity") +
  facet_grid(facets=partition_name~measure) +
  labs(x="Species", y="Diversity", fill="Species")
```



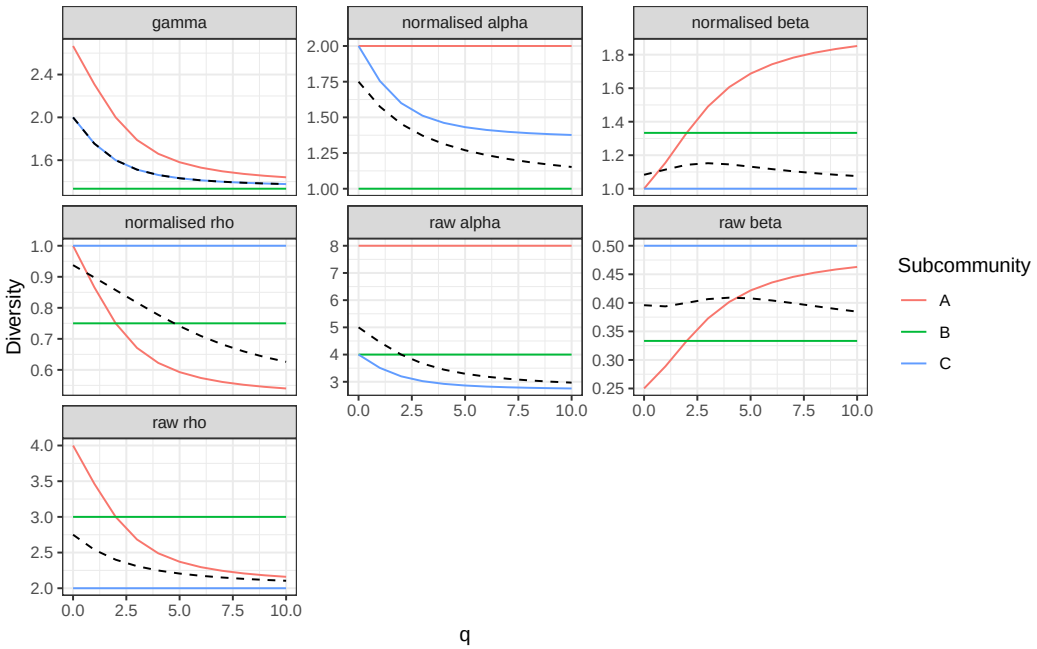
The `subdiv()` function is used to calculate subcommunity-level diversity. The `inddiv()`, `subdiv()`, and `metadiv()` functions all produce outputs in the same format as below.

```
# Calculate subcommunity-level diversity from q = 0 to q = 10
sres <- subdiv(data=mc, qs=0:10)

# Print subcommunity-level results
sres
```

```
# A tibble: 231 x 7
   measure       q type_level type_name partition_level partition_name diversity
   <chr>     <int> <chr>      <chr>     <chr>           <fct>                <dbl>
 1 raw alpha     0 types      ""        subcommunity    A                     8.00
 2 raw alpha     0 types      ""        subcommunity    B                     4.00
 3 raw alpha     0 types      ""        subcommunity    C                     4.00
 4 raw alpha     1 types      ""        subcommunity    A                     8.00
 5 raw alpha     1 types      ""        subcommunity    B                     4.00
 6 raw alpha     1 types      ""        subcommunity    C                     3.51
 7 raw alpha     2 types      ""        subcommunity    A                     8.00
 8 raw alpha     2 types      ""        subcommunity    B                     4.00
 9 raw alpha     2 types      ""        subcommunity    C                     3.20
10 raw alpha     3 types      ""        subcommunity    A                     8.00
# ... with 221 more rows
```

The `metadiv()` function is used to calculate metacommunity-level diversity.

```
# Calculate metacommunity-level diversity from q = 0 to q = 10
mres <- metadiv(data=mc, qs=0:10)

# Plot subcommunity and metacommunity diversities together
ggplot() + theme_bw() + facet_wrap(facets=~measure, scales="free_y") +
  geom_line(aes(x=q, y=diversity, colour=partition_name), sres) +
  geom_line(aes(x=q, y=diversity), mres,
            linetype="dashed", colour="black") +
  labs(x="q", y="Diversity", colour="Subcommunity")
```



For advanced functionality and detailed examples, see https://github.com/boydorr/rdiversity.

2.4 Conclusion

Reeve et al.'s (2016) framework of similarity-sensitive diversity measures is able to quantify the compositional variation present in a community from multiple perspectives and sensitivities. Using a simple caricature metacommunity, each of these measures were able to pick out distinguishing features for each subcommunity. In the simplest case, subcommunity ${}^{0}\bar{\alpha}_{B}^{\mathbf{Z}}$ correctly identified the presence of a single species in subcommunity B. Whereas the diversity profile plot from ${}^{q}\bar{\alpha}_{A}^{\mathbf{Z}}$ showed an evenly distributed subcommunity comprising two species. Subcommunity C comprised exactly half of each species in the metacommunity, which was identified from the redundancy profile, ${}^{q}\rho_{C}^{\mathbf{Z}}$.

More generally, the subcommunity measures allow subcommunities with high inherent diversity to be identified, per individual or overall (${}^{q}\alpha_{j}^{\mathbf{Z}}$ and ${}^{q}\bar{\alpha}_{j}^{\mathbf{Z}}$), with high distinctiveness or that are very redundant in the metacommunity (${}^{q}\beta_{j}^{\mathbf{Z}}$ or ${}^{q}\rho_{j}^{\mathbf{Z}}$), ones that are representative of large or only small, distinct parts of the metacommunity (${}^{q}\bar{\beta}_{j}^{\mathbf{Z}}$ and ${}^{q}\bar{\rho}_{j}^{\mathbf{Z}}$), and ones with strong per-individual influence on metacommunity diversity (${}^{q}\gamma_{j}^{\mathbf{Z}}$). Metacommunity gamma diversity (${}^{q}G^{\mathbf{Z}}$) and normalised subcommunity alpha diversity (${}^{q}\bar{\alpha}_{j}^{\mathbf{Z}}$) are equivalent to the usual notions of diversity for an undivided group, while the other metacommunity and subcommunity diversity measures in Tables 2.1, 2.3, and 2.2 are novel.

These measures all depend on the parameter q and the simple examples presented in this chapter highlight the benefits of calculating a diversity profile. Rather than having to select a single measure of diversity (*e.g.* Shannon diversity) and justify that choice, we can look at the whole profile of measures simultaneously and determine the result more generally than would be possible from a single diversity measure (see Hill, 1973; Tóthmérész, 1995; Leinster & Cobbold, 2012). Reeve et al.'s (2016) framework incorporates this concept for all measures of alpha, beta, and gamma. In particular, the exact interpretation of ${}^{q}\rho_{j}^{\mathbf{Z}}$ as the redundancy of subcommunity j varies with q (and similarly for ${}^{q}\beta_{j}^{\mathbf{Z}}$, ${}^{q}\bar{\beta}_{j}^{\mathbf{Z}}$ and ${}^{q}\bar{\rho}_{j}^{\mathbf{Z}}$. For instance, it may be useful to ask what it means for the redundancy of a subcommunity to achieve the minimal possible value, 1. When $q = 0$, anti-conservatively, this means that no species present in the subcommunity can be found anywhere else in the metacommunity, whereas at $q = \infty$, conservatively, it is sufficient for the subcommunity to have just one species not found anywhere else.

In the next chapter, I will investigate the diversity measures in a series of real-world examples, examining what diversity signals can be extracted from a set of previously studied datasets to understand their power to extract meaningful signals from data.

Acknowledgements: Derivation of similarity-sensitive diversity framework and some of the explanatory text for individual diversity measures have been taken from joint work with my supervisors (Reeve et al., 2016).

Chapter 3

Case studies

'[The] key question we should ask of a diversity metric is: does it measure the thing we are biologically interested in?'

— Ricotta (2010)

3.1 Abstract

In this chapter, three empirical case studies comprising several distinct problems were selected to examine the power and versatility of the family of similarity-sensitive diversity measures described in Chapter 2. These case studies were chosen to showcase disparate problems in distinct fields that can benefit from diversity analysis, each requiring very different signals to be detected. The overarching challenge was to determine whether the framework could be usefully applied to these very different datasets. And to this end, the framework is validated by identifying recognisable features of the populations being studied and comparing them to known results. The first case study illustrates the utility of these measures in a familiar setting, that of tree diversity in Barro Colorado Island, whilst each of the subsequent case studies applies new measures of beta diversity to less typical problems, to investigate population demographics and antimicrobial resistance.

Diversity measures are commonly used in ecological analyses of biodiversity. The first case study uses Reeve et al.'s (2016) framework to evaluate the compositional structure of the Barro Colorado Island (BCI) Forest dynamics plot. The framework enables us to detect areas of forest disturbance and a swamp, which has unique plant types. The second case study examines the population demographics of England and Wales from 2001 census data, where subcommunity *representativeness* (${}^{q}\bar{\rho}_j^{\mathbf{Z}}$) is used to identify towns with unusual age distributions such as the university town Keele. Finally, the third case study investigates the transmission of antimicrobial resistance (AMR) in sympatric human and animal host populations, where measures of subcommunity *redundancy* (${}^{q}\rho_j^{\mathbf{Z}}$) and *distinctiveness*, (${}^{q}\beta_j^{\mathbf{Z}}$) are used to detect emerging resistance in host populations.

3.2 Case study: Investigating the spatial and temporal biodiversity of the Barro Colorado Island Forest dynamics plot

3.2.1 Introduction

The Barro Colorado Island (BCI) Forest dynamics plot is a permanent 50 ha plot (1000 m × 500 m) of moist seasonal old-growth forest, established in 1980 by Steve Hubbell and Robin Foster at the Smithsonian Tropical Research Institute's field station on Barro Colorado Island, central Panama (Condit, 1998; Hubbell et al., 1999). The island itself[1] has been a biological reserve and laboratory since 1923, providing 95 years of historical data on the biotia, climate, and geology of the island. Consequentially, many scientific papers have been published on these data[2], with topics ranging from the maintenance of tree diversity (Hubbell et al., 1990; Condit et al., 2012a) to the impact of climate change (Condit et al., 1996; Feeley et al., 2011; Condit et al., 2017b).

The BCI dataset is well suited to investigate diversity measures. Over the last 37 years, 8 complete inventories of the BCI Forest dynamics plot have been recorded at approximately 5-year intervals (in 1981-83, 1985, 1990, 1995, 2000, 2005, 2010, and 2015). Data comprises all free-standing woody individuals (hereafter referred to as *trees*, although palms and shrubs are also included in the dataset) within the 50 ha site, having at least one stem $\geq$ 1 cm in diameter at least 1.3 m from ground level. Over 350,000 individual trees have been censused. The species' identity and spatial x-y coordinates of individual trees were originally collected with the support of the Center for Tropical Forest Science (CTFS) of the Smithsonian Tropical Research Institute (Condit et al., 2012b, 2017a). The BCI study site is relatively uniform, consisting almost entirely of well-drained upland soil. However, variation does exist: Harms et al. (2001) identifies six unique habitats from the 1990 census (see also Condit, 1998). The most distinct habitat is a seasonally inundated 1.5 ha swamp (recorded at the end of the wet season in 1992), which is surrounded by low-lying plateau (Figure 3.1).

The aim of this case study is to illustrate the properties of Reeve et al.'s (2016) framework of measures, by extracting signals in the data corresponding to regions of interest within the BCI study site. Variation in biodiversity is examined across space and time, first discounting similarity between species (to highlight the utility of each measure), and then by incorporating Shimatani's (2001) measure of taxonomic similarity (hereafter referred to as 'taxonomic diversity').

[1] A former hilltop, the Barro Colorado Island became an island when the Chagres River was dammed in the creation of the main reservoir for the Panama Canal, Gutan Lake (Hubbell & Foster, 1992).

[2] Hubbell & Foster (1992) report that 1500 were published as of 1992

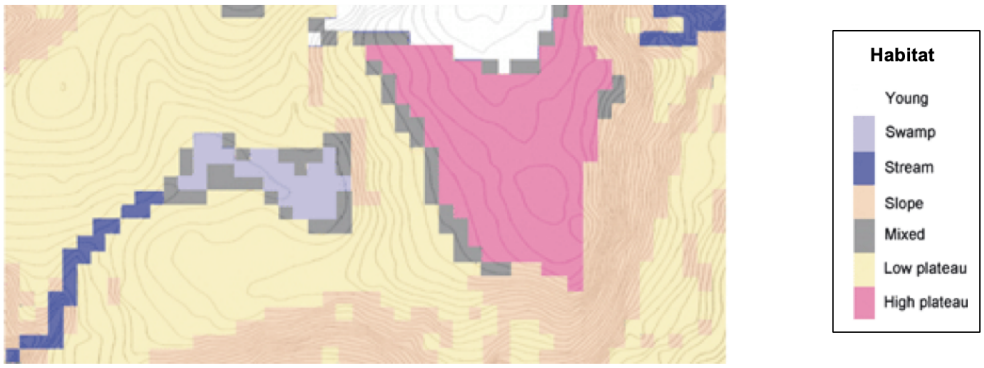


Figure 3.1: Distribution of habitats in the Barro Colorado Island Forest dynamics plot: adapted from Fig.1, Kanagaraj et al. (2011). Map shows the 50 ha plot (1000 m×500 m) divided into 20 m×20 m quadrats. Six distinct habitats were identified from the 1990 census. The remaining 66 quadrats that could not be unambiguously assigned are designated as 'Mixed' habitat.

Acknowledgements

The Barro Colorado Island Forest data was originally collected with the support of the Center for Tropical Forest Science (CTFS) of the Smithsonian Tropical Research Institute. The BCI forest dynamics research project was founded by S.P. Hubbell and R.B. Foster and is now managed by R. Condit, S. Lao, and R. Perez under the Center for Tropical Forest Science and the Smithsonian Tropical Research in Panama (Condit et al., 2012b, 2017a). Numerous organisations have provided funding, principally the U.S. National Science Foundation, and hundreds of field workers have contributed.

3.2.2 Methods

Spatial diversity

To assess the spatial diversity, the first census (1981/82) of the BCI Forest dynamics plot was examined, within which a mean of 188.24 trees (SD = 40.87) and 54.05 species (SD = 9.63) were recorded per quadrat. All trees contained within the 50 ha plot were considered to form a single metacommunity, which was partitioned into subcommunities defined as 20 m × 20 m quadrats, where $N = 1250$ (though a range of other grid sizes were also assessed). Figure 3.2 shows the abundance of trees in each quadrat.

In the simplest case, subcommunity-level diversity measures were used to investigate the biological variation of each quadrat in the study site. These measures calculate: the *effective number of tree-species* (${}^{q}\bar{\alpha}_j^{\mathbf{I}}$), an *estimate of naïve-community metacommunity diversity* (${}^{q}\alpha_j^{\mathbf{I}}$), the *contribution per tree toward metacommunity diversity* (${}^{q}\gamma_j^{\mathbf{I}}$), *representativeness* (${}^{q}\bar{\rho}_j^{\mathbf{I}}$), *redundancy* (${}^{q}\rho_j^{\mathbf{I}}$), an *estimate of the effective number of distinct subcommunities* (${}^{q}\bar{\beta}_j^{\mathbf{I}}$), and *distinctiveness* (${}^{q}\beta_j^{\mathbf{I}}$). The spatial diversity of the BCI dataset was calculated in the naïve-type case ($\mathbf{Z} = \mathbf{I}$), treating each species as completely distinct, and therefore allowing the utility of each measure to be examined with greater clarity. In order to fully discern the variation in

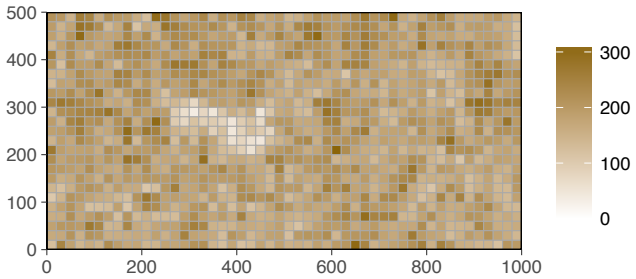


Figure 3.2: Tree abundance of the Barro Colorado Island Forest plot 1981/82 census: Coloured according to the number of trees per 20 m × 20 m quadrat across the 50 ha site.

community composition, each subcommunity measure was calculated at $q \in \{0, 1, 2, \infty\}$ and these results were combined to expose a complete description of the dataset.

Since groups of trees separated by family and genus intrinsically contain more diversity than those differing only by species, taxonomic spatial diversity was then calculated using taxonomic distance measures defined by Shimatani (2001). In this way, species and phenotypic diversity are combined to reveal a more accurate representation of subcommunity structure. Since species are classified into hierarchical taxonomic groups (based on characteristic similarities) pairwise taxonomic distance can be easily defined, as $d_{ii'} = 0$ when individuals belong to the same species, $d_{ii'} = 1$ when they belong to the same genus but different species, $d_{ii'} = 2$ when they belong to the same family but different genera, $d_{ii'} = 3$ if they belong to the same order[1] but different families, and $d_{ii'} = 4$ otherwise. These distances were converted into pairwise similarities, $Z_{ii'}$, and arranged as a similarity matrix $\mathbf{Z} = [Z_{ii'}]$ (hereafter denoted $\mathbf{Z}_{\mathbf{Z}_{tax}}$ to describe *taxonomic* similarity). Similarity was transformed such that $Z_{ii'} = 1 - (d_{ii'}/4)$, so: $Z_{ii'} = 1$ for conspecifics (when they belong to the same species), $Z_{ii'} = 0.75$ when they belong to the same genus, $Z_{ii'} = 0.5$ when they belong to the same family, $Z_{ii'} = 0.25$ when they belong to the same order, and $Z_{ii'} = 0$ otherwise. Shimatani's taxonomic diversity can be recovered from Reeve et al.'s (2016) framework by calculating ${}^{q}\gamma_j^{\mathbf{Z}_{tax}}$ when $q = 2$. However, as shown previously, higher values of q place more emphasis on their most conservative values (or least conservative in the case of ${}^{q}\beta_j^{\mathbf{I}}$ and ${}^{q}\bar{\beta}_j^{\mathbf{I}}$). Therefore here, taxonomic subcommunity diversities are calculated at $q = 1$: the *effective number of tree-species* (${}^{q}\bar{\alpha}_j^{\mathbf{Z}_{tax}}$), the *contribution per tree toward metacommunity diversity* (${}^{q}\gamma_j^{\mathbf{Z}_{tax}}$), *representativeness* (${}^{q}\bar{\rho}_j^{\mathbf{Z}_{tax}}$), and *distinctiveness* (${}^{q}\beta_j^{\mathbf{Z}_{tax}}$). As a default, $q = 1$ is used when there is no strong reason for another choice, as it corresponds to Shannon (1948) entropy and relative entropy Kullback & Leibler (1951), which are the most studied of the Rényi (1961) entropies.

Temporal diversity

By combining data across multiple time points, the dataset can be re-partitioned in such a way as to reveal the biological variation of each quadrat from a temporal perspective, and thus describe how the forest has changed over time. The 50 ha plot was again partitioned

[1] Shimatani (2001) used 'subclass', which is substituted for 'order' here, due to the data available

into 20 m × 20 m quadrats. In this case, a subcommunity was defined as a single quadrat at a single point in time, whereas a metacommunity was defined as a single quadrat over 8 time points (subcommunities), each corresponding to a single census (1981-82, 1985, 1990, 1995, 2000, 2005, 2010, and 2015). Metacommunity ${}^{q}B^{\mathbf{I}}$ diversity (the *average distinctiveness of subcommunities*) is indicative of a high turnover in species composition over time. This measure was calculated to reveal temporal changes in biological diversity within each quadrat. Results were plotted as a 2-dimensional heat map to allow each grid location to be compared across the BCI study site.

3.2.3 Results and discussion

3.2.3.1 Spatial diversity - varying grid size

Subcommunity diversity was examined at a range of grid sizes to determine the optimal resolution for subsequent analyses. Representativeness is calculated at $q = 1$ in the naïve-type case (${}^{1}\bar{\rho}_{j}^{\mathrm{I}}$), and compared at different spatial resolutions by partitioning the study site into $N \in \{50, 200, 1250, 5000, 20000\}$ quadrats. Values are plot with a standardised colour gradient in Figures 3.3a through 3.3f to show how representativeness changes as grid size is decreased.

From sampling theory, we know that classical measures of alpha diversity approach gamma diversity as quadrats increase in size and sampling resolution decreases (Tuomisto, 2010b). It follows then that, subcommunity representativeness (which takes a maximum value of one when the distribution of tree-species in a subcommunity is the same as the metacommunity as a whole) should also increase when spatial resolution is low. When quadrats are large (or few in number), their species distribution more closely matches that of the study site as a whole (as in Figures 3.3a & 3.3b), converging to complete representativeness when the metacommunity is undivided. In other words, as spatial resolution decreases, it is more difficult to delineate regions of interest as they are averaged across the metacommunity. Therefore, the lower the spatial resolution, the more uniform the metacommunity appears, as regions (and subcommunities) of interest are diffused.

When examined in greater detail (at higher spatial resolutions), each quadrat is much less representative of the study site as a whole. This is likely due to each quadrat comprising a smaller fraction of the total population and higher heterogeneity (at this resolution). The higher the spatial resolution, the more structural variation in species composition is revealed. In the most extreme case however, it is increasingly difficult to discern any region of interest, since each quadrat is equally interesting (Figure 3.3f). At this scale, all of the quadrants are unrepresentative of the metacommunity, but no single quadrant stands out as being the least representative. When scaled appropriately, interesting small-scale features can be identified. For example, 37 quadrats are found to be completely empty (coloured pink Figure 3.3g), 32

of which are in or near the swamp (Figure 3.3h). For the purposes of this case study, this kind of detail is unnecessary. Nevertheless, the exceptional resolution of this dataset allows each measure of diversity to be examined without concern about sampling error. For partially sampled plots, particular care should be taken when selecting an appropriate spatial resolution as unusual areas may be due to stochastic sampling effects (see Chapter 5).

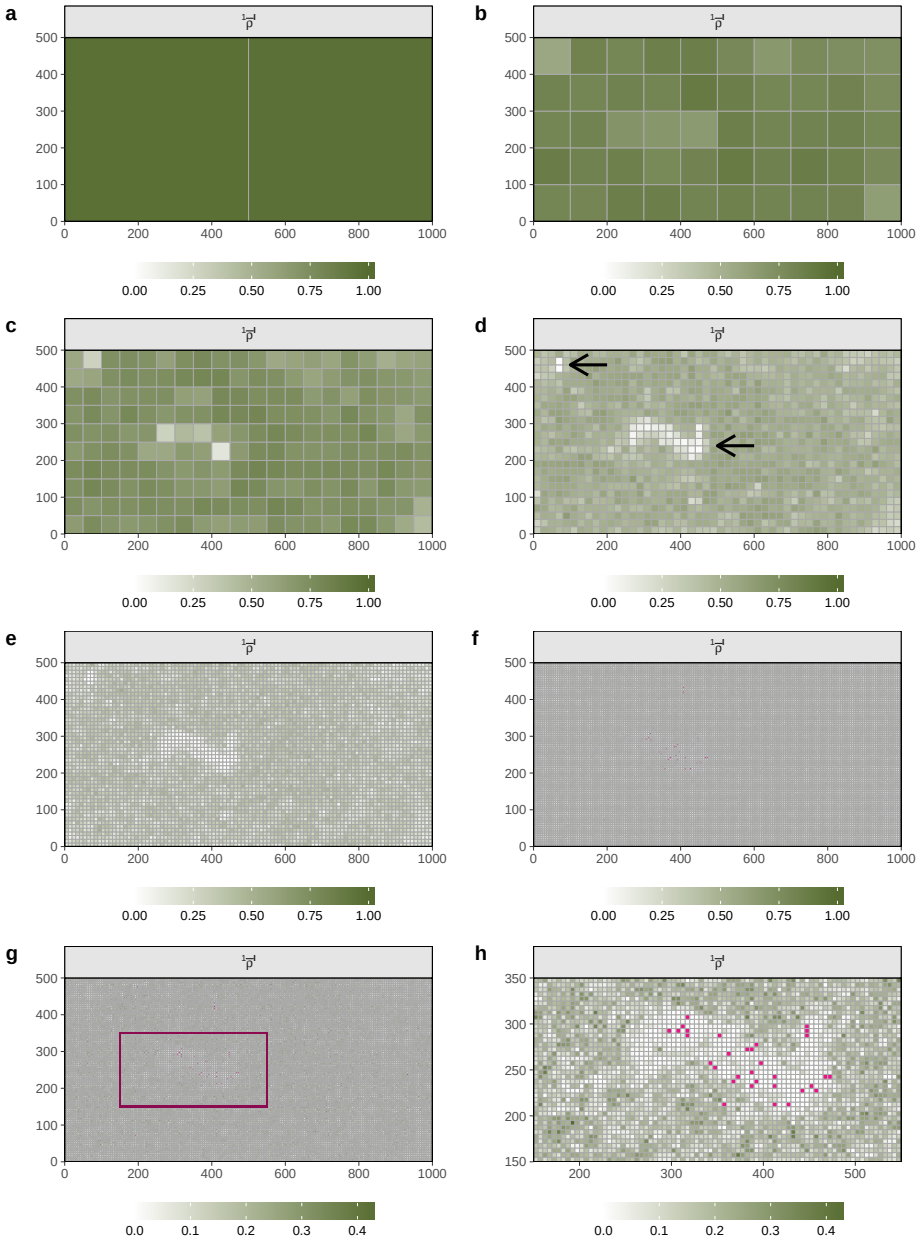


Figure 3.3: Heatmap showing the naïve-type spatial representativeness of the Barro Colorado Island Forest plot 1981/82 census: Figures are coloured according to the representativeness (subcommunity ${}^1\bar{\rho}_j^{\mathbf{I}}$) of each quadrat across the 50 ha site, where quadrats are sized: (a) 500 m × 500 m, (b) 100 m × 100 m, (c) 50 m × 50 m, (d) 20 m × 20 m, (e) 10 m × 10 m, and (f) 5 m × 5 m on a standardised colour scale. Arrows indicate regions of interest. Colours are rescaled for maximum contrast in figures (g) 5 m × 5 m and (h) 5 m × 5 m, which expands the red box. In figure (h), 32 quadrats within the swamp (37 across the study site as a whole) are found to be completely devoid of trees (shaded pink).

In summary then, the BCI Forest dynamics plot is a fully sampled dataset, which allows the study site to be examined at an exceptionally high resolution. Intrinsically, when grid sizes are too large, it is difficult to pinpoint the precise location of any regions of interest (Figures 3.3a & 3.3b). Conversely, when grid sizes are too small, it is difficult to identify clear features of subcommunity structure amongst the variation (Figures 3.3e & 3.3f). Therefore, a resolution of 20 m × 20 m was selected as being useful for further analysis (Figure 3.3d). At this scale, there are two areas (highlighted with arrows) that have low representativeness, which suggests that the distribution of the species at these sites differs from what is observed across the whole plot – a swamp in the centre-left and an area comprising invasive species in the top-left corner (Rick Condit, personal communication). The significance of these results is discussed in more detail later in this chapter.

3.2.3.2 Spatial diversity - varying the q parameter

To observe community assembly across the study site, subcommunity diversity measures were calculated at each 20 m × 20 m quadrat for $q \in \{0, 1, 2, \infty\}$ with similarity defined in the naïve-type case (where species are considered to be completely distinct, *i.e.* $\mathbf{Z} = \mathbf{I}$).

Normalised subcommunity alpha, ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$

The simplest and most commonly used measure of diversity is Whittaker's alpha diversity, which is the mean species diversity (effective number of species) per subcommunity, and can be calculated using metacommunity ${}^{q}\bar{A}^{\mathbf{I}}$. The constituent components of this measure can be quantified using subcommunity ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$, which measures the effective number of species in each subcommunity in isolation (${}^{q}\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.4). This value depends on how many species are present and their relative abundance, but is highest when a subcommunity contains a large number of species that are evenly distributed.

Species richness is calculated at $q = 0$, which reveals that the swamp contains noticeably fewer species than the surrounding areas (${}^{0}\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.4a). Averaging across subcommunities, that is ${}^{0}\bar{\alpha}_{\text{swamp}}^{\mathbf{I}} = 37.04$ in the swamp compared to ${}^{0}\bar{\alpha}_{\text{forest}}^{\mathbf{I}} = 54.43$ across the surrounding forest. However, at higher values of q, where ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ is less sensitive to species of low abundance, the effective number of species within the swamp is comparable to that of the surrounding area (${}^{1}\bar{\alpha}_j^{\mathbf{I}}$ to ${}^{\infty}\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.4b - 3.4d), where the effective number of species drops to ${}^{1}\bar{\alpha}_{\text{swamp}}^{\mathbf{I}} = 24.33$, ${}^{2}\bar{\alpha}_{\text{swamp}}^{\mathbf{I}} = 16.39$, and ${}^{\infty}\bar{\alpha}_{\text{swamp}}^{\mathbf{I}} = 6.62$ for the swamp, and ${}^{1}\bar{\alpha}_{\text{forest}}^{\mathbf{I}} = 28.00$, ${}^{2}\bar{\alpha}_{\text{forest}}^{\mathbf{I}} = 15.78$, and ${}^{\infty}\bar{\alpha}_{\text{forest}}^{\mathbf{I}} = 6.03$ for the surrounding forest. From these results, it is clear that the swamp has a much more even distribution of species than the surrounding forest, since the drop in the effective number of species as q increases is much less. In the upper-left corner of the study site there is another region of interest (see Figure 3.3d). Here, as q increases ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ drops quite considerably, indicating that this area contains a skewed distribution of species.

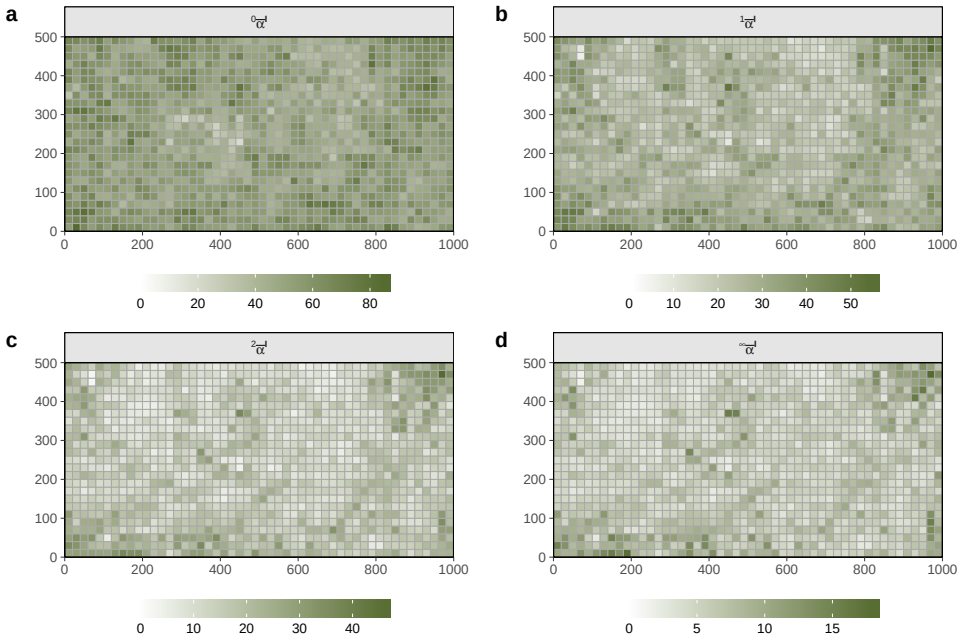


Figure 3.4: Heatmap showing the naïve-type spatial ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ diversity (normalised) of the Barro Colorado Island Forest dynamics plot during the 1981/82 census: Figures are coloured according to the the inherent biodiversity of each 20 m × 20 m quadrat across the 50 ha site, showing: (a) ${}^{0}\bar{\alpha}_j^{\mathbf{I}}$, (b) ${}^{1}\bar{\alpha}_j^{\mathbf{I}}$, (c) ${}^{2}\bar{\alpha}_j^{\mathbf{I}}$, and (d) ${}^{\infty}\bar{\alpha}_j^{\mathbf{I}}$.

Raw subcommunity alpha, ${}^{q}\alpha_j^{\mathbf{I}}$

Raw subcommunity alpha diversity is high when the subcommunity is diverse (in the $\bar{\alpha}$ sense), but also when it is only a small fraction of the metacommunity. The high values of ${}^{q}\alpha_j^{\mathbf{I}}$ in the swamp (${}^{q}\alpha_j^{\mathbf{I}}$, Figure 3.5) is therefore not surprising, since the swamp contains considerably fewer trees than anywhere else (tree count, Figure 3.2), yet a comparable number of species (${}^{q}\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.4).

Subcommunity gamma, ${}^{q}\gamma_j^{\mathbf{I}}$

Subcommunity gamma diversity calculates the average contribution of each individual in a subcommunity toward the diversity of the metacommunity. This is high when a subcommunity contains species that are rare across the metacommunity as a whole. At $q = 0$, four quadrats are identified as contributing most strongly to the diversity of the metacommunity (${}^{0}\gamma_j^{\mathbf{I}}$, highlighted in Figure 3.6). To identify which species are responsible for these high values of ${}^{0}\gamma_j^{\mathbf{I}}$, the weighted species-level components of gamma diversity ($P_{ij}g_{ij}^{\mathbf{I}}$) are calculated for each of these subcommunities (Figure 3.7). These results show that the high values of ${}^{0}\gamma_j^{\mathbf{I}}$ found at grid coordinates (350, 270), (270, 290), (450, 290), and (130, 350) are most strongly influenced by *Maclura Tinctoria*, *Chimarrhis Parviflora*, *Ficus Maxima*, and *Pavonia Dasypetala*, respectively. Then unsurprisingly, examining the spatial distribution of each of these species highlights their rarity across the BCI study site as a whole (Figure 3.8, where crosses indicate the original subcommunities).

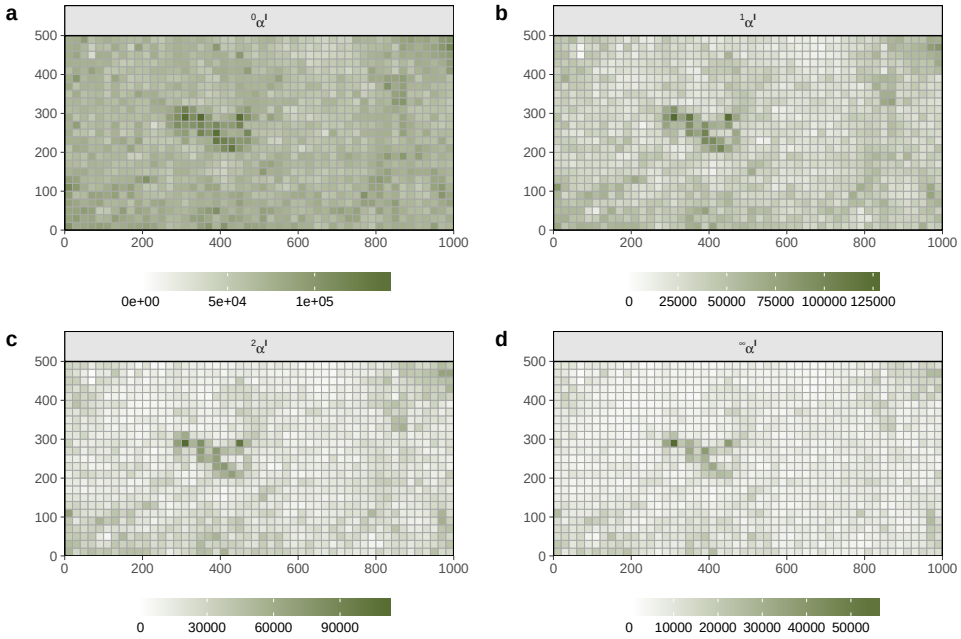


Figure 3.5: Heatmap showing the naïve-type spatial ${}^{q}\alpha_j^{\mathbf{I}}$ diversity (raw) of the Barro Colorado Island Forest dynamics plot during the 1981/82 census: Figures are coloured according to the estimate of naïve-community metacommunity diversity of each 20 m × 20 m quadrat across the 50 ha site, showing: (a) ${}^{0}\alpha_j^{\mathbf{I}}$, (b) ${}^{1}\alpha_j^{\mathbf{I}}$, (c) ${}^{2}\alpha_j^{\mathbf{I}}$, and (d) ${}^{\infty}\alpha_j^{\mathbf{I}}$.

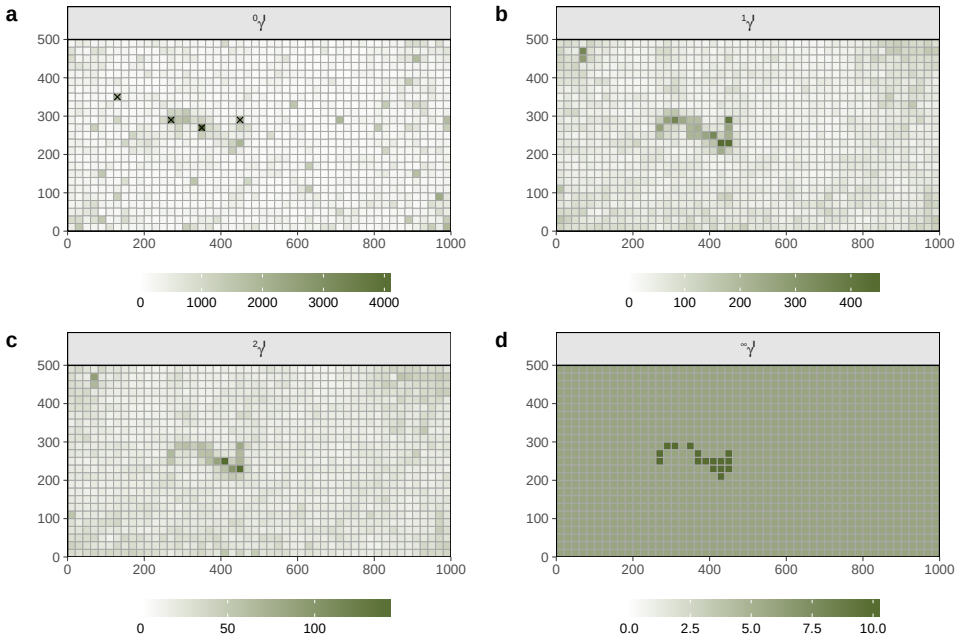


Figure 3.6: Heat map showing the naïve-type spatial ${}^{q}\gamma_j^{\mathbf{I}}$ of the Barro Colorado Island Forest dynamics plot during the 1981/82 census: Figures are coloured according to the subcommunity contribution toward metacommunity diversity of each 20 m × 20 m quadrat across the 50 ha site, showing: (a) ${}^{0}\gamma_j^{\mathbf{I}}$, (b) ${}^{1}\gamma_j^{\mathbf{I}}$, (c) ${}^{2}\gamma_j^{\mathbf{I}}$, and (d) ${}^{\infty}\gamma_j^{\mathbf{I}}$.

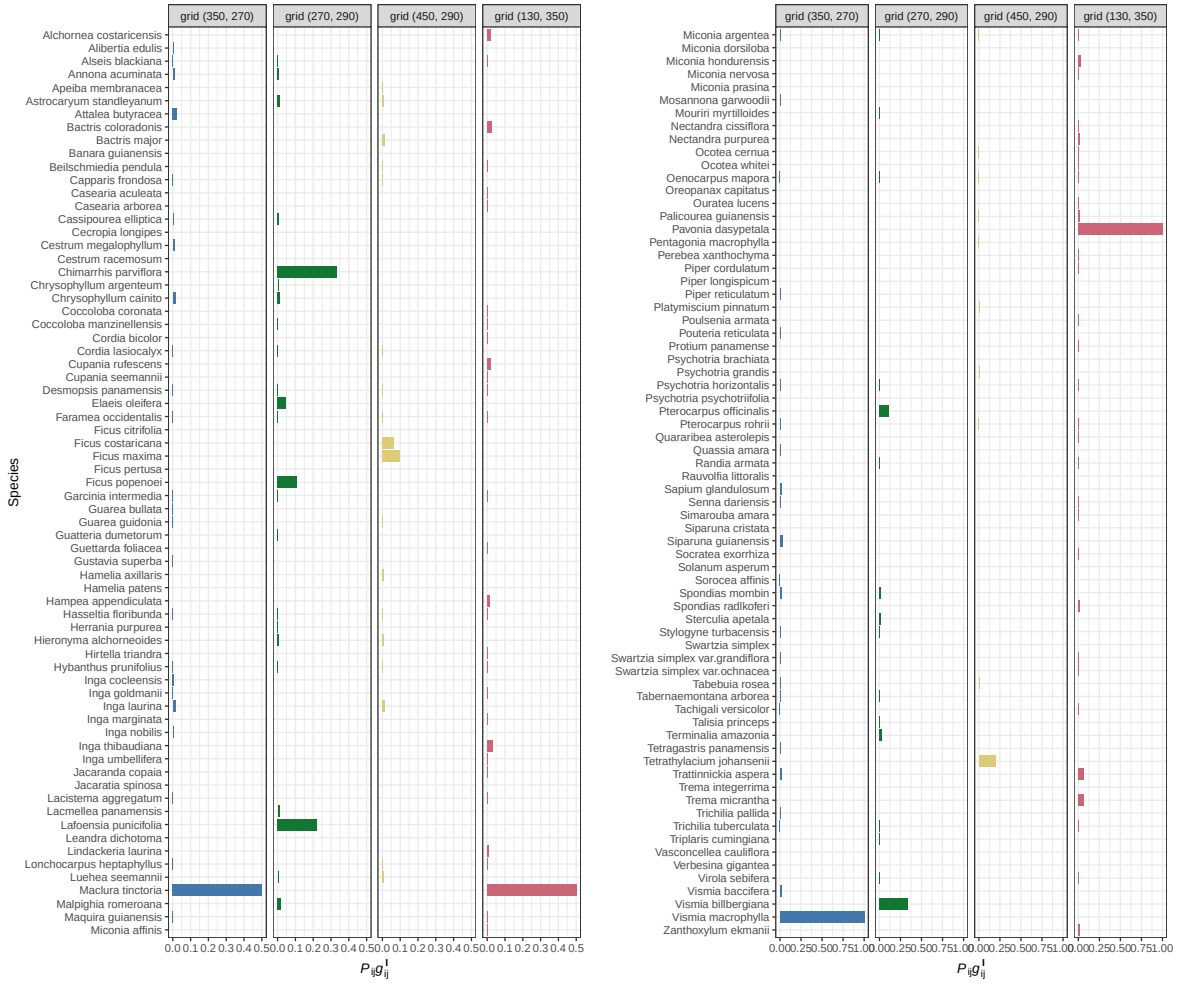


Figure 3.7: Histogram of $P_{ij}g^{\mathbf{I}}_{ij}$ (the weighted species-level component of gamma diversity) of each species in the four subcommunities that contribute most to the diversity of the Barro Colorado Island Forest dynamics plot (highest ${}^{0}\gamma^{\mathbf{I}}_{j}$): These subcommunities are highlighted (with crosses) in Figure 3.6a.

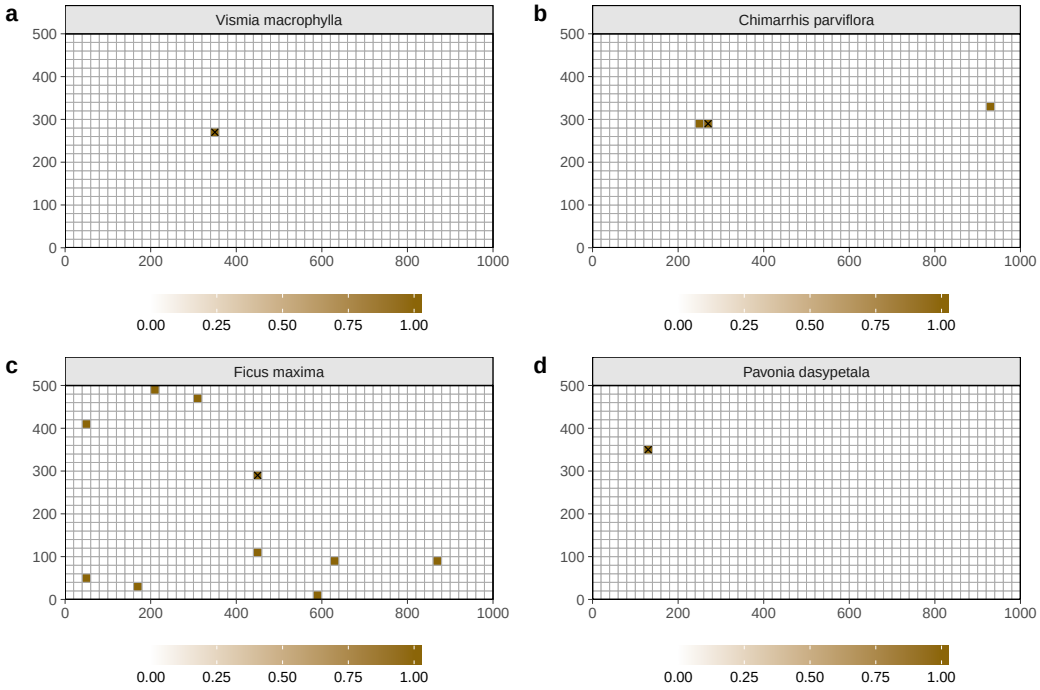


Figure 3.8: Heat map showing the distribution of (a) *Chimarrhis Parviflora*, (b) *Lafoensia Punicifolia*, (c) *Maclura Tinctoria*, and (d) *Pavonia Dasypetala* across the Barro Colorado Island Forest dynamics plot: Figures are coloured according to $P_{ij}g^{\mathbf{I}}_{ij}$ within each 20 m × 20 m quadrat across the 50 ha site. The crosses highlight the location of each of the five subcommunities with the highest ${}^{0}\gamma^{\mathbf{I}}_{j}$ values, whilst each species distribution corresponds to the species within those subcommnuities with the highest $P_{ij}g^{\mathbf{I}}_{ij}$ values.

At $q = 1$, when ${}^{q}\gamma^{\mathbf{I}}_{j}$ is weighted exactly by the relative abundance of each species in each subcommunity, both the swamp and the two quadrats in the top-left corner are identified as being of interest (${}^{1}\gamma^{\mathbf{I}}_{j}$, Figure 3.6b). Though sparsely populated (tree count, Figure 3.2) and comprising a lower species richness than the surrounding areas (${}^{0}\bar{\alpha}^{\mathbf{I}}_{j}$, Figure 3.4a), the swamp contains a greater proportion of species that don't exist in the surrounding metacommunity compared to the rest of the BCI forest plot, and therefore contributes strongly to the diversity of the study site. Likewise, the top-left corner, whose tree-abundance (tree count, Figures 3.2) and species richness (${}^{0}\bar{\alpha}^{\mathbf{I}}_{j}$, Figure 3.4a) is comparable to the surrounding area, contains species that are rare across the metacommunity as a whole, and therefore contributes strongly to metacommunity diversity. At $q = 2$, ${}^{q}\gamma^{\mathbf{I}}_{j}$ is the inverse of the encounter rate between conspecifics in the subcommunity and the metacommunity (see also, Section 1.2.4). Therefore, higher values of ${}^{2}\gamma^{\mathbf{I}}_{j}$ in the swamp (and top-left corner) means that when one individual is sampled from the swamp, it is unlikely to be of the same species than another individual sampled from elsewhere in the metacommunity (${}^{2}\gamma^{\mathbf{I}}_{j}$, Figure 3.6c).

As q increases, ${}^{q}\gamma^{\mathbf{I}}_{j}$ becomes more conservative, as species that are common to the metacommunity (that contribute a much smaller fraction to the diversity of the metacommunity) are weighted more strongly. At $q = \infty$, the swamp contrasts strongly against the rest of the study site (${}^{\infty}\gamma^{\mathbf{I}}_{j}$, Figure 3.6d). These subcommunities are referred to as *swamp subcommunities* in subsequent analyses. Examining the relevant species-level components ($\arg\min_i g_{ij}$) from each

subcommunity reveals that *Bactris major* and *Hybanthus prunifolius* contribute least to the diversity of the study site in the swamp and surrounding forest, respectively. Unsurprisingly then, each species is well distributed across their respective areas (Figure 3.9). Interestingly, the opposite is true for ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ (Figure 3.4), where as q increases the swamp boundary becomes more difficult to discern. This is because ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ is determined from the relative abundance of the species within each subcommunity in isolation (and therefore ${}^{\infty}\bar{\alpha}_j^{\mathbf{I}}$ considers only the most common species in subcommunity j), whilst ${}^{q}\gamma_j^{\mathbf{I}}$ takes into account the relative abundance of species across the metacommunity as a whole (such that ${}^{\infty}\gamma_j^{\mathbf{I}}$ only considers the species in subcommunity j that is most common in the metacommunity).

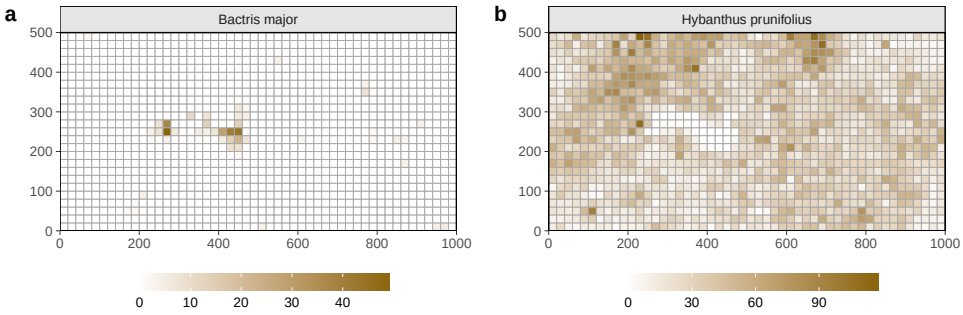


Figure 3.9: Heat map showing the distribution of (a) *Bactris Major* and (b) *Hybanthus Prunifolius* across the Barro Colorado Island Forest dynamics plot: Figures are coloured according to the number of individuals within each 20 m × 20 m quadrat across the 50 ha site.

Representativeness (${}^{q}\bar{\rho}_j^{\mathbf{I}}$)

Normalised subcommunity rho diversity is a measure of representativeness, or of how much the proportional abundance of species within each subcommunity represents the distribution of those same species across the metacommunity as a whole (Figure 3.10). At $q = 0$, ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ is calculated as the arithmetic mean of species representativeness ($\bar{r}_{ij}^{\mathbf{I}}$), which is equal to the proportion of individuals in the metacommunity that have any representation in the subcommunity j at all. Figure 3.10a shows that the composition of species within the swamp is unrepresentative of the composition of species across the wider metacommunity. On average, 34% of the metacommunity is represented by the swamp (${}^{0}\bar{\rho}_{\text{swamp}}^{\mathbf{I}} = 0.340$), whereas 56% is represented by the area in the top-left corner (${}^{0}\bar{\rho}_{\text{top left}}^{\mathbf{I}} = 0.561$).

At $q = 1$, representativeness decreases such that on average only 9% of the metacommunity is represented by the swamp and 2% is represented by the top-left quadrats (Figure 3.10b). Since ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ is low when species within a subcommunity are rare or absent across the rest of the metacommunity, it is reasonable to say that the species contained within the swamp (and the top-left corner) exist in much lower abundances throughout the rest of the study site. At $q = 2$, the swamp is even less representative of the surrounding forest and is therefore more difficult to delineate (Figure 3.10c), and at $q = \infty$ it is no longer possible to distinguish the swamp from the surrounding area (Figure 3.10d); likewise, the area in the top-left corner. At this value of q, the least representative species in the metacommunity is identified at the

grid cell with centre (510, 190) as *Vismia Macrophylla* (Figure 3.11a). The usefulness of this measure decreases in subcommunities with higher representativeness, for example the least representative species in the most representative subcommunity turns out to be *Coccoloba Manzinellensis* in grid (350, 270), which might be of minimal interest (Figure 3.11b).

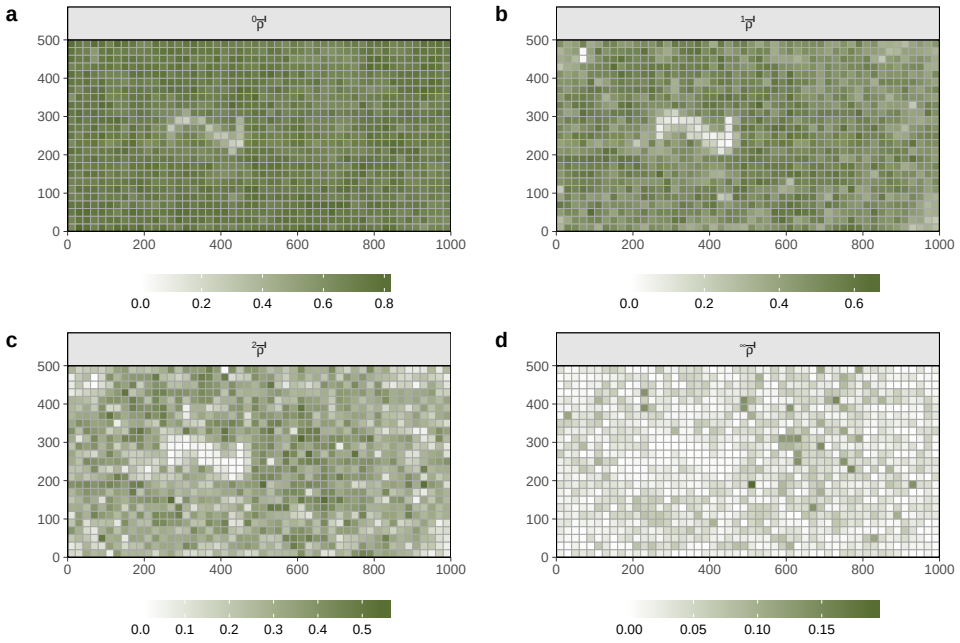


Figure 3.10: Heat map showing the naïve-type spatial ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ of the Barro Colorado Island Forest dynamics plot during the 1981/82 census: Figures are coloured according to the representativeness of species within each 20 m × 20 m quadrat across the 50 ha site, showing: (a) ${}^{0}\bar{\rho}_j^{\mathbf{I}}$, (b) ${}^{1}\bar{\rho}_j^{\mathbf{I}}$, (c) ${}^{2}\bar{\rho}_j^{\mathbf{I}}$, and (d) ${}^{\infty}\bar{\rho}_j^{\mathbf{I}}$.

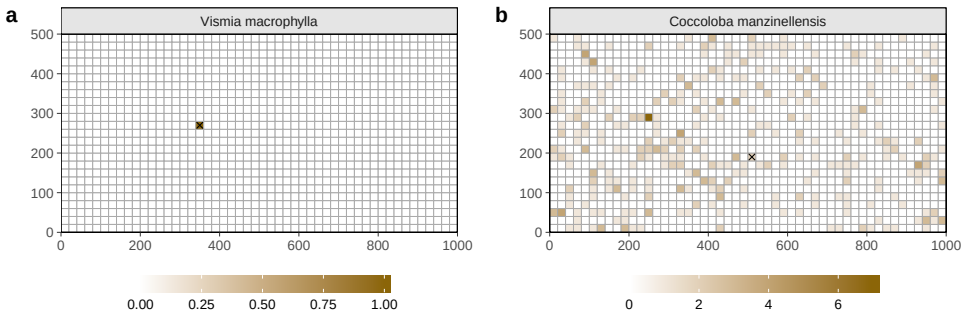


Figure 3.11: Heat map showing the distribution of the least representative species at $q = \infty$, from the least representative and most representative subcommunities across the Barro Colorado Island Forest dynamics plot: (a) *Vismia Macrophylla* and (b) *Coccoloba Manzinellensis*. Crosses highlight the postition of (a) the most representative subcommunity, at grid coordinates (510, 190), and (b) the least representative subcommuntiy, at (350, 270). Figures are coloured according to the number of individuals within each 20 m × 20 m quadrat across the 50 ha site, which in (a) is either 0 or 1.

Redundancy (${}^{q}\rho_j^{\mathbf{I}}$)

Raw subcommunity rho diversity describes the redundancy of species within a subcommunity (Figure 3.12) – a measure of the capacity of the metacommunity to replicate the distribution of species within each subcommunity. As q increases, ${}^{q}\rho_j^{\mathbf{I}}$ becomes less conservative as focus shifts toward species of high distinctiveness (since ${}^{q}\rho_j^{\mathbf{I}} = {}^{1}/_{{}^{q}\beta_j^{\mathbf{I}}}$).

At $q = 0$, when ${}^{q}\rho_j^{\mathbf{I}}$ is calculated as the arithmetic mean of species redundancies ($r_{ij}^{\mathbf{I}}$), the most redundant subcommunities are found near the swamp (Figure 3.12a), indicating that these subcommunities contain species that are common throughout the rest of the metacommunity. Given that the swamp is also primarily responsible for the greatest contribution to the diversity of the study site (Figure 3.6), it may be inferred that the areas of high redundancy in the swamp must be due to the presence of a low abundance of species that are common to the surrounding BCI forest plot. This is confirmed at $q = 1$, where the swamp (and the two quadrats in the top-left corner of the study site) appears comparatively less redundant than the surrounding forest (Figure 3.12b). At this value of q less weighting is given to the most redundant species in each subcommunity (species-level components are weighted exactly by their subcommunity-abundance). Likewise for $q = 2$ (Figure 3.12c).

At $q = \infty$, only the least redundant (or most distinct) species in each subcommunity is captured by the weighting (Figure 3.12d). That is, species that are common within the subcommunity, but rare across the metacommunity as a whole. To illustrate this, Figure 3.13 maps the spatial distribution of the least redundant species in the subcommunity with the highest ${}^{\infty}\rho_j^{\mathbf{I}}$, and 1 of the 22 subcommunities with the lowest ${}^{\infty}\rho_j^{\mathbf{I}}$, which each have a redundancy of one.

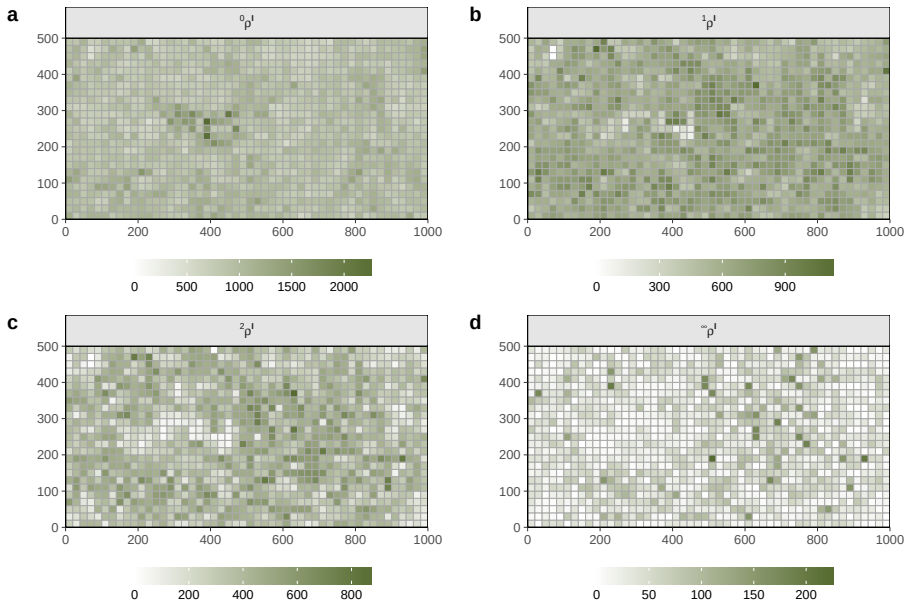


Figure 3.12: Heat map showing the naïve-type spatial ${}^{q}\rho_j^{\mathbf{I}}$ of the Barro Colorado Island Forest dynamics plot during the 1981/82 census: Figures are coloured according to the redundancy of species within each 20 m × 20 m quadrat across the 50 ha site, showing: (a) ${}^{0}\rho_j^{\mathbf{I}}$, (b) ${}^{1}\rho_j^{\mathbf{I}}$, (c) ${}^{2}\rho_j^{\mathbf{I}}$, and (d) ${}^{\infty}\rho_j^{\mathbf{I}}$.

The largest value of ${}^{\infty}\rho_j^{\mathbf{I}}$ is measured in the quadrat located at grid coordinates (510, 190), corresponding to the species-level redundancy of *Coccoloba Manzinellensis* (Figure 3.13a). Given that *Coccoloba Manzinellensis* is quite well distributed across the study site, grid (510, 190) must not contain any particularly rare species. Conversely, the lowest value of ${}^{\infty}\rho_j^{\mathbf{I}}$ is measured in grid (990, 290), which equals the species-level redundancy of *Annona Hayesii*, a species unique to this subcommunity (Figure 3.13b).

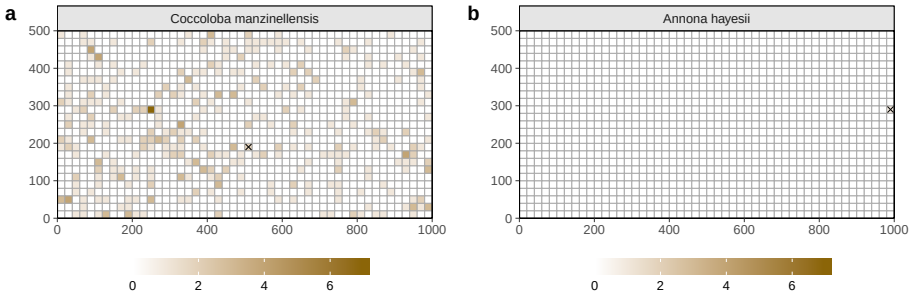


Figure 3.13: Heat map showing the distribution of (a) *Coccoloba Manzinellensis* and (b) *Annona Hayesii* across the Barro Colorado Island Forest dynamics plot: Figures are coloured according to the number of individuals within each 20 m × 20 m quadrat across the 50 ha site. The crosses highlight the location of the subcommunity with the (a) highest and (b) one of the joint equal lowest values of ${}^{\infty}\rho_j^{\mathbf{I}}$, where ${}^{\infty}\rho_j^{\mathbf{I}} = r_{ij}^{\mathbf{I}}$ for *Coccoloba manzinellensis* in grid coordinates (510, 190) and *Annona hayesii* in grid (990, 290), respectively.

The effective number of subcommunities (${}^{q}\bar{\beta}_j^{\mathbf{I}}$)

Subcommunity ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ – the inverse of ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ – is an estimate of the effective number of distinct subcommunities in the metacommunity (Figure 3.14). This value is high when the average representativeness of species within a subcommunity is low. At $q = 0$, only the swamp is identified as having high ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ compared to the rest of the study site, with a maximum effective number of 4.778 distinct subcommunities and an average of 3.095, compared to 1.401 in the surrounding forest (Figure 3.14a). Since the estimate of the effective number of distinct subcommunities like those in the swamp is relatively high, the species within these subcommunities must be, on average, quite distinct. At $q = 1$ and $q = 2$, the upper-left corner is also identified as having a high value of ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ (Figures 3.14b & 3.14c). Since as q increases, ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ weights less representative species more strongly, the area in the top-left corner must contain a higher abundance of rare species than the swamp.

At $q = \infty$, ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ reaches its *least* conservative value, which for each subcommunity, is calculated as the highest species-level estimate of the effective number of subcommunities. In other words, the least representative species in each subcommunity (species with high subcommunity abundances but low metacommunity abundances). The four subcommunities with the highest values of ${}^{\infty}\bar{\rho}_j^{\mathbf{I}}$ are highlighted in Figure 3.14d and as expected, the spatial distribution of each species is unrepresentative of the BCI study site as a whole, since they each only exist in a single quadrat (Figure 3.15).

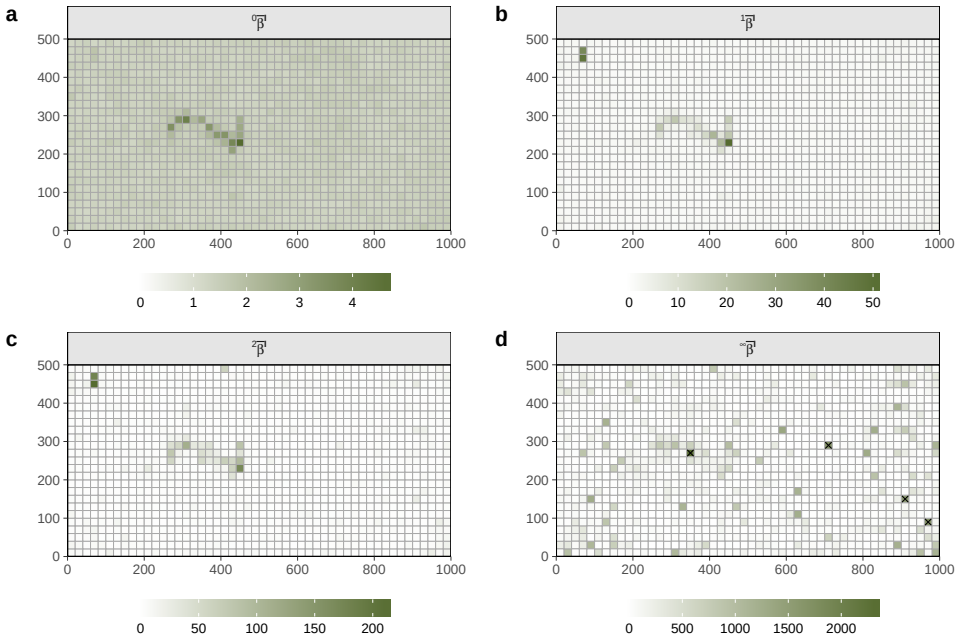


Figure 3.14: Heat map showing the naïve-type spatial ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ of the Barro Colorado Island Forest dynamics plot during the 1981/82 census: Figures are coloured according to the effective number of subcommunities in each 20 m × 20 m quadrat across the 50 ha site, showing: (a) ${}^{0}\bar{\beta}_j^{\mathbf{I}}$, (b) ${}^{1}\bar{\beta}_j^{\mathbf{I}}$, (c) ${}^{2}\bar{\beta}_j^{\mathbf{I}}$, and (d) ${}^{\infty}\bar{\beta}_j^{\mathbf{I}}$.

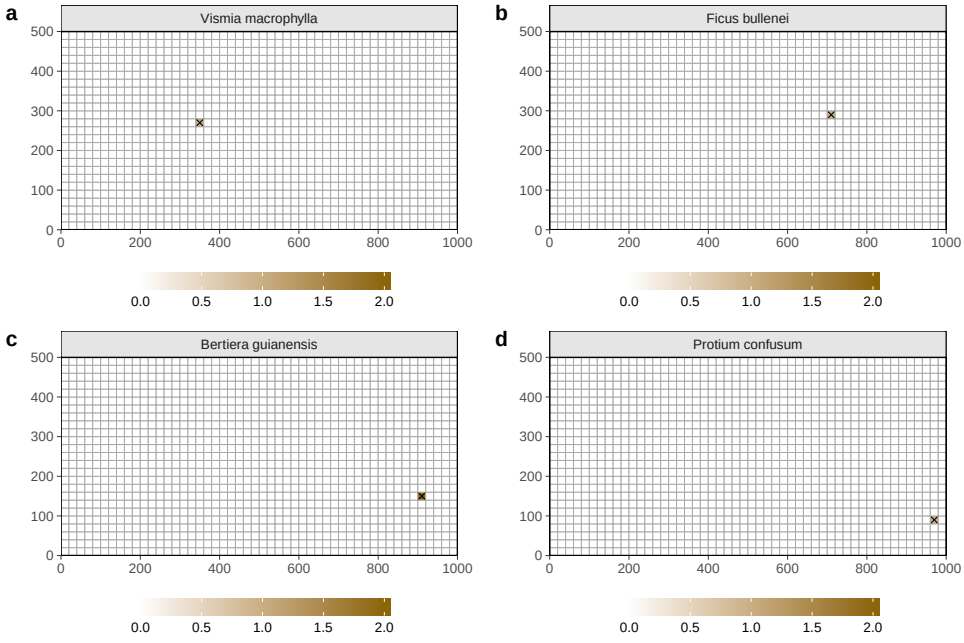


Figure 3.15: Heat map showing the distribution of (a) *Vismia macrophylla*, (b) *Ficus bullenei*, (c) *Bertiera guianensis*, and (c) *Protium confusum*, and across the Barro Colorado Island Forest dynamics plot: Figures are coloured according to the number of individuals within each 20 m × 20 m quadrat across the 50 ha site. The crosses highlight the locations of the four subcommunities with the highest ${}^{\infty}\bar{\beta}_j^{\mathbf{I}}$ values.

Distinctiveness (${}^{q}\beta_j^{\mathbf{I}}$)

Subcommunity ${}^{q}\beta_j^{\mathbf{I}}$ – the inverse of ${}^{q}\rho_j^{\mathbf{I}}$ – describes the average distinctiveness of species in each subcommunity relative to the metacommunity (Figure 3.16), which is high when ${}^{q}\rho_j^{\mathbf{I}}$ is low. That is, when a subcommunity contains a large number of individuals whose species have a low metacommunity-abundance.

Overall, ${}^{q}\beta_j^{\mathbf{I}}$ is highest in the top-left corner of the study site, and low in the swamp and surrounding areas. The distinctiveness of the swamp is low despite the low representativeness (Figure 3.10) due to the relatively low abundance of trees in this area (Figure 3.2). Whereas the distinctiveness of the surrounding areas is low because few rare species reside there. As q increases, focus is increasingly placed on species of high redundancy (species with low subcommunity-abundance and high metacommunity-abundance). At $q = \infty$, only the most redundant species in each subcommunity is considered. The darkest quadrats therefore, are weighted only by their most redundant species.

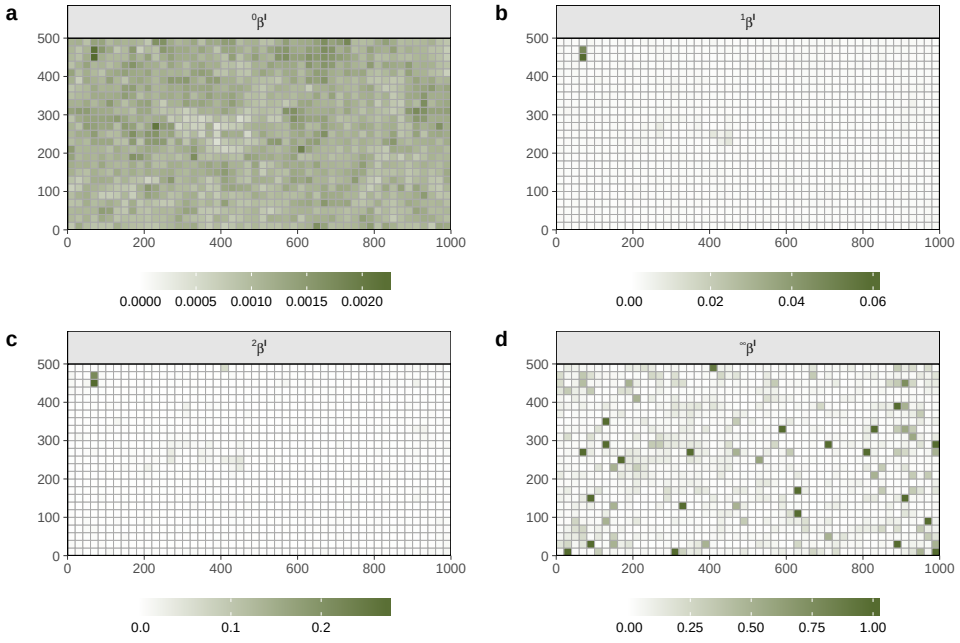


Figure 3.16: Heat map showing the naïve-type spatial ${}^{q}\beta_j^{\mathbf{I}}$ diversity of the Barro Colorado Island Forest dynamics plot during the 1981/82 census: Figures are coloured according to the distinctiveness of species within each 20 m × 20 m quadrat across the 50 ha site: (a) ${}^{0}\beta_j^{\mathbf{I}}$, (b) ${}^{1}\beta_j^{\mathbf{I}}$, (c) ${}^{2}\beta_j^{\mathbf{I}}$, and (d) ${}^{\infty}\beta_j^{\mathbf{I}}$.

3.2.3.3 Taxonomic diversity

In contrast to the naïve-type case, where species are considered completely distinct from one another, taxonomic diversity measures can be used to incorporate information about the taxonomic relatedness of species. Forest ecosystems are often analysed using taxonomic diversity measures (Desrochers & Anand, 2003; Moreno et al., 2009; Culmsee & Leuschner, 2013) and so here, spatial diversity is assessed using transformations of Shimatani's (2001) parameters of taxonomic distance at $q = 1$.

Results show that the maximum effective number of species in a subcommunity drops by half, from ${}^1\bar{\alpha}_j^{\mathbf{I}} = 56.690$ in the naïve-type case (Figure 3.17a) to ${}^1\bar{\alpha}_j^{\mathbf{Z}_{tax}} = 24.926$ when taxonomic similarity is included (Figure 3.17b). In both cases, the quadrats in the top-left corner are consistently less diverse than the rest of the study site. The maximum contribution of a subcommunity to overall metacommunity diversity drops by 70% from ${}^1\gamma_j^{\mathbf{I}} = 441.150$ to ${}^1\gamma_j^{\mathbf{Z}_{tax}} = 60.664$ (Figures 3.17c & 3.17d). However, in both cases, the swamp is identified as contributing highly to the diversity of the study site. This is because the swamp contains a (taxonomically) distinct set of indicator species that are only found in this continuously wet part of the study site. Conversely, the quadrats in the top-left area identified as distinct in Figure 3.17g, are dominated by *A. panamensis*, a species of the same family (Annonaceae) as many other trees in the plot. This invasive species (Rick Condit, personal communication) is therefore taxonomically similar to the rest of the BCI Forest plot.

The range in values of representativeness increases from ${}^1\bar{\rho}_j^{\mathbf{I}} = 0.019 - 0.669$ to ${}^1\bar{\rho}_j^{\mathbf{Z}_{tax}} = 0.086 - 0.915$ (Figures 3.17e & 3.17f) across the study site. Both the swamp and the area in the top-left corner are much less clearly delineated, suggesting the species within these quadrats are closely related to those in the rest of the study site. The maximum distinctiveness of a subcommunity drops from 6.1% in the naïve-type case (${}^1\beta_j^{\mathbf{I}} = 0.061$) to 0.5% when taxonomic similarity is included (${}^1\beta_j^{\mathbf{Z}_{tax}} = 0.005$) (Figures 3.17g & 3.17h). The top-left quadrats continue to be identified as the most distinct subcommunities. However, the relative difference in distinctiveness between these quadrats and the rest of the study site is reduced as species are shown to share taxonomic similarities.

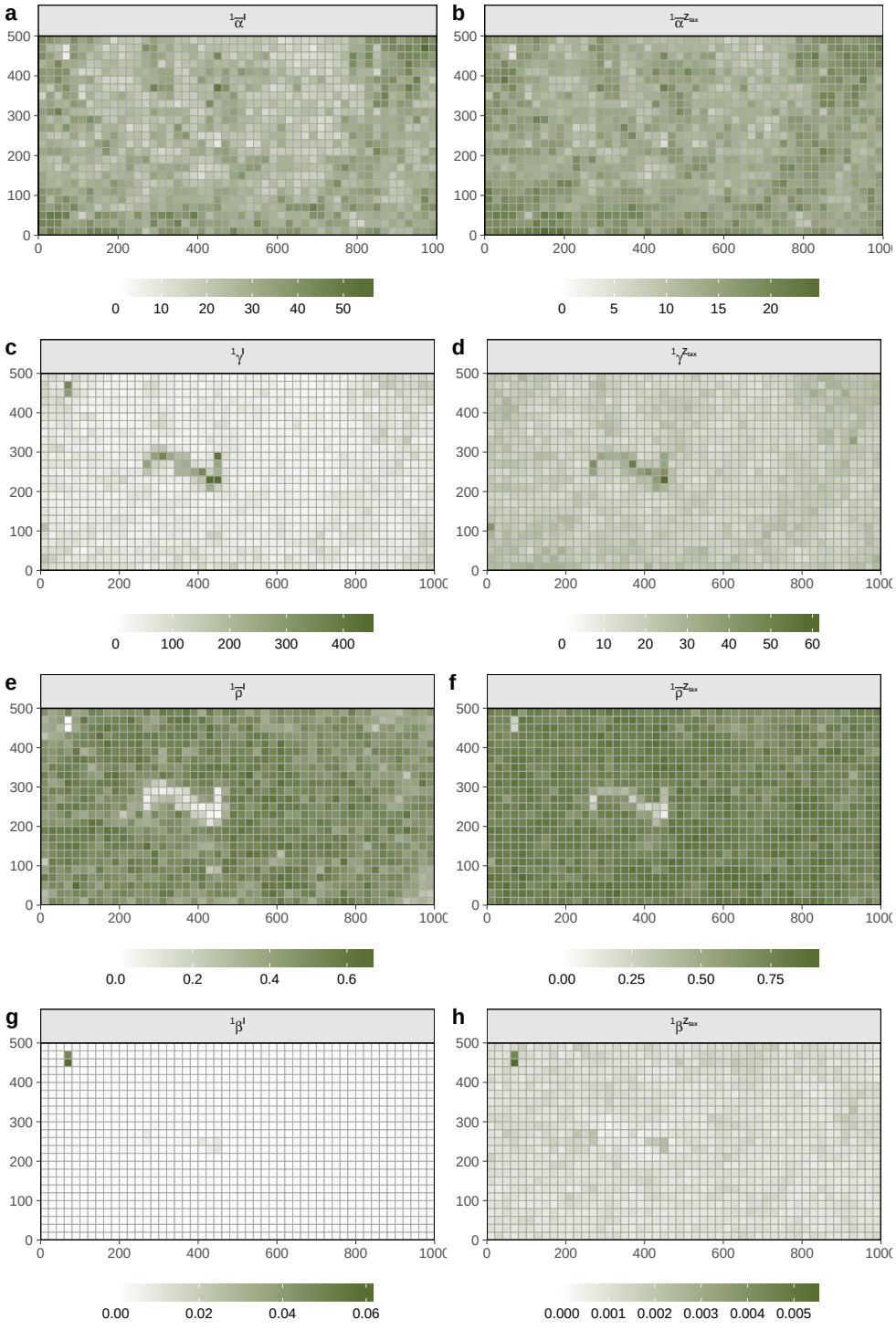


Figure 3.17: Heat map comparing the naïve-type and taxonomic spatial diversity of the Barro Colorado Island Forest plot 1981/82 census: Figures are coloured according to the subcommunity diversity of each 20 m × 20 m quadrat across the 50 ha site: (a) ${}^1\bar{\alpha}_j^{\mathbf{I}}$ and (b) ${}^1\bar{\alpha}_j^{\mathbf{Z}_{tax}}$ - the effective number of species, (c) ${}^1\gamma_j^{\mathbf{I}}$ and (d) ${}^1\gamma_j^{\mathbf{Z}_{tax}}$ - the contribution to metacommunity diversity, (e) ${}^1\bar{\rho}_j^{\mathbf{I}}$ and (f) ${}^1\bar{\rho}_j^{\mathbf{Z}_{tax}}$ - representativeness, and (g) ${}^1\beta_j^{\mathbf{I}}$ and (h) ${}^1\beta_j^{\mathbf{Z}_{tax}}$ - distinctiveness.

3.2.3.4 Temporal diversity

The Barro Colorado Island Forest dataset contains 8 tree censuses, taken at approximately 5-year intervals from 1981-2015. The temporal dynamics of this data can be observed with exactly the same tools that were used to study spatial structure.

The swamp – previously identified in the centre-left of the plot – has a high temporal beta diversity (${}^{q}B^{\mathbf{I}}$, Figure 3.18b), which is indicative of a high turnover in species composition over time (see Section 2.2.6.4). This can be observed in more detail by comparing the change in spatial representativeness across the 8 censuses. In 1981/82, the swamp is unrepresentative of metacommunity diversity (${}^{q}\bar{\rho}_j^{\mathbf{I}}$, Figure 3.19c) and is relatively sparsely populated (tree count, Figure 3.19g), but by 2015, the swamp is much more representative of the metacommunity (${}^{q}\bar{\rho}_j^{\mathbf{I}}$, Figure 3.19d), and the tree density has increased (tree count, Figure 3.19h). However, despite the gain in representativeness, the swamp area still contributes strongly to metacommunity diversity (${}^{q}\gamma_j^{\mathbf{I}}$, Figure 3.19f) because some of the original species are still present and are not found elsewhere in the plot.

The area in the top left, on the other hand, has low turnover over the study period (${}^{q}B^{\mathbf{I}}$, Figure 3.18) and is unrepresentative of the BCI plot for the whole time series (${}^{q}\bar{\rho}_j^{\mathbf{I}}$, Figures 3.19c & 3.19d). This is the result of the colonizing species (from outside the study area) maintaining its dominance in this small area without spreading further across the plot over the years.

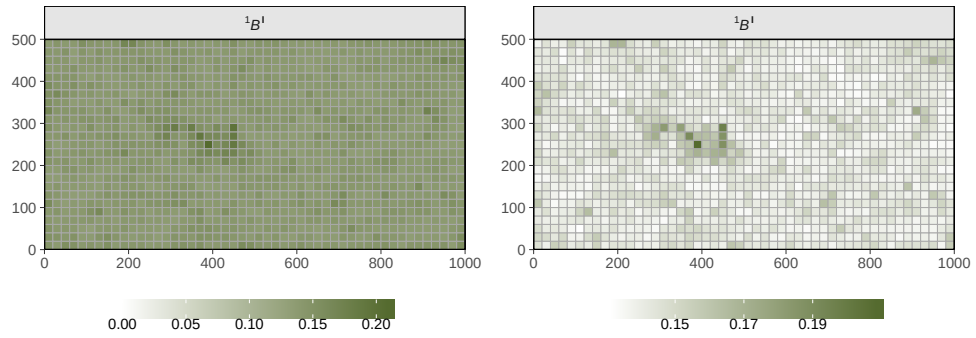


Figure 3.18: Naïve-type temporal beta diversity (metacommunity ${}^{1}B^{\mathbf{I}}$) of the Barro Colorado Island Forest plot: Figures are coloured according to the metacommunity diversity of each 20 m × 20 m quadrat across the 50 ha site between 1981-2015. Values are presented (a) relative to a baseline of 0, and (b) unscaled.

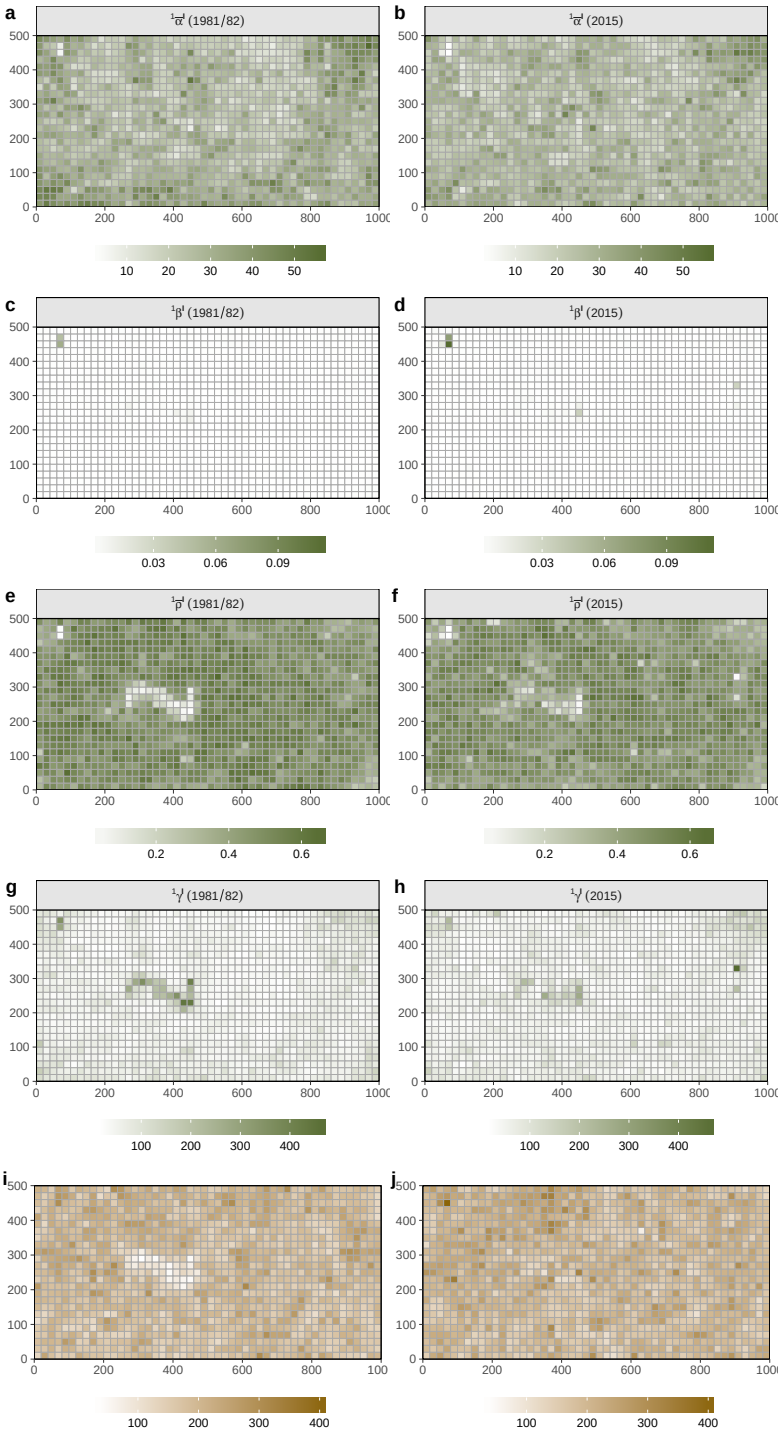


Figure 3.19: Naïve-type spatial diversity of the Barro Colorado Island Forest plot at different time points: Figures are coloured according to the subcommunity diversity of each 20 m × 20 m quadrat: (a) ${}^1\bar{\alpha}^{\mathbf{I}}_{1981/82}$ and (b) ${}^1\bar{\alpha}^{\mathbf{I}}_{2015}$ - effective number of species, (c) ${}^1\bar{\rho}^{\mathbf{I}}_{1981/82}$ and (d) ${}^1\bar{\rho}^{\mathbf{I}}_{2015}$ - representativeness, (e) ${}^1\gamma^{\mathbf{I}}_{1981/82}$ and (f) ${}^1\gamma^{\mathbf{I}}_{2015}$ - contribution to overall metacommunity diversity, and (g) and (h) - tree counts, during the 1981/82 and 2015 censuses, respectively.

3.2.4 Summary

This case study examined the variation in spatial biodiversity within each quadrat using a range of subcommunity measures. After which, the variation in temporal biodiversity was investigated by re-partitioning the data and calculating the metacommunity diversity of each quadrat as a cross section through time. The spatial biodiversity of the BCI Forest dynamics plot (during the first census, 1981-1982) was first investigated using naïve-type diversity measures. These results are summarised in Figure 3.20, which brings together key images from earlier in the chapter. The most basic measure of diversity, the effective number of species, revealed no clear spatial structure (${}^1\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.20b). This remained true at a variety of quadrat sizes (plots not shown) and values of q (${}^q\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.4). However, when beta and gamma diversities are examined, strong signals emerged and two sites – a large area in the centre-left and a small area in the upper-left – were identified as being of particular interest.

Although both areas stand out as being unrepresentative of the BCI forest plot (${}^1\bar{\rho}_j^{\mathbf{I}}$, Figure 3.20e), the swamp had a high tree diversity (${}^1\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.20b) but fewer trees per m^2 than the rest of the plot (tree count, Figure 3.20a), resulting in low distinctiveness (${}^1\beta_j^{\mathbf{I}}$, Figure 3.20f) but high contribution to overall metacommunity diversity (${}^1\gamma_j^{\mathbf{I}}$, Figure 3.20d). The second area in the upper-left had low representativeness (${}^1\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.20b) but high tree abundance per m^2 (tree count, Figure 3.20a), resulting in high distinctiveness (${}^1\beta_j^{\mathbf{I}}$, Figure 3.20f) and also a high overall contribution to metacommunity diversity (${}^1\gamma_j^{\mathbf{I}}$, Figure 3.20d). Some of the difference between the two sites may be observed in terms of contribution to metacommunity diversity at $q = 0$ – which discounts relative abundance and looks at per tree contribution to species richness – which is relatively low in the upper-left quadrats (${}^0\gamma_j^{\mathbf{I}}$, Figure 3.20c). These subcommunity diversity measures therefore help identify two interesting areas of the plot, which are not identifiable from the quadrats in isolation (${}^1\bar{\alpha}_j^{\mathbf{I}}$, Figure 3.20b), and could not even in principle be identified from metacommunity-level diversity measures. These results show that, Reeve et al.'s (2016) subcommunity diversity measures are capable of providing clear and immediate insight into the differences between the sites through simple summary statistics before going into a detailed analysis of the underlying data.

Taxonomic diversity was investigated, using transformations of Shimatani's (2001) parameters of taxonomic distance at $q = 1$. It was found that the contribution to metacommunity diversity remained high in the swamp, but not the area in the top-left (${}^1\gamma_j^{\mathbf{Z}_{tax}}$, Figure 3.17d). An examination of the metadata and species information revealed that the area in the top-left has been colonised by a single tree species (*Anaxagorea Panamensis*) from outside the study site (Rick Condit, personal communication) that has developed a high local density in this single area. This invasive species is taxonomically similar to the rest of the BCI Forest plot whereas, the central area is a swamp with a distinct species composition. These results agree with those obtained by Kembel & Hubbell (2006), who investigate how the phylogentic structure of tree communities (estimated with phylogenetic distance metrics, MPD and MNND) varies among

spatial scales and habitats. They observe that tree communities are more distantly related than expected in swamp and slope habitats[1]. In this current study, the same results were obtained using much simpler methods – by incorporating taxonomic similarity into Reeve et al.'s (2016) framework, however it might be interesting to extend these analyses to examine the dynamics of phylogenetic diversity across space and time.

Temporal dynamics were examined by rearranging the dataset such that a subcommunity is a single quadrat (as before), but a metacommunity extends this subcommunity over 8 tree censuses (from 1981-2015). Results identified a low temporal beta diversity in the top-left quadrats, suggesting a low turnover in species composition over the study period (${}^{q}\bar{B}^{\mathbf{I}}$, Figure 3.18). Furthermore, these quadrats were found to be unrepresentative of the BCI plot for the whole time series (${}^{q}\bar{\rho}_{j}^{\mathbf{I}}$, Figures 3.19c & 3.19d), a result of the colonizing species (from outside the study area) maintaining its dominance in this small area without spreading further across the plot over the years. On the other hand, the swamp was identified as having a high temporal beta diversity (${}^{q}\bar{B}^{\mathbf{I}}$, Figure 3.18) over the study period, indicative of a high turnover in species composition over time. This was confirmed by comparing the spatial diversity at each time point, where an increase in representativeness (${}^{q}\bar{\rho}_{j}^{\mathbf{I}}$, Figures 3.19c & 3.19d) and tree density (${}^{q}\bar{\rho}_{j}^{\mathbf{I}}$, Figures 3.19i & 3.19j) was revealed. Ecologically speaking, an increase in the representativeness of the swamp suggests that species common to the surrounding areas are spreading to these quadrats. This change could be explained by the swamp drying out, but there is no evidence that this is the case (Rick Condit, personal communication). In fact, Condit et al. (2017b) show that prior to 1992, there were multiple extreme dry seasons, after which, census intervals have been wetter than the long-term average. These increased moisture levels coincide with a change in the pattern of recruits (trees absent from the previous census) observed by Kanagaraj et al. (2011). They used multivariate regression trees to group quadrats with similar species composition according to the topographical characteristics of the BCI study site. Interestingly, the swamp was identified as a distinct habitat type in the 1985 and 1990 censuses, after which it homogenises with the low plateau habitat type in the surrounding area. Despite this, the swamp still contributed to metacommunity diversity (${}^{q}\bar{\rho}_{j}^{\mathbf{I}}$, Figures 3.19h), because species are present (albeit in low numbers) that are not found elsewhere in the plot. If the plot represented a larger landscape that was being considered for management, results such as these would help identify unique and diverse parts of the plot for conservation. This "hidden" diversity provided by a few trees of rare species against a background of common species is only demonstrated by Reeve et al.'s (2016) new subcommunity gamma diversity.

[1]They also find that tree communities are more closely related than expected in young forest and plateau habits, which they hypothesise is due to environmental filtering of phylogenetically conserved traits (see also Swenson et al., 2012)

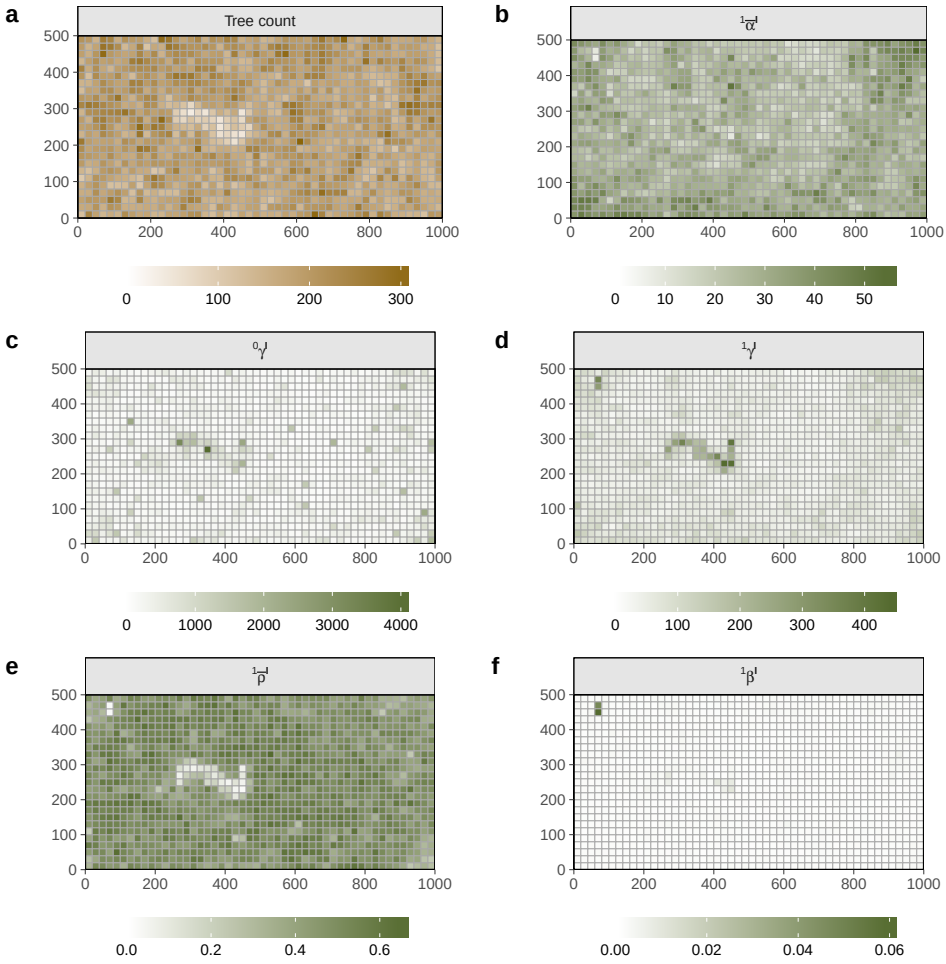


Figure 3.20: Heat map showing the naïve-type spatial diversity of the Barro Colorado Island Forest dynamics plot 1981/82 census: Figures are coloured according to the subcommunity diversity of each 20 m × 20 m quadrat across the 50 ha site, showing: (a) tree counts, (b) ${}^{1}\bar{\alpha}_j^{\mathbf{I}}$ - the effective number of species, (c) ${}^{0}\gamma_j^{\mathbf{I}}$ and (d) ${}^{1}\gamma_j^{\mathbf{I}}$ - contribution to overall metacommunity diversity, (e) ${}^{1}\bar{\rho}_j^{\mathbf{I}}$ - representativeness, and (f) ${}^{1}\beta_j^{\mathbf{I}}$ - distinctiveness.

3.3 Case study: Identifying communities with the least representative demographic profiles in the 2001 census of England and Wales

3.3.1 Introduction

In this case study, Reeve et al.'s (2016) framework of diversity measures are used to investigate the underlying demographic structure and variability of the human population of England and Wales during the 2001 census. This dataset was selected as a means to validate the framework, since results could be easily verified. Two different breakdowns of census data are available, each describing the same population from a different viewpoint, differing in both spatial resolution and distribution of age classes. Diversity measures were used to identify which geographical areas were least representative of England and Wales. That is, areas comprising unusual age class distributions compared to England and Wales as a whole. This was achieved by incorporating carefully constructed similarity matrices, based on the natural similarity of age classes. The aim of this work was to identify features of population structure and show that quantitatively similar results could be obtained despite differences in resolution.

Acknowledgements
The 2001 census of England and Wales was obtained from the Office for National Statistics.

3.3.2 Methods

Dataset
Data were obtained from the Office for National Statistics 2001 census of England and Wales (Office for National Statistics, 2001). Two datasets were examined, each comprising the total population of England and Wales, differing only in the way data was partitioned (Table 3.1). These data comprised the age structure of the population in established geographical areas (10370 parishes and 8850 Census Area Statistics (CAS) wards) at different levels of age resolution (16 and 81 age classes, respectively, comprising ages from 0-110 years old).

Defining similarity
Each dataset was defined as a separate metacommunity, comprising N subcommunities, defined naturally within each dataset as distinct parishes or CAS wards. Diversity was calculated over *types* defined as age classes, categorically divided into S classes with individuals ranging from 0-110 years. Pairwise similarity between age classes was calculated as a transformation of distance, $Z_{ii'} = e^{-kd_{ii'}}$ (a standardised age-related similarity metric), where k is a scaling factor and $d_{ii'}$ is the age difference between the mean of age classes i and i'.

Table 3.1: Tabulated summaries of the parish and CAS ward datasets: showing the total number of subcommunities, the size of the largest subcommunity, the total number of individuals, and the total number of age classes (in each dataset).

Dataset	Subcommunity		Metacommunity	
	Total number	Largest population	Total population	Number of age classes
CAS wards	8850	35102	52042026	81
Parishes	10370	969197	52041915	16

Similarity was calibrated by altering the parameter k and fitting each metacommunity to an effective number of age classes (${}^{q}G^{\mathbf{Z}}$, Section 2.2.5). This was necessary to standardise the two datasets in order to sensibly compare results between them, since when $\mathbf{Z} = \mathbf{I}$, each age classes would be considered completely distinct. The problem is made clear in Figure 3.21, which shows the distribution of individuals in each age class. In the naïve-type case, each age class is considered distinct, such that the parish and CAS ward datasets comprise 16 and 81 distinct age classes, respectively. Subsequently, in the parish dataset (Figure 3.21a), there are around 3,400,000 individuals aged 25-29 that are 100% similar to each other, and have no similarity to the 12,000,000 individuals aged 30-44. On the other hand, in the CAS ward dataset (Figure 3.21b), the 620,000 individuals aged 25 have no similarity to the 650,000

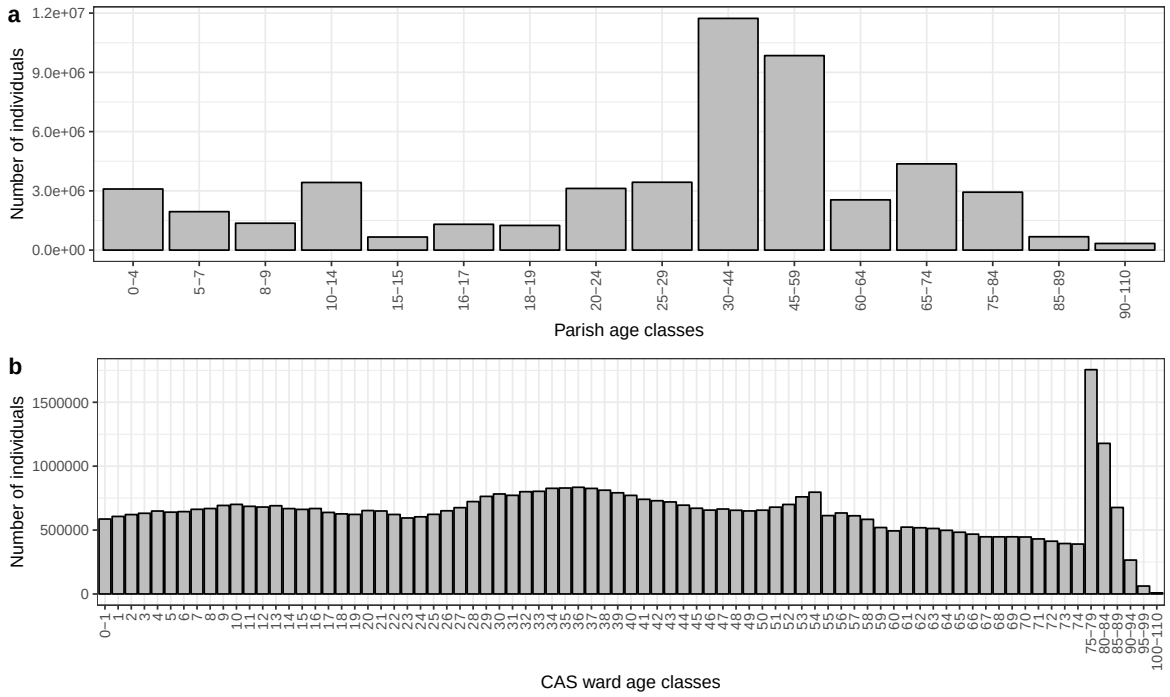


Figure 3.21: Histogram showing the number of individuals in each age class: for the (a) parish and (b) CAS ward datasets. Note that the peak at 75-79 in the CAS ward dataset is due to the amalgamation of individuals aged 75-79 into the same age class.

individuals aged 26, who are in turn completely dissimilar to the 680,000 individuals aged 27, and so on. The diversity of age classes in the parish dataset is therefore not comparable to the diversity of age classes in the CAS ward dataset without some means of standardisation.

To standardise these age classes, metacommunity ${}^{1}G^{\mathbf{Z}}$ was used to calculate the effective number of age classes in each dataset. Recalling that the actual number of age classes in the parish and CAS ward datasets is 16 and 81, respectively (in the naïve-type case), then an effective number ≤ 16 should be selected for standardisation. In the first instance, the parameter k was altered until metacommunity ${}^{1}G^{\mathbf{Z}} \approx 8$ (which was selected as an adequate number of age classes for a human lifetime), yielding $k_{p8} = 0.194$ and $k_{c8} = 0.163$ for the parish and CAS ward datasets, respectively. However ${}^{1}G^{\mathbf{Z}} \approx 4$ was also tested, yielding $k_{p4} = 0.082$ and $k_{c4} = 0.078$. Figure 3.22 shows how varying k determines how similar different age classes are considered to be, at different distances. A lower value of k requires a greater distance for an age class to be considered completely distinct. In the second part of this study, the high number of age classes in the CAS ward dataset was used to compare how results were affected by varying the k parameter. As before, k was calibrated to an effective number of age classes $\in \{2, 4, 8, 16, 32\}$, corresponding to $k_{c2} = 0.031$, $k_{c4} = 0.078$, $k_{c8} = 0.163$, $k_{c16} = 0.330$, and $k_{c32} = 0.728$, respectively.

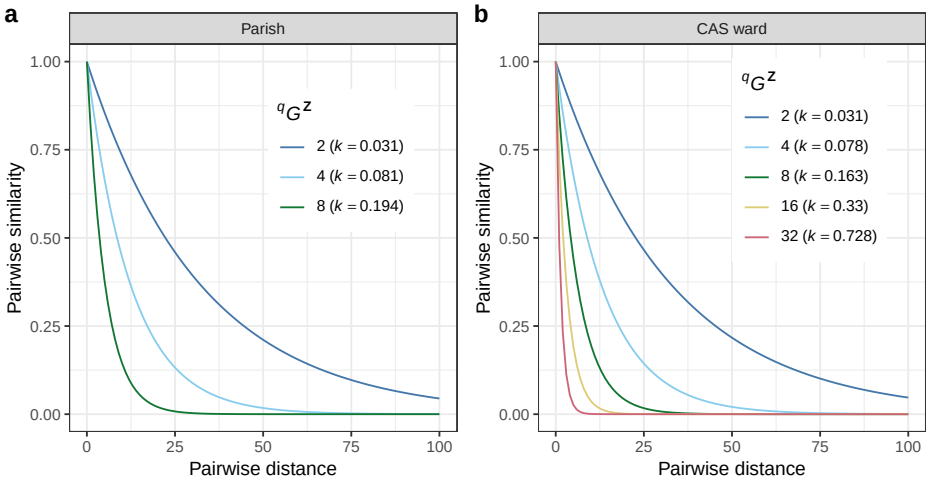


Figure 3.22: Distance versus similarity at different values of k: where distance is defined between age classes (or the median of age classes), and similarity is calculated as $Z_{ii'} = e^{-kd}$ with a scaling parameter k, for the (a) parish and (b) CAS ward datasets. Lower values of k require a greater age difference before age classes are considered completely distinct. The legend includes the effective number of age classes corresponding to each scaling parameter, as described in the methods section.

Calculating diversity

To identify which areas were least representative of England and Wales (in 2001), subcommunity ${}^{1}\bar{\rho}_{j}^{\mathbf{Z}}$ was calculated (Section 2.2.6.1). Representativeness is low when a subcommunity contains a high abundance of species (in this case, age classes) that are rare across the metacommunity as a whole. The least representative areas were identified by ranking each parish (or CAS ward) by ${}^{1}\bar{\rho}_{j}^{\mathbf{Z}}$ in descending order. These results were assessed by plotting the relative proportional abundance of individuals within each age class, thus observing their demographic structure. The relative proportional abundance was calculated by dividing the proportional abundance of each age class by its bin length (the number of ages within that age class), which was then normalised by dividing by the total proportional abundance across England and Wales. In this way, any subcommunity with a distribution exactly proportional to the average of England and Wales (*e.g.* England and Wales itself) would yield a relative proportional abundance equal to one, across all age classes. Therefore, subcommunities comprising age classes of relatively low national abundance (such as the elderly) could be clearly identified.

3.3.3 Results and discussion

Aggregate census data were examined by finding the lowest values of ${}^{1}\bar{\rho}_{j}^{\mathbf{Z}}$, to expose distinct features (subcommunities with particularly unusual age distributions) in the underlying demographic structure of England and Wales (Figure 3.23). Similarity was first calibrated to ${}^{q}G^{\mathbf{Z}} \approx 8$. The 10 least representative parishes of England and Wales are shown in Figure 3.24a and listed in Table 3.2. These include parishes dominated by young adult populations (aged 18-24) such as Keele, Heslington, and Cathays, corresponding to Keele University, the extended campus of the University of York, and Cardiff University, respectively; also of interest are parishes inhabited by young families (adults aged 18-29 and their children, aged

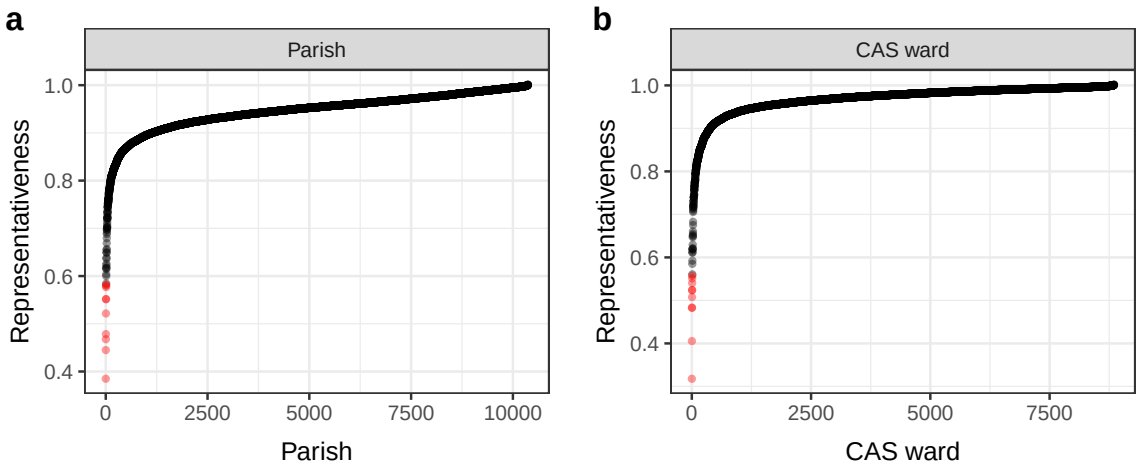


Figure 3.23: The representativeness of each community in the England and Wales 2001 census: Subcommunity ${}^{1}\bar{\rho}_{j}^{\mathbf{Z}}$ was calculated over: (a) 10370 parishes and (b) 8850 CAS wards, where red dots highlight the 10 least representative communities in each assemblage.

<4) such as West Thorney; parishes containing boarding schools populated by teenagers such as Bryanston and Acton Burnell; and parishes occupied by the elderly (aged 75+), such as Tixover and Tabley Inferior.

These analyses were validated with CAS ward data using different age classes but the same underlying age-based similarity measure. Similar results were obtained (Figure 3.24b) and listed in Table 3.2. The 10 least representative CAS wards of England and Wales include those inhabited by young adult populations, such as Holywell, Keele, Elvet, Heslington, Carfax, Headingley, Menai, St Nicholas, and Highcliffe; and CAS wards populated by the elderly, such as Cathays. Keele, Heslington, and Cathays were identified in both the 10 least representative parishes and CAS wards. Of the remaining locations, differences in results are presumed to be caused by population variations within spatial boundaries. However, both analyses were able to detect interesting populations comprising particularly unusual demographies.

Table 3.2: The 10 least representative areas of England and Wales, where similarity is calibrated to $^1G^{\mathbf{Z}} \approx 8$ ($k_8 = 0.1785$): for parishes (left), and CAS wards (right). Features of each parish and CAS ward are given below each location in brackets. The age column highlights age classes with unusually high abundance (*i.e.* with a relative proportional abundance greater than 2, where 1 denotes the relative proportional abundance of England and Wales as a whole).

Parish (Feature)	Age	CAS ward (Feature)	Age
Keele/non-parished area (Keele University)	18-24	Holywell (Oxford University)	18-24
Tixover (Retirement village)	75+	Keele (Keele University)	18-23
Heslington (University of York)	18-24	Elvet (Durham University)	18-24
Newton St Loe (Bath Spa University)	18-24	Heslington (University of York)	18-24
Cathays (Cardiff University)	18-24	Carfax (Oxford University)	18-26
Acton Burnell (Concord College)	15-19	Headingley (Leeds Beckett U., Headingley)	19-25
Bryanston (Bryanston School)	10-19	Menai, Bangor (Bangor University, Ffiddoedd)	18-23
West Thorney (Unknown)	<4, 18-29	St Nicholas (Giles Infant & Primary School)	18-22
Tabley Inferior (Retirement village)	75+	Highcliffe (Retirement village)	67+
Stowe (Unknown)	10-19	Cathays (Cardiff University)	19-25

Calibrating similarity to ${}^{q}G^{\mathbf{Z}} \approx 4$ produced comparable results, with 4 replacements in the parish dataset Figure 3.25a and only 1 in the CASward dataset Figure 3.25b. In the parish dataset, Acton Burnell (15-19), Tabley Inferior (75+), and Stowe (10-19) were replaced with Tarrant Monkton (16-24), Torksey/Brampton/Hardwick (60-84), and South Tedworth (18-29). Note that the values in brackets highlight age classes with unusually high abundance (*i.e.* with a relative proportional abundance greater than 2, where 1 denotes the relative proportional abundance of England and Wales as a whole). Whereas in the CAS ward dataset, Cathays (19-25) was replaced with South Downham (61-94), identifying a new demographic feature. These results show that the CAS ward dataset, with the higher age-class resolution and lower spatial resolution, is more robust to changes in the scaling parameter k.

Extending this further, Figure 3.26 shows the distribution of age classes within the 10 least representative CAS wards, with similarity calibrated to ${}^{q}G^{\mathbf{Z}} \in \{2, 4, 16, 32\}$. As already mentioned, decreasing ${}^{q}G^{\mathbf{Z}}$ from 8 to 4 effective age classes resulted in Cathays (19-25) being replaced with South Downham (61-94). Decreasing again to ${}^{q}G^{\mathbf{Z}} \approx 2$ showed no change. On the other hand, when ${}^{q}G^{\mathbf{Z}}$ was increased from 8 to 16 effective age classes, Highcliffe (67+) was replaced with Market (18-25), and increasing again to ${}^{q}G^{\mathbf{Z}} \approx 32$, results were again unchanged. Overall, 8 (out 10) CAS wards were identified within the 10 least representative subcommunities across all values of k.

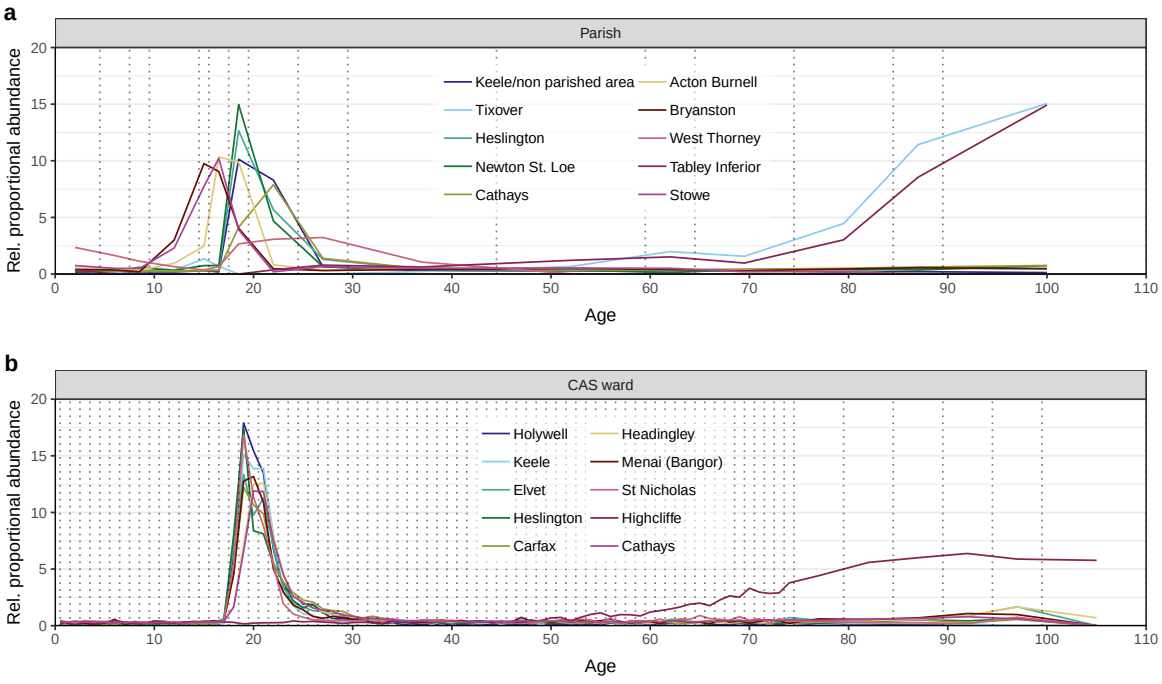


Figure 3.24: The 10 least representative areas of England and Wales, where similarity is calibrated to ${}^{1}G^{\mathbf{Z}} \approx 8$: for (a) parishes and (b) CAS wards. The y-axis is normalised within age classes, by dividing the proportional abundances by the number of ages contained within each class. These values are normalised against the metacommunity, by dividing by the total proportional abundance across England and Wales. Any community with a distribution proportional to the average of England and Wales would have a relative proportional abundance of 1, for all age classes. The vertical dotted lines denote age class cut-offs.

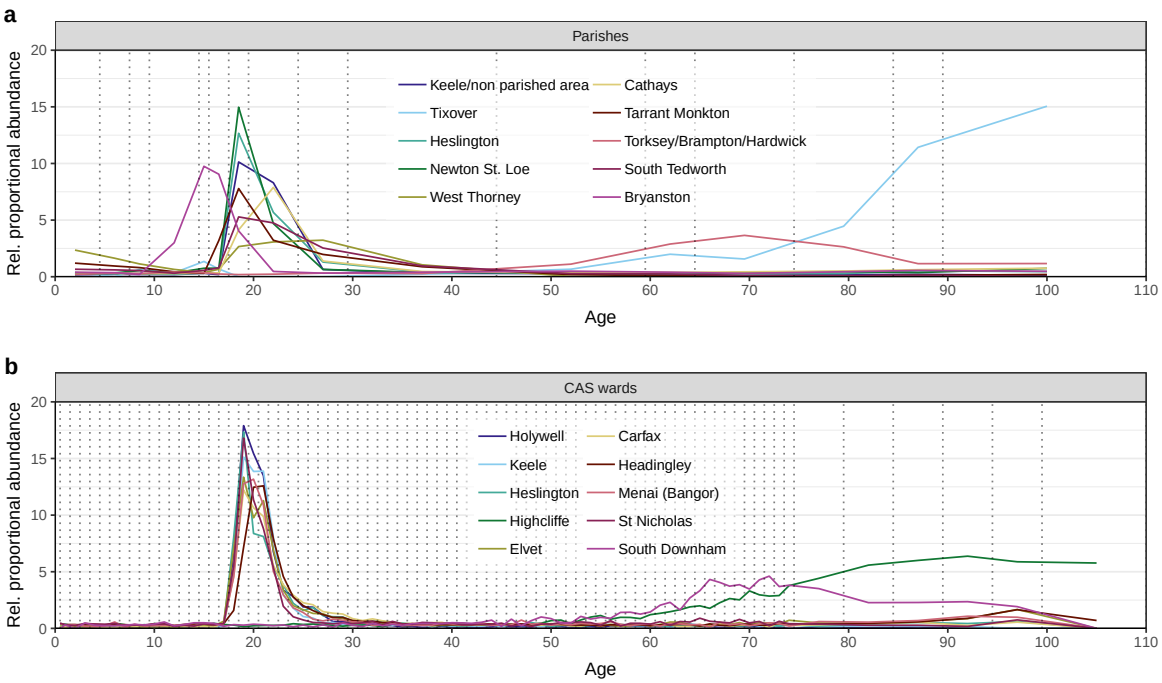


Figure 3.25: The 10 least representative areas of England and Wales, where similarity is calibrated to ${}^{1}G^{\mathbf{Z}} \approx 4$: for (a) parishes and (b) CAS wards. The y-axis is normalised within age classes, by dividing the proportional abundances by the number of ages contained within each class. These values are normalised against the metacommunity, by dividing by the total proportional abundance across England and Wales. Any community with a distribution proportional to the average of England and Wales would have a relative proportional abundance of 1, for all age classes. The vertical dotted lines denote age class cut-offs.

3.3.4 Summary

The focus of this case study was explicitly chosen to illustrate the flexibility and utility of Reeve et al.'s (2016) framework of diversity measures. These measures were used to develop new methods with which to identify communities with the least representative demographic profiles (*i.e.* unusual age class distributions) in the 2001 census of England and Wales. To do this, subcommunity ${}^1\bar{\rho}_j^{\mathbf{Z}}$ was calculated over *types* defined as age classes. Results were compared between two datasets (metacommunities), each comprising the total population of England and Wales: (1) the parish dataset had a higher spatial resolution but lower number of age classes, whereas (2) the CAS ward dataset had a lower spatial resolution but a much higher number of age classes. To allow a useful comparison to be made, the similarity between age-classes was standardised between the two datasets. This was done by varying the parameter k (effectively, scaling the similarity between age classes) until the calculated effective number of age classes, ${}^qG^{\mathbf{Z}} \approx 8$, which was pre-selected as an adequate number of age classes for a human lifetime; ${}^qG^{\mathbf{Z}} \approx 4$ was also tested, as well as 2, 16, and 32, for the CAS ward dataset.

When ${}^qG^{\mathbf{Z}} \approx 8$, Keele, Heslington, and Cathays were identified within the 10 least representative subcommunities of both the parish and CAS ward datasets Figure 3.24. Amongst the remaining communities, the same distinct age profiles were observed. That is, young adults in proximity to universities, and retirement communities dominated by the elderly. When the effective number of age classes was decreased to ≈ 4, results were mostly stable: 6 (out of 10) parishes and 9 (out of 10) CAS wards continued to be identified as being unrepresentative (Figure 3.25). These results identify the CAS ward dataset as being more robust to varying the scaling parameter k (though a new demographic feature was identified: individuals aged 61-94). The CAS ward dataset was then used to investigate to what extent these results depend on k, which was calibrated against an effective number of age classes $\in \{2, 4, 8, 16, 32\}$. Within this dataset, 8 (of the 10) CAS wards were consistently identified as being least representative.

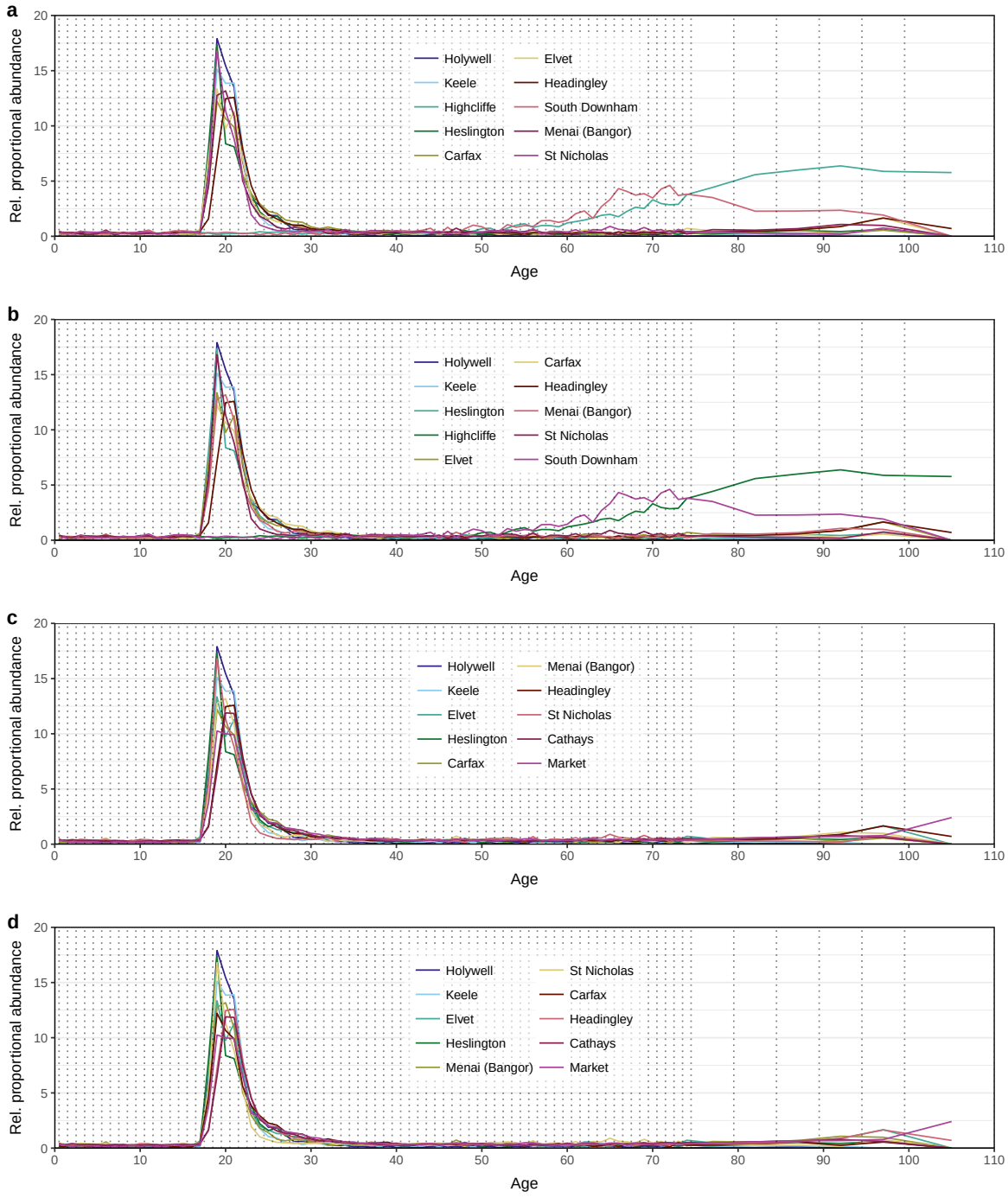


Figure 3.26: The 10 least representative CAS wards of England and Wales, where similarity is calibrated to different vaule of ${}^1G^{\mathbf{Z}}$: (a) $k_{c2} = 0.031$ (${}^1G^{\mathbf{Z}} \approx 2$), (b) $k_{c4} = 0.078$ (${}^1G^{\mathbf{Z}} \approx 4$), (c) $k_{c16} = 0.330$ (${}^1G^{\mathbf{Z}} \approx 16$), and (d) $k_{c32} = 0.728$ (${}^1G^{\mathbf{Z}} \approx 32$). The y-axis is normalised within age classes, by dividing the proportional abundances by the number of ages contained within each class. These values are normalised against the metacommunity, by dividing by the total proportional abundance across England and Wales. Any community with a distribution proportional to the average of England and Wales would have a relative proportional abundance of 1, for all age classes. The vertical dotted lines denote age class cut-offs.

3.4 Case study: Examining the flow of AMR phenotypes in a sympatric population of human and animal hosts

> 'Without urgent, coordinated action, the world is heading towards a post-antibiotic era, in which common infections and minor injuries, which have been treatable for decades, can once again kill.'
>
> — World Health Organization (2015)

3.4.1 Introduction

Antimicrobial resistance (AMR) is the acquired resistance of a microorganism (bacteria, fungi, viruses and parasites) to drugs that are used to treat or prevent infection caused by that organism. It is spreading at an alarming rate and is an important global concern due to its involvement in animal and human disease. This has been blamed, in part, on the use and misuse of antimicrobials (Levy & Marshall, 2004). Under- or over-exposure to these drugs and the subsequent evolutionary selective pressure of resistance in host populations causes susceptible strains to die off, leaving resistant strains to proliferate (Andersson & Hughes, 2010, 2011). Under continued selection, resistance will spread.

The acquisition of resistance by a single population may occur through a number of different mechanisms (for a review, see Alanis, 2005). Sources of infection, particularly by human hosts, are numerous. Imported food and foreign travel increase risk of exposure to isolated (*allopatric*) populations. Within co-located (*sympatric*) populations, local environmental reservoirs and cohabitation by multiple host species lead to the threat of cross-species transfer. Even within a single host, unique strains from different bacterial species may exchange genetic resistance via mobile genetic elements. New resistance phenotypes are emerging each year and in doing so, AMR may be conferred both within and across multiple host species, through various modes of transmission. These may be vertical (*i.e.* inheritance of chromosomal genes across generations, as described in Doss, 1994), or horizontal (*e.g.* insertion or deletion of multiple resistance phenotypes via plasmid exchange, as described in Bennett, 2008) with potential additional spontaneous point mutations in either case. Plasmid-mediated resistance is known to be associated with linkage, which allows multiple resistances to be acquired simultaneously. Genetic linkage between multiple AMR sites maintain the persistence of plasmids so long as at least one antimicrobial, which the bacterium is resistant to, is present in the environment. The probabilities associated with AMR resistance groupings (genetic linkages or likelihood of co-resistance) are beyond the scope of this study and are therefore not described.

Salmonella is a virulent pathogen, and a global health concern, made worse by the antimicrobial

resistance of several strains. In 1984, the epidemic strain, *Salmonella enterica* serovar Typhimurium definitive phage type 104 – hereafter referred to as *Salmonella* DT104 – was isolated in humans (Threlfall et al., 1994). These strains are typically characterised as being multi-drug resistant to ampicillin, chloramphenicol, streptomycin, sulfonamides, and tetracyclines (ACSSuT resistance type), however variants have been identified that are also resistance to fluoroquinolones, trimethoprim, and kanamycin (Boyd et al., 2001). *Salmonella* confers resistance through both vertical and horizontal mechanisms, such that phenotypic precursors may be one or multiple resistances away from the emerging phenotypic group (Carattoli, 2003; Miriagou et al., 2006; Alcaine et al., 2007). Resistance genes are encoded in the multidrug-resistant (MDR) region of *Salmonella* genomic island 1 (SGI1), which is a 43-kb region of the bacterial chromosome (Boyd et al., 2000, 2001; Mulvey et al., 2006). This is of great concern, since unlike plasmid-mediated resistance, which may be lost (albeit slowly) in the absence of selective pressure, encoding resistant traits within the bacterial chromosome allows AMR to be maintained through subsequent generations.

Mather et al. (2012) examined the antimicrobial resistance (AMR) of *Salmonella* DT104 in sympatric human and animal populations in Scotland. The directionality of AMR transmission between the two host populations was assessed by calculating the number of differences in resistance between resistance profiles. To identify from which host population each resistance profile was likely to have originated from, the most probable precursor was defined as the resistance profile with the fewest resistance changes, that was sampled earliest. To identify whether resistance profiles were mainly circulated within a population or transmitted between host populations, the most probable precursor was defined as the resistance profile with the fewest resistance changes, that was sampled closest to the isolate of interest. To assess the ecological diversity of resistance profiles, species richness, Shannon diversity, Simpson diversity, and Berger-Parker diversity (see Section 1.4) were calculated, hypothesising that the most probable precursor should be the population with the most diversity. It was found that the effective number of resistance profiles was greater in human isolates compared to animal isolates and that some resistance profiles in animal isolates were distinct from those found in humans. This showed – somewhat controversially – that in terms of the Scottish dataset, antibiotic resistance was unlikely to have originated from the sympatric animal population.

In this case study, Mather et al.'s (2012) work is advanced by: (1) using beta diversity measures rather than alpha diversity measures, and (2) incorporating the similarity between resistance profiles rather than considering them to be completely distinct. To assess the transmission of antimicrobial resistance between human and animal populations, Reeve et al.'s (2016) beta diversity measures are used to calculate *phenotypic* diversity (a measure of the diversity of antimicrobial phenotypes). This necessitates the inclusion of tailored similarity measures to describe potential evolutionary relationships between AMR resistance profiles. These methods are validated by revisiting the questions asked in Mather et al. (2012) and comparing each set of results.

Acknowledgements
The phenotypic antimicrobial resistance data was kindly provided by Alison Mather and originally obtained from the Scottish *Salmonella*, *Shigella* and *Clostridium difficile* Reference Laboratory (SSSCDRL).

3.4.2 Methods

Dataset
As part of a previous study (Mather et al., 2012), DT104 isolate data from 1990-2004 were obtained from the Scottish *Salmonella, Shigella* and *Clostridium difficile* Reference Laboratory (SSSCDRL); *Salmonella* is a reportable human and animal pathogen in the UK and as such, all veterinary and medical diagnostic laboratories in Scotland are required to forward suspect isolates to the SSSCDRL. Each sample was tested against 13 antimicrobials to produce a record of resistances. The antimicrobials and the percentage of human- and animal-origin isolates resistant to them is listed in Table 3.3. Mather et al.'s (2012) dataset comprised 2439 animal and 2761 human isolates, where animal isolates were collated across a number of host species (Mather et al., 2012, supplementary materials). For the purpose of this work, data were categorised as originating from either animal or human hosts.

Isolates in each host population were characterised phenotypically by the presence or absence of resistance to each antimicrobial. Thus, isolates were attributed a 13-digit binary code corresponding to a distinct resistance profile, with a theoretical total of 2^{13} unique phenotypic groups (assuming every permutation were possible). For example, a code of:

AP	CL	CP	FZ	GM	KA	NAL	NE	SP	ST	SX	TE	TM
1	0	0	0	0	0	0	0	0	1	1	1	0

represents an isolate with resistance to Ampicillin, Streptomycin, Sulphamethoxazole, and Tetracycline. In reality, the number of distinct profiles is constrained by gene linkage and the predominance of certain resistance phenotypes, leaving 65 unique resistance profiles in total. The animal-origin isolates comprised 35 distinct profiles and human-origin isolates comprised 52 distinct profiles, with 22 profiles common to both host groups and 13 and 32 unique, respectively. A connectivity diagram, generated by Mather et al. (2012) using eBURST is shown in Figure 3.27, where 95% of all profiles are connected. Though samples were collected through passive surveillance, over a period of 15 years, sample coverage appears comprehensive enough to have encountered most phenotypic combinations.

Table 3.3: Antimicrobial resistance: List of antimicrobials and the percentage of human- and animal- origin DT104 isolates that are resistant to them.

Antimicrobial	Animal isolates (%)	Human isolates (%)
Ampicillin	16.390	15.957
Chloramphenicol	16.274	15.551
Ciprofloxacin	0.007	0.025
Furazolidone	0.034	0.038
Gentamicin	0.151	0.013
Kanamycin	0.130	0.150
Nalidixic acid	0.370	1.313
Netilmicin	0.116	0.006
Spectinomycin	16.438	16.170
Streptomycin	16.473	16.239
Sulphamethoxazole	16.534	16.595
Tetracycline	16.370	15.626
Trimethoprim	0.713	2.319

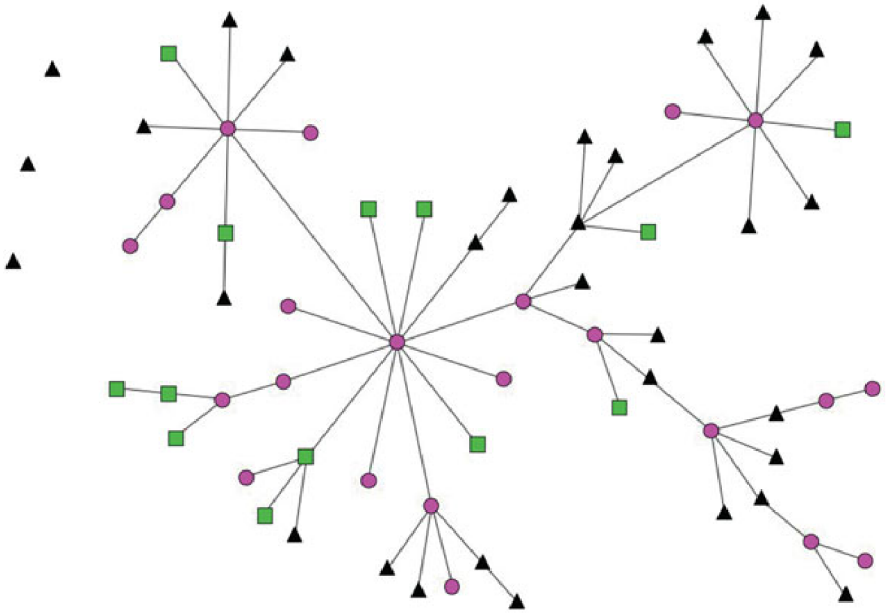


Figure 3.27: Connectivity diagram for human and animal phenotypic resistance of *Salmonella* Typhimurium DT104: Figure taken from (Mather et al., 2012, figure 3). Green squares denote profiles unique to animal isolates, black triangles denote profiles unique to human isolates, profiles found in both human and animal isolates are pink circles. Each profile is connected by lines, which represent the loss or acquisition of resistance to a single antimicrobial.

Calculating diversity

The aim of this work was to use Reeve et al.'s (2016) beta diversity measures to identify the directionality of resistance transmission between animal and human populations. This was done in two parts. The first question sought to determine the most probable phenotypic precursor of the following year's resistance, and complementing this, the second question was posed to identify the emergence of phenotypic novelty. Phenotypic diversity was calculated over time to investigate the transmission of epidemic strains of *Salmonella Typhimurium* DT104 between host populations.

To investigate the antigenic diversity of DT104 on a temporal scale, data were arranged into overlapping 6-year blocks. The amount of historical data available for each analysis was standardised to avoid boundary effects and increase the temporal resolution of these measurements. Each block was defined as a metacommunity, within which emerging resistance was compared to the preceding 5 years of recorded data (Figure 3.28). The emergent year was defined iteratively through 1995-2004, which provided 10 sets of data points for the analysis. It was found that a 6-year time period was both small enough to allow analysis of most of the outbreak – describing the change in AMR through 6 time points – whilst being large enough to provide a good sample size for comparison (though 4-year blocks were also tested). Resistance profiles were considered to constitute *types*, over which diversity was calculated.

Question 1: from which prior host population are existing phenotypes of antimicrobial resistance likely to have originated. As mentioned previously, new resistance phenotypes may arise from various modes of transmission. If resistance traits are transferred through chromosomal mutation, it is likely that genetic change will occur incrementally. On the other hand, if strain resistance is plasmid-mediated, then multiple phenotypes may be gained or lost instantaneously. It is known that DT104 confers resistance through both of these mechanisms, such that phenotypic precursors may be one or multiple resistances away from

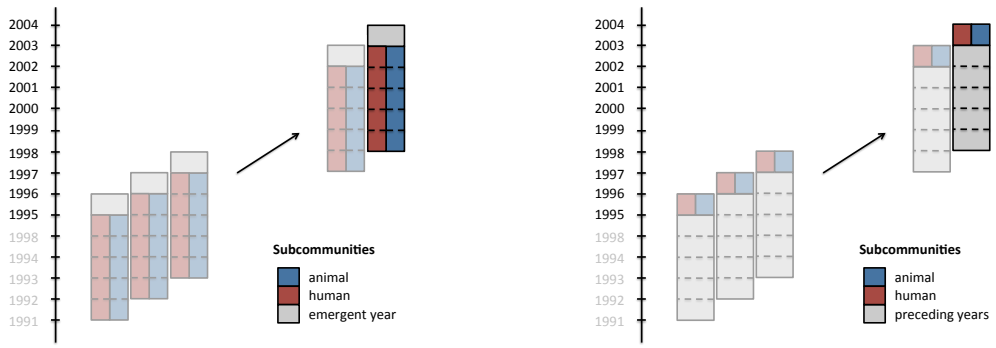


Figure 3.28: Experimental setup to investigate the flow of antimicrobial resistance in human- and animal-origin DT104 isolates: (a) *Question 1* - from which host population does AMR originate? and (b) *Question 2* - in which host population does novelty arise?

the emerging phenotypic group. Although the possible sources of phenotypic variance are extensive, the system was simplified by assuming that resistance in DT104 is primarily conferred via mutation, and that genetic change occurs incrementally. Similarity was defined as $Z_{ii'} = e^{-kd_{ii'}}$ for $k = 1$, where k is a scaling factor and $d_{ii'}$ is the number of differences in resistance between isolates i and i'. In this way, an isolate is completely similar to itself when $Z_{ii'} = 1$, strains differing by a single resistance have a similarity of $Z_{ii'} = e^{-1}$, and no commonality confers a value of $Z_{ii'} = e^{-13}$. Hence, a strain that differs by one resistance phenotype is more likely to originate from a genotypic precursor than one that differs by multiple phenotypes. Each metacommunity (6-year block of historical data) was partitioned into 3 distinct subcommunities: (1) emergent year resistance profiles of human and animal origin, (2) preceding 5 years of human resistance profiles, and (3) preceding 5 years of animal resistance profiles (Figure 3.28a). Data were partitioned in such a way as to permit the comparison of human and animal historic phenotypes. Subcommunity ${}^1\rho_j^{\mathbf{I}}$ was calculated[1], to determine which historical population – of human or animal origin – was the most redundant (*i.e.* which phenotypic precursor was the most comparable to emerging phenotypic resistance).

Question 2: which host population is the most likely source of phenotypic novelty. New resistance phenotypes are emerging each year and in doing so, AMR may be conferred both within and across multiple host species. As mentioned previously, there are a number of different avenues through which strain diversity may emerge. A new strain may arise, during a single event, by horizontal or vertical transfer with potential additional point mutations. To identify the emergence of phenotypic novelty, all sources are considered to be equally distinct, and all resistance profiles – no matter the number of differences – are considered equally dissimilar. To do this, similarity was defined in the naïve-type case (where species are considered to be completely distinct, *i.e.* $\mathbf{Z} = \mathbf{I}$). The dataset was rearranged to focus on emerging resistance, where each metacommunity (6-year block of historical data) was partitioned into 3 distinct subcommunities: (1) emerging animal resistance profiles; (2) emerging human resistance profiles; and (3) the collective human and animal resistances from the preceding 5 years (Figure 3.28b). Subcommunity ${}^1\beta_j^{\mathbf{I}}$ was calculated to determine which emerging population – human or animal – was the most distinct (*i.e.* which was the greater source of emerging phenotypic novelty).

Statistical analysis

Though sample coverage seems good (Figure 3.27). It is unlikely that these data represent a complete sample of DT104 in Scotland and therefore, to account for differential sampling effort and to maintain consistency between human and animal populations, for each year, animal and human data were subsampled down to the smallest number of resistance profiles across hosts (Table 3.4). The redundancy (${}^1\rho_j^{\mathbf{I}}$) and distinctiveness (${}^1\beta_j^{\mathbf{Z}}$) of resistance profiles

[1]Though, subcommunity representativeness (${}^1\bar{\rho}_j^{\mathbf{I}}$) might equally have been used to compare precursory resistance phenotypes, since animal and human subcommunity weights had been standardised

were calculated for each host populations (for 1000 subsampled iterations), over which, mean values and 95% confidence intervals were calculated for each time point. These values were compared within each metacommunity, contrasting the emergent year with 5 years prior.

Table 3.4: Number of recorded isolates per year: Scottish DT104 isolate data from 1990-2004, taken from sympatric human- and animal-host populations. For each year of data, human and animal data were subsampled down to the smallest number of resistance profiles across hosts.

Year	Animal	Human	Subsample
1990	8	10	8
1991	117	90	90
1992	123	144	123
1993	164	234	164
1994	509	265	265
1995	479	397	397
1996	359	511	359
1997	278	308	278
1998	143	219	143
1999	117	150	117
2000	57	145	57
2001	41	85	41
2002	11	81	11
2003	12	49	12
2004	21	73	21
Total	2436	2761	-

3.4.3 Results and discussion

This case study extended the work of Mather et al. (2012), to investigate the transmission of antimicrobial resistance of *Salmonella* DT104 in sympatric human and animal populations in Scotland. Reeve et al.'s (2016) framework was used to calculate phenotypic diversity by incorporating the similarity of resistance profiles and two questions were posed: (1) from which host population was antimicrobial resistance likely to have originated? and (2) from which host population did novelty first arise?

In order to determine from which host population antimicrobial resistance originated, subcommunity ${}^{1}\rho_j^{\mathbf{I}}$ was used to calculate the redundancy of prior human- and animal-origin resistance profiles against the combined resistance of emerging phenotypes (isolates associated with emerging phenotypes were used in the calculation, but not included in the plot). Subcommunity redundancy (in the naïve-type case) is high when species that are rare within the

subcommunity are common throughout the metacommunity as a whole. Therefore, if all isolates were evenly mixed across both host populations, no difference in redundancy should be observed between human and animal populations. Instead, human resistance profiles were found to be consistently more redundant than their animal counterparts, providing no support that AMR originated in the sympatric animal population at any point during the outbreak ($^1\rho_j^{\mathbf{I}}$, Figure 3.29a).

To determine from which emerging host population novelty first arose, subcommunity $^1\beta_j^{\mathbf{Z}}$ was used to calculate the distinctiveness of emerging human- and animal-origin resistance profiles against the combined resistance of the preceding 5 years of phenotypes (omitted from the plot). This measure takes a maximum value when a subcommunity only contains species that aren't present elsewhere in the metacommunity, and therefore the host population with the highest distinctiveness should contain the most phenotypic novelty. Results show that for each emergent year, except 1998 and 2003, resistance profiles taken from human-origin isolates have a higher distinctiveness than those taken from the animal population ($^1\beta_j^{\mathbf{Z}}$, Figure 3.29b). Though the number of samples in each emergent year is relatively low, this provides evidence that phenotypic novelty is arising more in human, than in animal populations, and reinforce previous conclusions that the animal population is unlikely to be the source of the following year's resistance.

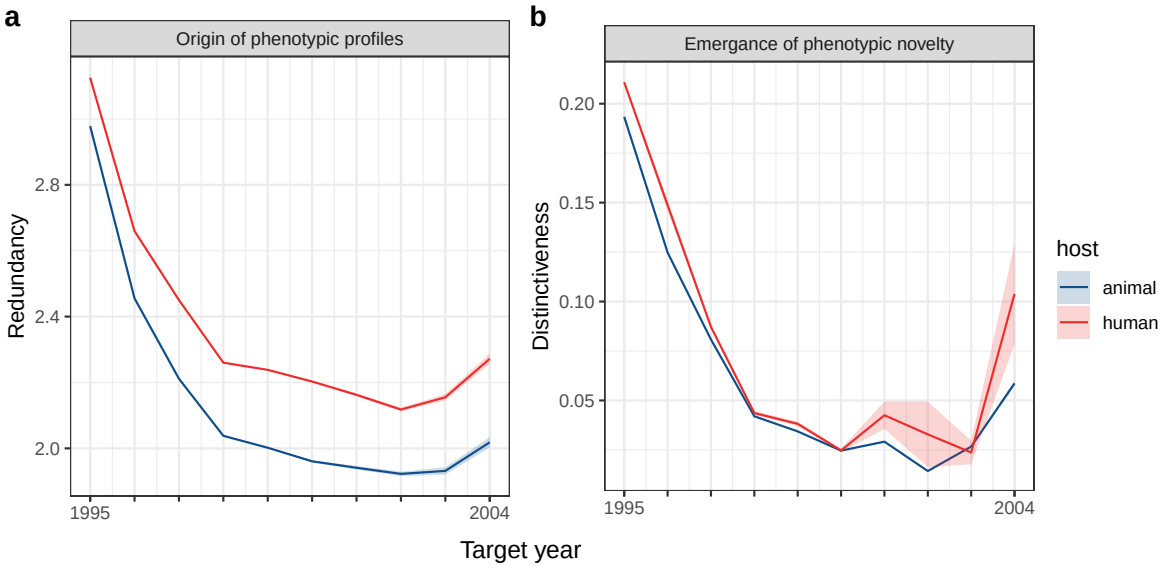


Figure 3.29: Calculating diversity to investigate the flow of antimicrobial resistance in *Salmonella* DT104 from sympatric human and animal populations: (a) Subcommunity redundancy ($^1\rho_j^{\mathbf{I}}$) is used to determine from which prior host populations existing phenotypes of antimicrobial resistance originate, and (b) subcommunity distinctiveness ($^1\beta_j^{\mathbf{Z}}$) is used to determine from which emerging host population phenotypic novelty first arises. The shaded ribbon (where present) denotes a 95% confidence interval.

3.5 Conclusion

This chapter investigated three distinct case studies to illustrate the flexibility and utility of Reeve et al.'s (2016) framework of diversity measures. In the first case study, the functionality of each measure was illustrated with a classic analysis of forest biodiversity, whereas in the second and third case studies novel diversity-based solutions were developed to showcase more unusual applications.

The first case study was focused on examining the biodiversity of the Barro Colorado Island Forest dynamics plot, a fully sampled 50 ha study site. The considerable resolution of this site is useful for demonstrating the properties of Reeve et al.'s (2016) framework of measures, whilst also meaning that sampling errors need not be considered (because most datasets are not this well sampled). The issue of undersampling is considered in Chapter 5. In contrast to other traditional measures of diversity, this framework can be used to investigate subcommunity structure. This was showcased in the first case study where subcommunity-level measures were used to differentiate regions of particular interest. In the naïve-type case (where species are considered completely distinct) clear signals were identified within the BCI Forest dynamics plot. Specifically, two areas were highlighted as being unrepresentative (${}^{q}\bar{\rho}_{j}^{\mathbf{I}}$) of the BCI study site: a swamp in the centre-left and a smaller patch of invasive species in the top-left.

When taxonomic diversity was examined, by incorporating Shimatani's (2001) measure of taxonomic distance, both areas were identified as being unrepresentative of the study site. However, only the swamp was identified as contributing to the biodiversity of the BCI Forest plot (${}^{q}\gamma_{j}^{\mathbf{Z}_{tax}}$, Figure 3.17d). Taxonomic diversity reflects the evolutionary relatedness between tree species (by incorporating information about taxonomic rank), and so, these results highlight the evolutionary uniqueness of swamp-based species. On the other hand, the subcommunities in the top left, though distinct (${}^{q}\beta_{j}^{\mathbf{Z}_{tax}}$, Figure 3.17g), no longer contributed highly to the biodiversity of the study site. These subcommunities are dominated by *A. panamensis*, a species of the same family (Annonaceae) as many other trees in the plot. In contrast to most traditional measures, Reeve et al.'s (2016) framework is sensitive to similarity (be it taxonomic similarity, the similarity between age classes, or the similarity between antimicrobial resistance profiles). Uniquely, it can encompass all of these similarity types within a single framework, rather than being specific to a single one.

The temporal diversity was assessed to reflect how the diversity of the BCI study site changes over time. This was done by rearranging the dataset and using the same framework as above. The turnover in species composition (metacommunity ${}^{q}\bar{B}^{\mathbf{I}}$) over the 8 tree censuses was found to be highest in the swamp (Figure 3.18). These results were confirmed by examining spatial diversity at each individual time point, where the swamp became less sparsely populated and more representative of the study site through time (Figure 3.19).

Another interesting dataset is that of the 2001 census of England and Wales, which was the focus of the second case study. Diversity measures were used to identify the least representative human populations in England and Wales, where similarity was naturally defined between age classes. Two datasets were used in this study, each described the population from different viewpoints. The parish dataset had a high spatial resolution but fewer age classes, and the CAS ward dataset had a lower spatial resolution but a higher number of age classes. Despite these differences, it was possible to design the similarity matrices to behave equivalently within the two different datasets, by varying a scaling parameter to control the amount of similarity between age classes. Within each dataset, this parameter was calibrated to a pre-defined effective number of age classes, ${}^{q}G^{\mathbf{Z}} = 8$. As a result, Reeve et al.'s (2016) measures were able to identify Keele, Heslington, and Cathays within the 10 least representative subcommunities of both datasets, whilst the remaining communities shared the same distinct age profiles (populations inhabited by the young adults in proximity to universities, and retirement communities dominated by the elderly).

Repeating this analysis for ${}^{q}G^{\mathbf{Z}} = 4$, suggested that the CAS ward dataset was more robust to changes in k, and extending these results for ${}^{q}G^{\mathbf{Z}} \in \{2, 4, 8, 16, 32\}$ showed the same trend. Though some variability was observed for ${}^{q}G^{\mathbf{Z}} \in \{4, 8, 16\}$ – where a new demographic feature was detected – 8 (of the 10) CAS wards were consistently identified as being least representative. It seems likely that this robustness was a result of increasing the similarity between age classes (rather than being anything to do with the low spatial resolution), effectively smoothing the demographic profile of each CAS ward, however this requires further investigation. On the other hand, the parish dataset showed consistency in 6 (out of 10) parishes being identified as being unrepresentative for ${}^{q}G^{\mathbf{Z}} \in \{4, 8\}$. This highlighted the importance of carefully considering how similarity is quantified and consequentially, how measured diversity is affected, particularly where few age classes are available. Incorporating these kinds of standardisation techniques into diversity-based methods might prove useful, where possible applications might include a meta-analysis of the functional diversity of different sites where the functional diversity is categorised slightly differently in multiple datasets collected by different research groups.

The focus of the last case study, was to investigate the directionality of transmission of antimicrobial resistance of *Salmonella* DT104 in a sympatric population of humans and animals in Scotland. To examine antimicrobial resistance, a specific genetic or phylogenetic measure is usually required, but the same measures that can detect more "classical" notions biodiversity in the BCI study site can detect phenotypic diversity in antimicrobial resistance. In order to achieve this, the similarity between phenotypic resistance profiles was incorporated into Reeve et al.'s (2016) framework. Measures of beta diversity were then used to determine the flow of resistance between populations. As far as I am aware, this is the first time this method has been used. However, results agreed with those obtained by Mather et al. (2012), showing that antimicrobial resistance was unlikely to have originated from the animal

population.

In summary then, this chapter has shown that Reeve et al.'s (2016) framework is able to examine diversity across a range of different applications and generate results validated by other methods. It also shows the value of investigating subcommunity structure and the ease with which it is possible to generate new diversity measures for new problems by tailoring the similarity matrix. However it does not show that these measures should necessarily be preferred over existing measures that are already being used to solve the same problems. In the next chapter, I take a well established diversity problem, I create diversity measures within this framework to tackle it, and I investigate how these measures compare to those currently used in the literature.

Chapter 4

Phylogenetic beta diversity: Comparisons with traditional measures of phylogenetic diversity

'A limitation of traditional metrics of community similarity is that they do not account for the possible similarity among species that are not shared between communities yet might nonetheless share traits through a common ancestry.'

— Ives & Helmus (2010)

4.1 Abstract

Recently, with advances in sequencing technology, the study of phylogenetics has permeated almost every branch of biology, with applications ranging from the study of the evolutionary relatedness between species, to the epidemiological dynamics of pathogens (Yang & Rannala, 2012). Phylogenetic diversity measures capture this information and thus provide a greater understanding of how evolutionary factors influence patterns of diversity, providing insight into subcommunity structure in terms of species composition and evolutionary relatedness. Thus allowing evolutionary distinctiveness to be measured and preserved.

Phylogenetic beta diversity compares phylogenetic diversity between communities. These measures can be broadly split into two categories – those that look at the phylogeny as a whole, such as Faith's (1992) phylogenetic diversity (Faith's PD), and those that look at pairwise tip distances, such as mean pairwise distance (MPD; Webb, 2000). In this chapter, new measures are developed to quantify phylogenetic diversity from both of these perspectives, providing alternative means with which to study the links inherent in ecological communities. These methods are compared to those commonly used in literature and their robustness is investigated using a variety of different phylogenetic simulations. Overall, results show that compared to measures commonly used in the literature, the new measures of phylogenetic beta diversity are better able to detect phylogenetic signal in community structure, under nearly all circumstances.

To conclude this chapter, the practical utility of these measures is highlighted in a simple case study, which re-investigates the transmission of antimicrobial resistance (AMR) in sympatric

populations of humans and animals, using phylogenetic data to determine whether or not epidemic strains of DT104 and its resistance genes were maintained separately or transmitted extensively between host populations.

4.2 Introduction

To understand the scale of variation in a system, there are many factors to consider. Species richness is the simplest measure of diversity, which counts the number of species in a community. Incorporating relative abundance data reveals information regarding evenness or heterogeneity in abundance. But consider for example, a community comprising three species of crow, and another containing a bear, a sheep, and a goat. Most people would consider the second community to be clearly more diverse. If two communities each have the same species richness, then it is intuitive that the most diverse community will be that which contains the most evolutionary history. Similarity-sensitive measures, such as taxonomic diversity measures, capture this information to reveal a richer representation of the variation in species composition[1].

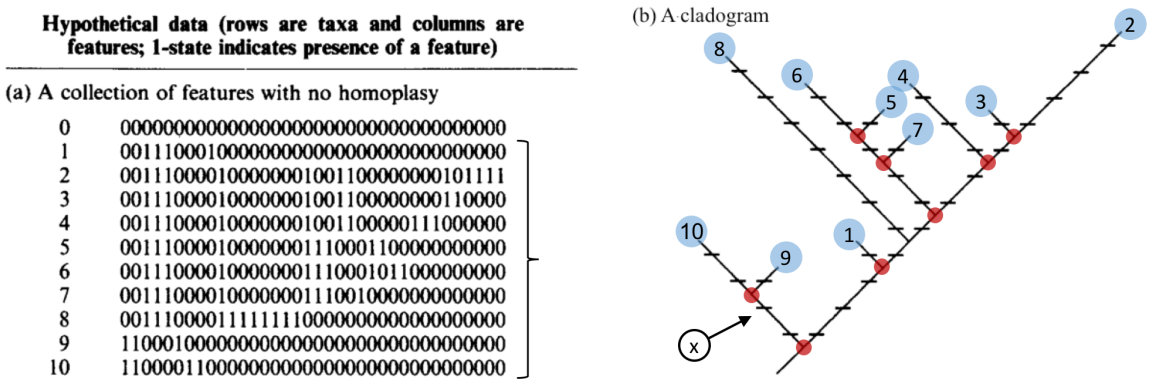
Phylogenetic diversity metrics first appeared in conservation biology when it was noted that traditional measures of biological diversity were "inadequate for the task in hand" (Vane-Wright et al., 1991; Weitzman, 1993): inadequate, because difficulties in estimating the true number of species at a given location made it impossible to accurately represent the biological diversity of a community. To tackle this problem, Vane-Wright et al. (1991) proposed a measure of *taxonomic distinctness*, based on the number of nodes in a taxonomic tree (but see, May, 1990). This was quickly followed by Faith (1992), who proposed a measure of *phylogenetic diversity* with which to assess conservation priorities in terms of the phylogenetic differences between species (by incorporating known branch lengths). In contrast to traditional abundance-based diversity metrics that treat species as being taxonomically distinct (Simpson and Shannon diversities, for example), these metrics ignore relative abundances, focusing instead on cladistic classifications and relationships between species.

Faith's phylogenetic diversity (commonly known as *Faith's PD*) is defined as the sum of branch lengths, written:

$$PD = \sum_{b \in \boldsymbol{b}_T} \ell(b) \tag{4.1}$$

where $\boldsymbol{b}_T$ is a set of branches in the time interval $[-T, 0]$ and $\ell(b)$ is the length of branch b. Faith's PD describes the cumulative evolutionary history of a community from the most recent common ancestor (MRCA). To understand what this means, consider the following example,

[1] These measures rely on complete knowledge of taxonomic information, which might not always be available, and assume that species in each taxonomic unit (*e.g.* two different genera) are equally dissimilar



Hypothetical data (rows are taxa and columns are features; 1-state indicates presence of a feature)

(a) A collection of features with no homoplasy

0	00000000000000000000000000000000000000
1	00111000100000000000000000000000000000
2	00111000010000000100110000000101111
3	00111000010000000100110000000110000
4	00111000010000000100110000000111000000
5	00111000010000000111000110000000000000
6	00111000010000000111000101100000000000
7	00111000010000000111001000000000000000
8	00111000011111111100000000000000000000
9	11000100000000000000000000000000000000
10	11000011000000000000000000000000000000

Figure 4.1: Distribution of features across a set of ten taxa (adapted from Faith, 1992): (a) presence or absence of features, denoted as 1's and 0's, respectively. (b) Cladogram showing the relationships between these taxa, where taxa are grouped according to shared features and tick marks correspond to the derivation of each feature.

adapted from Faith (1992). Figure 4.1a lists the presence or absence of a set of features in the absence of homoplasy[1], where presence and absence are denoted 1 and 0, respectively. These features (which may be phenotypic, behavioural, functional, or ecological) are represented in Figure 4.1b in the form of a cladogram (a phylogeny where only topology is defined). The cladogram comprises a set of external nodes (also known as tips, leaves, or terminal taxa – highlighted by blue circles), which correspond to the set of taxa (taxonomic groups of any rank, such as species, family, class, *etc.*) under study. These taxa are connected to internal nodes (highlighted by red circles), which represent putative ancestors in the phylogeny. Lineages are defined as lines of descent from ancestral nodes to a particular taxon of interest. Typically, a cladogram is used to show the relative recency of common ancestry between species and therefore, the branches (connective lines between nodes) are of arbitrary length. However, in Figure 4.1b, each branch denotes an evolutionary chain between taxa, where tick marks correspond to relevant feature changes. The length of each branch, therefore, corresponds to the number of unique features derived from descendant nodes since the last putative ancestor.

Modern phylogenies are typically represented as bifurcating trees, where each node branches into two descendent lineages, and branch lengths denote the amount of genetic change (measured as the proportion of nucleotide substitutions), or evolutionary time, since divergence. These can be represented as: *non-ultrametric trees* (also known as additive trees, metric trees, or phylograms) or *ultrametric trees* (also known as dendrograms), where terminal taxa are equidistant from the root.

In Figure 4.1b, feature x is shared by taxa 9 and 10, stemming from the second internal node from the root. These taxa are said to form a clade, or monophyletic group (a group of taxa originating from a shared common ancestor). When Faith's PD is calculated, the total amount

[1]Homoplasy occurs when a feature evolves independently in multiple species and is therefore not present in their common ancestor

of evolutionary history represented by this clade is 5, equal to the sum of branch lengths (the total number of tick marks) connecting taxa 9 and 10 to the root. By calculating diversity in this way – where branch lengths are accurate and in the absence of homoplasy – Faith's PD is able to provide much more information than species richness alone. This is useful as a means of prioritising species for conservation, for example, by identifying which set of species represents the greatest amount of evolutionary history.

Phylogenetic beta diversity is identical to beta diversity when species are equally related to each other. Commonly used measures of phylogenetic beta diversity can be categorised into those that consider the phylogeny as a whole (tree-based measures) and those that incorporate phylogenetic relationships as pairwise tip distances (distance-based measures). In this study, I develop new measures of phylogenetic beta diversity (by extending Reeve et al.'s (2016) framework) that use both of these strategies. These new measures are included in the `rdiversity` package (Mitchell & Reeve, 2017), which is available on CRAN. For advanced functionality and detailed examples, see https://github.com/boydorr/rdiversity.

The aim of this chapter is to determine how well these new measures of phylogenetic beta diversity compare to those commonly used in literature (for which packages are available in R packages R Core Team, 2016), in their ability to detect phylogenetic signal in community structure. In the following sections, I randomly generate phylogenies (both ultrametric and non-ultrametric) and evolve traits along them to create subcommunity structure (incidence-based and abundance-based), based on different data types (qualitative and quantitative). In the first instance, I investigate how each measure of diversity performs in the two subcommunity case, since most traditional measures are designed for this purpose. After which, I explore how these results are affected by nestedness, tree size, number of subcommunities, and evolutionary rate.

4.3 Methods for measuring phylogenetic beta diversity

4.3.1 General notation

Consider a metacommunity comprising only two subcommunities, j and k. As described above, the relatedness between S terminal taxa (across the metacommunity as a whole) can be represented as a phylogenetic tree, $\mathfrak{T}$. Then the set of branches associated with each subcommunities j and k are subsets of $\mathfrak{T}$, denoted $\mathfrak{T}_j$ and $\mathfrak{T}_k$, respectively. The total number of terminal taxa in subcommunities j and k are denoted S_j and S_k, respectively. And the number of species descended from branch b in subcommunity j is $S_j(b)$. Likewise for subcommunity k. As described in Section 2.2.2, the relative abundance of species in subcommunity j is represented by the vector $\boldsymbol{P}_{.j} = (P_{1j}, \ldots, P_{Sj})$, where P_{ij} denotes the relative abundance of

species i in subcommunity j. Also, $\bar{\boldsymbol{P}}_{.j} = (\bar{P}_{1j}, \ldots, \bar{P}_{Sj})$ is the normalised relative abundance of species in subcommunity j in isolation. Likewise for subcommunity k.

4.3.2 Tree-based measures of phylogenetic beta diversity

4.3.2.1 Common Branch Length (CBL)

Common Branch Length – a beta diversity analogue of Faith's PD – is the total length of branches shared by $\mathfrak{T}_j$ and $\mathfrak{T}_k$ (Tsirogiannis & Sandel, 2016), written:

$$CBL = \sum_{b \in \mathfrak{T}_j \cap \mathfrak{T}_k} \ell(b) \tag{4.2}$$

where $\mathfrak{T}_j \cap \mathfrak{T}_k$ denotes the set of branches shared by $\mathfrak{T}_j$ and $\mathfrak{T}_k$.

Software: CBL is provided in the `PhyloMeasures` package v2.1 (Tsirogiannis & Sandel, 2015).

4.3.2.2 PD resemblance

As described in Section 1.6.1, measures of compositional similarity (and differentiation) – the most well known being the Jaccard (1901, 1912) and Sørensen (1948) indices – are commonly used to quantify variation between communities. Ferrier et al. (2007) extended these measures to incorporate phylogenetic information, developing a new class of incidence-based measures referred to by Nipperess et al. (2010) as *PD resemblance*. This approach can be applied to any measure (of compositional similarity or differentiation) based on the matching and mismatching components of a 2×2 contingency table, where a is the number of species shared by two communities, b is the number of species distinct to one community, c is the number of species distinct to the other, and d the number of species present in communities outwith the two under study (Koleff et al., 2003). Ferrier et al. (2007) incorporate phylogenetic relationships by redefining a, b, c, and d in terms of total branch length, where a is the total branch length shared by two communities, b is the total branch length distinct to one community, c is the total branch length distinct to the other, and d is the branch lengths absent from the communities under study. Nipperess et al. (2010) then extend this approach to incorporate species-abundance information, using a modification of Tamás et al.'s (2001) framework. Independently, phylogenetic diversity analogues of the Jaccard index (UniFrac) and the Bray-Curtis index (PhyloSor) were derived by Lozupone & Knight (2005) and Bryant et al. (2008), respectively. These are discussed in more detail below.

4.3.2.3 Unique fraction distance metric (UniFrac)

In their seminal paper, Catherine Lozupone and Rob Knight proposed a qualitative measure of beta diversity to compare the distribution of microbial lineages between samples (Lozupone & Knight, 2005). *UniFrac* (the unique fraction distance[1]) between two subcommunities is a qualitative measure that captures the amount of evolutionary history unique to each subcommunity as the fraction of unshared branch lengths.

$$UF = \sum_{i=1}^{n} \frac{b_i \left| I_{P_{ij}>0} - I_{P_{ik}>0} \right|}{\sum_{i=1}^{n} b_i} \tag{4.3}$$

where $I_{P_{ij}>0}$ is an indicator variable (likewise for $I_{P_{ik}>0}$):

$$I_{P_{ij}>0} = \begin{cases} 1 & \text{if } P_{ij} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.4}$$

where P_{ij} and P_{ik} denote the proportion of taxa descended from branch i in subcommunities j and k, respectively. In other words,

$$UF = \frac{PD_{j \cup k} - PD_{j \cap k}}{PD_{j \cup k}} \tag{4.5}$$

where PD is Faith's (1992) phylogenetic diversity. The UniFrac measure is illustrated in Figure 4.2a (taken from Fig.1 Lozupone et al., 2007) where the distance between the circle subcommunity and the square subcommunity is the fraction of branches that are not grey (note that duplicate sequences contribute no additional branch length). This metric captures patterns of diversity related to the terminal nodes of a phylogeny, and when species are equally related (in the case of a star phylogeny), it is equivalent to the Jaccard (1901, 1912) index of dis(similarity).

Weighted UniFrac (wUF) extends this concept such that branches are weighted by the abundance of information they contain (Lozupone et al., 2007):

$$wUF = \frac{\sum_{i=1}^{n} b_i \left| P_{ij} - P_{ik} \right|}{\sum_{i=1}^{n} b_i \left(P_{ij} + P_{ik} \right)} \tag{4.6}$$

This is clearly demonstrated in Figure 4.2b where branch thickness corresponds to the relative abundance of sequences in square or circle subcommunities attributed to each branch. Squares are weighted twice as much as circles, because there are twice as many circles than squares. For example, branch 3 is completely balanced and contributes no weight to the calculations. However, branches 1 and 2 have a thickness of 7 units and are therefore weighted very strongly.

Chen et al. (2012) argue that unweighted and weighted (raw and normalised) UniFrac distances place too much emphasis on rare and highly abundant lineages, respectively. They propose

[1] Equivalent to Δ_T (Bacaro et al., 2007)

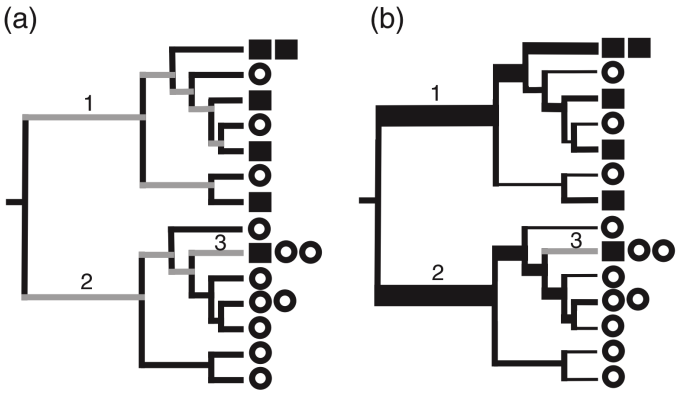


Figure 4.2: Unique fraction distance metric: (a) unweighted UniFrac, and (b) weighted UniFrac. Figure taken from Fig.1 Lozupone et al. (2007). Squares and circles represent sequences from different subcommunities. Lines denote the fraction of branch length with descendants from either the square or the circle subcommunities (black), or both (gray). Line thickness is proportional to how much each branch is weighted in the calculation, where grey branches have no weight.

a new distance measure, Generalised UniFrac (gUF), which is more sensitive to changes in moderately abundant lineages. This measure avoids bias by controlling the weighting on abundant lineages, by incorporates a parameter, α:

$$gUF^{\alpha} = \frac{\sum_{i=1}^{n} b_i \left(P_{ij} + P_{ik}\right)^{\alpha} \left|\frac{P_{ij}-P_{ik}}{P_{ij}+P_{ik}}\right|}{\sum_{i=1}^{n} b_i \left(P_{ij} + P_{ik}\right)^{\alpha}} \tag{4.7}$$

where α is an unnamed parameter[1] that controls the contribution of high-abundance branches towards UniFrac distance. This parameter takes any value between 0 and 1, where as α tends to 1, more emphasis is placed on highly abundant branches. When $\alpha = 0$,

$$gUF^{0} = \frac{\sum_{i=1}^{n} b_i \left|\frac{P_{ij}-P_{ik}}{P_{ij}+P_{ik}}\right|}{\sum_{i=1}^{n} b_i} \tag{4.8}$$

which takes unweighted UniFrac as a special case when abundance data is converted into presence-absence data (whereas weighted UniFrac does not). At the other extreme, when $\alpha = 1$, generalised UniFrac equals weighted UniFrac. Chen et al. (2012) recommends that $\alpha = 0.5$ is used, (being more robust than both weighted and unweighted UniFrac in simulated studies) and so this value is tested alongside $\alpha \in \{0, 1\}$.

Software: weighted and unweighted UniFrac measures are provided in the `phyloseq` package v1.22.3 (McMurdie & Holmes, 2013), on Bioconductor. The generalised UniFrac measures (Chen, 2012) are provided in the `GUniFrac` package v1.1, on CRAN.

[1]Note that Chen et al.'s (2012) α parameter is completely unrelated to alpha diversity

4.3.2.4 Phylogenetic Sørensen index (PhyloSor)

Jessica Bryant and colleagues developed a measure of compositional and phylogenetic similarity, derived from Sørensen's (1948) similarity coefficient. *PhyloSor*, the phylogenetic analogue of the Bray Curtis index (Bray & Curtis, 1957) (and therefore also Sørensen's similarity coefficient), is a similarity metric that describes the fraction of branch length shared by two communities (Bryant et al., 2008),

$$PhyloSor = \frac{PD_{j\cap k}}{\frac{1}{2}(PD_j + PD_k)} \tag{4.9}$$

where PD_j and PD_k are the total lengths of branches descended from taxa contained in subcommunities j and k, respectively, and $PD_{j\cap k}$ is the total branch length shared by communities j and k. In contrast to UniFrac, this metric is sensitive to turnover occurring deeper in the phylogeny.

Software: PhyloSor is provided in the `betapart` package v1.5.0 (Baselga & Orme, 2012)

4.3.2.5 Partitioning turnover and nestedness

Using an additive framework, Andrés Baselga (2010) proposed a new approach to separate Sørensen's (1948) dissimilarity index into species turnover and nestedness components. Using Koleff et al.'s (2003) notation, Sørensen's index can be written:

$$sor = \frac{b + c}{2a + b + c} \tag{4.10}$$

where b and c denote the total number of species distinct to each community, and a is the total number of species shared by them both. This index incorporates both turnover and richness components of diversity. Conversely, Simpson's dissimilarity index describes spatial turnover without the influence of species richness (Simpson, 1943; Baselga, 2010), and is written

$$sim = \frac{\min(b, c)}{a + \min(b, c)} \tag{4.11}$$

When communities have the same number of species, b and c are equal, and therefore *sor* and *sim* are equal. Baselga (2010) states that any dissimilarity occurring between these communities must be due to turnover, since nestedness is not possible when communities share the same species. Therefore, when *sor* and *sim* are not equal, their difference, $sor - sim$, must then be a measure of the nestedness,

$$sne = \frac{\max(b, c) - \min(b, c)}{2a + b + c} \times \frac{a}{a + \min(b, c)} \tag{4.12}$$

Multiple (more than two) community analogues of Sørensen's dissimilarity (SOR), the turnover component of Sørensen's dissimilarity (SIM; Simpson's dissimilarity), and the

nestedness-resultant component of Sørensen's dissimilarity (SNE) are derived by substituting the multiple-community a-component, $\left(\sum_{j<k} a_{jk} - \sum_{j<k<l} a_{jkl} + \sum_{j<k<l<m} a_{jklm} - \dots\right)$ by $\left(\sum_j S_j - S_T\right)$, alongside pairwise b- and c-component analogues, where S_j and S_T denote the total number of species in subcommunity j and across both subcommunities, respectively. These are described in Baselga (2010) and listed here:

$$SOR = \frac{\left(\sum_{j<k} \min(b_{jk}, b_{kj})\right) + \left(\sum_{j<k} \max(b_{jk}, b_{kj})\right)}{2\left(\sum_j S_j - S_T\right) + \left(\sum_{j<k} \min(b_{jk}, b_{kj})\right) + \left(\sum_{j<k} \max(b_{jk}, b_{kj})\right)} \tag{4.13}$$

$$SIM = \frac{\left(\sum_{j<k} \min(b_{jk}, b_{kj})\right)}{\left(\sum_j S_j - S_T\right) + \left(\sum_{j<k} \min(b_{jk}, b_{kj})\right)} \tag{4.14}$$

$$\begin{aligned} SNE = {} & \frac{\left(\sum_{j<k} \max(b_{jk}, b_{kj})\right) - \left(\sum_{j<k} \min(b_{jk}, b_{kj})\right)}{2\left(\sum_i S_j - S_T\right) + \left(\sum_{j<k} \min(b_{jk}, b_{kj})\right) + \left(\sum_{j<k} \max(b_{jk}, b_{kj})\right)} \\ & \times \frac{\sum_j S_j - S_T}{\left(\sum_j S_j - S_T\right) + \left(\sum_{j<k} \min(b_{jk}, b_{kj})\right)} \end{aligned} \tag{4.15}$$

where b_{jk} and b_{kj} are the number of species exclusive to subcommunities j and k, respectively.

Two years later, Baselga (2012) proposed a similar decomposition, partitioning Jaccard's (1901) dissimilarity into turnover and nestedness components. Jaccard's index can be written:

$$jac = \frac{b + c}{a + b + c} \tag{4.16}$$

The turnover component of Jaccard's dissimilarity index is

$$jtu = \frac{2\min(b, c)}{a + 2\min(b, c)} \tag{4.17}$$

Since $jac = jtu + jne$, the nestedness component of Jaccard's dissimilarity index is

$$jne = \frac{\max(b, c) - \min(b, c)}{a + b + c} \times \frac{a}{a + 2\min(b, c)} \tag{4.18}$$

Multiple community analogues of Jaccard dissimilarity (JAC), the turnover component of Jaccard dissimilarity (JTU), and the nestedness-resultant component of Jaccard dissimilarity

(JNE), are described in Baselga (2012) and listed here:

$$JAC = \frac{\left(\sum_{i<j} \min(b_{ij}, b_{ji})\right) + \left(\sum_{i<j} \max(b_{ij}, b_{ji})\right)}{\left(\sum_i S_i - S_T\right) + \left(\sum_{i<j} \min(b_{ij}, b_{ji})\right) + \left(\sum_{i<j} \max(b_{ij}, b_{ji})\right)} \tag{4.19}$$

$$JTU = 2\frac{\left(\sum_{i<j} \min(b_{ij}, b_{ji})\right)}{\left(\sum_i S_i - S_T\right) + 2\left(\sum_{i<j} \min(b_{ij}, b_{ji})\right)} \tag{4.20}$$

$$JNE = \frac{\left(\sum_{i<j} \max(b_{ij}, b_{ji})\right) - \left(\sum_{i<j} \min(b_{ij}, b_{ji})\right)}{\left(\sum_i S_i - S_T\right) + \left(\sum_{i<j} \min(b_{ij}, b_{ji})\right) + \left(\sum_{i<j} \max(b_{ij}, b_{ji})\right)} \times \frac{\sum_i S_i - S_T}{\left(\sum_i S_i - S_T\right) + 2\left(\sum_{i<j} \min(b_{ij}, b_{ji})\right)} \tag{4.21}$$

Leprieur et al. (2012) extend the pairwise-community partitioning framework to explicitly define the phylogenetic equivalents of the Sørensen (PhyloSor) and Jaccard (UniFrac) coefficients.

If each branch b in the phylogeny has a length of $\ell(b)$, then $PD = \sum_{\mathfrak{T}} \ell(b)$, $PD_{tot} = \sum_{\mathfrak{T}_j \cup \mathfrak{T}_k} \ell(b)$, $PD_k = \sum_{\mathfrak{T}_k} \ell(b)$, and $PD_j = \sum_{\mathfrak{T}_j} \ell(b)$. Therefore: $a = PD_k + PD_j - PD_{tot}$, $b = PD_{tot} - PD_k$, and $c = PD_{tot} - PD_j$. PhyloSor and UniFrac can then be written with their turnover components:

$$PhyloSor = \frac{2PD_{tot} - PD_k - PD_j}{PD_k + PD_j} \tag{4.22}$$

$$PhyloSor_{turn} = \frac{\min(PD_{tot} - PD_k, PD_{tot} - PD_j)}{PD_k + PD_j - PD_{tot} + \min(PD_{tot} - PD_k, PD_{tot} - PD_j)} \tag{4.23}$$

$$UniFrac = \frac{2PD_{tot} - PD_k - PD_j}{PD_{tot}} \tag{4.24}$$

$$UniFrac_{turn} = \frac{2\min(PD_{tot} - PD_k, PD_{tot} - PD_j)}{PD_k + PD_j - PD_{tot} + 2\min(PD_{tot} - PD_k, PD_{tot} - PD_j)} \tag{4.25}$$

Software: The phylogenetic analogues of the Jaccard and Sørensen indices for pairwise and multiple subcommunities, alongside their spatial turnover and nestedness components, are provided in the `betapart` package v1.5.0 (Baselga & Orme, 2012).

4.3.2.6 Phylogenetic diversity as 'effective numbers'

The study of phylogenetic diversity has gained considerable interest in recent years and many methods of assessment have been explored. Most commonly, Faith's PD (total phylogenetic distance; Faith, 1992), MPD (mean phylogenetic distance; Webb, 2000), and MNTD (mean nearest taxon distance; Webb et al., 2002). For reviews, see Vellend et al. (2007) and Tucker

et al. (2016). Recently however, Chao et al. (2010) derived a new class of phylogenetic measures based on Hill (1973) numbers (described in Section 1.4). This family of measures quantify phylogenetic diversity in terms of effective numbers (described in Section 1.3) and therefore satisfy sensible mathematical properties such as the replication principle (also described in Section 1.3). Furthermore, like Hill numbers, Chao et al.'s (2010) phylogenetic measures are expressed as a function of the parameter q, which allows the variability of a community to be quantified along a continuum of values known as a 'diversity profile'.

Chao's phylogenetic diversity

Chao et al. (2010) proposed that the diversity of a phylogenetic tree should be considered in terms of the Hill number of the "entire virtual assemblage of ancestral species". Consider the following example. Figure 4.3 shows an ultrametric tree, representing the evolutionary history of four present-day species (terminal taxa) with relative abundances $\boldsymbol{p} = (p_1, p_2, p_3, p_4)$. For any point in time, t, in the interval $[-T, 0]$, the *importance* attributed to ancestral species is calculated by summing the relative abundance of their present-day descendants. For example, the ancestral species connecting nodes A and B has a *relative abundance* of $p_2 + p_3$, which reflects how important this ancestral species is to the present-day assemblage. The *lineage diversity at time t* is calculated by dividing this value by the total abundance at time t and inserting these values into the equation for Hill (1973) numbers of order q. For example, the lineage diversity of slice 2 (during the interval T_2) is calculated as $D(t) = \mathrm{M}_{q-1}(\boldsymbol{p}, \boldsymbol{p})^{-1}$, where $\boldsymbol{p} = (\frac{p_1}{T}, \frac{p_2+p_3}{T}, \frac{p_4}{T})$ and $T = (p_1 + p_2 + p_3 + p_4)$. For a given community, these measures generalise traditional distance-based approaches to account for species relatedness over taxonomic and phylogenetic distance. Therefore, as with traditional measures of spatial diversity, phylogenetic diversity can be calculated within- (alpha), between- (beta), and across- (gamma) multiple subcommunities. However, in contrast to spatial diversity, these measures quantify community structure in terms of species composition and evolutionary relatedness.

Branch diversity, or 'phylogenetic diversity of order q through T years ago', is calculated as the effective number of lineage years:

$$ {}^{q}PD(T) = T \times {}^{q}\bar{D}(T). \tag{4.26} $$

where ${}^{q}\bar{D}(T)$ is described as the 'mean diversity of order q over T years' or the *effective number of maximally distinct lineages* for the time interval $[-T, 0]$, and T is the interval length (which when the phylogeny is non-ultrametric, is replaced with $\bar{T}$, the average distance from root to tip for all terminal taxa). Branch diversity takes Faith's (1992) PD (the total phylogenetic length) as a special case when $q = 0$, (when only species richness is considered).

The effective number of maximally distinct lineages is calculated as:

$$ {}^{q}\bar{D}(T) = \begin{cases} \left(\sum_{b \in \boldsymbol{b}_T} \frac{\ell(b)}{T} p(b)^q \right)^{1/(1-q)} & q \geq 0,\ q \neq 1 \\ \exp\left(\sum_{b \in \boldsymbol{b}_T} \frac{\ell(b)}{T} p(b) \log p(b) \right) & q = 1 \end{cases} \tag{4.27} $$

where $\boldsymbol{b}_T$ is the set of branches in the time interval $[-T, 0]$, $\ell(b)$ is the length of branch b, $p(b)$ is the total relative abundance of species descended from branch b, and – as with Hill numbers – the parameter q determines how sensitive these measures are to ancestral species of differing abundance. This value is high when there are many deep branches that are well represented in the present-day assemblage, and low when branches emerged recently and are poorly represented in the present-day assemblage (Chao et al., 2010). When all branch lengths are equal to T, the effective number of maximally distinct lineages reaches an upper limit equal to Hill numbers ${}^qD$, the effective number of species.

Software: At the time of writing, only the gamma diversity measures are available as an R package and Chao et al.'s (2010) phylogenetic beta diversity measures are therefore not included in the analyses in this chapter.

Leinster & Cobbold's phylogenetic similarity

Subsection 1.5.2 describes Leinster & Cobbold's (2012) measure of similarity-sensitive diversity. This metric can be adapted to calculate phylogenetic diversity by considering a particular period of evolutionary history, rather than the species themselves (as in Chao et al., 2010). A phylogeny represents the inferred evolutionary relationships of a set of species, the tips represent the species themselves, and the root, their most recent common ancestor (MRCA). Internal nodes correspond to speciation events, and these events are connected by internal

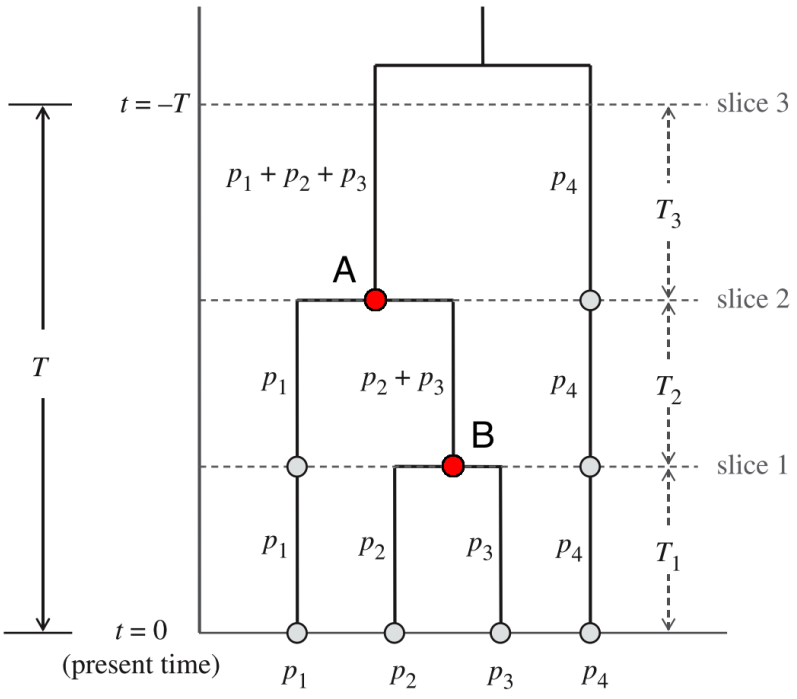


Figure 4.3: An ultrametric rooted phylogenetic tree used to describe how the concept of 'branch diversity' is defined: (Adapted from Figure 1 in Chao et al., 2010).

branches representing periods of shared evolutionary history. Each lineage therefore contains both shared (internal branches) and distinct (terminal branches) periods of evolutionary history, termed *historical species*. Historical species, denoted (i, b), exist during a particular period in the evolutionary history of species i, equivalent to the length of branch b.

Leinster & Cobbold (2012) calculate phylogenetic diversity by substituting the relative abundance of historical species and the similarity between historical species into their measure of similarity-sensitive diversity. Consider a community, $\boldsymbol{p} = (p_1, \ldots, p_S)$, where p_i defines the relative abundance of the i^{th} species. The *relative abundance of historical species* is determined by weighting p_i by the amount of evolutionary history being occupied. Thus the relative abundance of the historical species '(i, b)' – ancestral to species i, occupying branch b – is calculated as:

$$\pi_{(i,b)} = \frac{\ell(b)}{\bar{T}} p_i \qquad \text{for } i \in I_b \tag{4.28}$$

otherwise historic species (i, b) does not exist, and where $\bar{T}$ is the mean evolutionary change per terminal taxa, given as:

$$\bar{T} = \sum_i p_i \ell_i = \sum_b p(b)\ell(b) = \sum_{i,b:i\in \boldsymbol{I}_b} p_i \ell(b) \tag{4.29}$$

where ℓ_i is the total evolutionary change of terminal taxa undergone by species i back to the root of the tree, $p(b)$ is the total relative abundance of species descended from b, and $\boldsymbol{I}_b$ is the set of species descended from branch b.

Using this notation, phylogenetic gamma diversity is therefore calculated as usual:

$${}^{q}D^{\mathbf{Z}}(\boldsymbol{\pi}) = \begin{cases} \left(\sum_{i:\pi_{i,b}>0} \pi_{i,b} \left(\mathbf{Z}\boldsymbol{\pi}\right)_i^{q-1}\right)^{1/1-q} & \text{if } q \notin \{1, \infty\} \\ \prod_{i:\pi_{i,b}>0} \left(\mathbf{Z}\boldsymbol{\pi}\right)_i^{-\pi_{i,b}} & \text{if } q = 1 \\ \left(\max\limits_{i:\pi_{i,b}>0} \left(\mathbf{Z}\boldsymbol{\pi}\right)_i\right)^{-1} & \text{if } q = \infty \end{cases} \tag{4.30}$$

where $\boldsymbol{\pi}$ is a vector of the relative abundance of historical species for $\sum_b \sum_{i:i\in \boldsymbol{I}_b} \pi_{(i,b)} = 1$, and the *similarity between historical species* is defined as

$$Z_{(i,b),(i',b')} = \begin{cases} \bar{T}/\ell_{i'} & \text{if } i' \in \boldsymbol{I}_b \text{ (1 if tree is ultrametric)} \\ 0 & \text{otherwise} \end{cases} \tag{4.31}$$

where for each branch b, $\boldsymbol{I}_b$ is the set $\{1, 2, \ldots, S\}$ of species descended from b, i is the species descended from branch b, i' is the species descended from branch b', and $Z_{(i,b),(i',b')}$ is the pairwise similarity between (i, b) and (i', b'). When i' is part of the set of species descended from branch b, the pairwise similarity between (i, b) and (i', b') is the mean evolutionary change per species relative to the length of evolutionary history descended from species i'. This breaking down of the data into historical species with the associated similarity matrix defined above was done to fit Chao et al.'s (2010) phylogenetic diversity within Leinster &

Cobbold's (2012) broader framework. They therefore give identical results for phylogenetic gamma diversity.

By incorporating this methodology into Reeve et al.'s (2016) framework, we can calculate phylogenetic alpha, beta, and gamma diversities. This is described in the following section.

New tree-based phylogenetic beta diversity

Reeve et al.'s (2016) framework of similarity-sensitive diversity measures can be extended to incorporate the phylogenetic relatedness between species described above. Following Leinster & Cobbold (2012), phylogenetic diversity is calculated across types defined as *historical species*. The relative abundance of the historical species in a single community is written as $\pi_{(i,b)} = \frac{\ell(b)}{\bar{T}} p_i$ (Leinster & Cobbold, 2012). Then, incorporating this into the framework, the relative abundance of the historical species (i, b) in subcommunity j can be written as:

$$\Pi_{(i,b),j} = \frac{\ell(b)}{\bar{T}} P_{ij} \qquad \text{for } i \in I_b \tag{4.32}$$

where $\sum_{b,j} \sum_{i:i\in \boldsymbol{I}_b} \Pi_{(i,b),j} = 1$. The pairwise similarity of historical species is calculated as described above, that is:

$$Z_{(i,b),(i',b')} = \begin{cases} \bar{T}/\ell_{i'} & \text{if } i' \in \boldsymbol{I}_b \text{ (1 if tree is ultrametric)} \\ 0 & \text{otherwise} \end{cases} \tag{4.33}$$

where species i is descended from branch b, species i' is descended from branch b', and $Z_{(i,b),(i',b')}$ is the pairwise similarity between (i, b) and (i', b'). When i' is part of the set of species descended from branch b, the pairwise similarity between (i, b) and (i', b') is the mean evolutionary change per species relative to the length of evolutionary history descended from species i'.

If necessary, a parameter, T, can be used to define how much of the evolutionary history is preserved (effectively allowing the similarity between historical species to be scaled). When $T = 0$, no evolutionary history is preserved, but some terminal taxa remain. For ultrametric trees this is all species, but for non-ultrametric trees, only the most recent species are preserved. As T increases, more evolutionary history is captured and in the case of non-ultrametric trees, older species are gradually included. At $T = 1$, the entire phylogeny presented is preserved and for $T > 1$, the root(s) of the tree is/are extended proportionately.

These newly developed tree-based phylogenetic beta diversity measures can be used to calculate the standard metacommunity-level measures of beta diversity[1], and are written using the following notation: ${}^{q}B^{\mathbf{Z}_{tree}}$ (*average distinctiveness*), ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$ (*the effective number of subcommunities*), ${}^{q}R^{\mathbf{Z}_{tree}}$ (*average redundancy*), and ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$ (*average representativeness*).

[1]In fact, we can use the same methodology to calculate tree-based phylogenetic gamma and alpha diversities from metacommunity-level measures of alpha and gamma

Software: These new measures of phylogenetic diversity are provided in the `rdiversity` package v1.2.1, developed by myself.

4.3.3 Distance-based measures of phylogenetic beta diversity

4.3.3.1 Mean pairwise distance (MPD)

Mean pairwise distance (MPD) is measure of dissimilarity, which gives the average phylogenetic distance between each taxa in a single community (Webb, 2000),

$$MPD = \frac{\sum_{i'} \sum_{i:i<i'} \delta_{ii'} p_i p_{i'}}{\sum_{i'} \sum_{i:i<i'} p_i p_{i'}} \tag{4.34}$$

where p_i and $p_{i'}$ denote the relative abundances of taxa i and i', respectively, and $\delta_{ii'}$ is the phylogenetic distance between them. This metric is commonly considered a basal metric of phylogenetic relatedness, and may be used to determine whether species in a given community are more closely related than expected by chance.

There is a beta diversity measure derived directly from the mean pairwise distance. In a two community sample, $\boldsymbol{j}$ and $\boldsymbol{k}$, the mean pairwise distance separating taxa in these communities is calculated by measuring, for each taxon in sample $\boldsymbol{j}$, the average distance to all taxa in sample $\boldsymbol{k}$, and determining the mean of these values (Webb et al., 2008):

$$MPD_{\boldsymbol{jk}} = \frac{\sum_{i' \in \boldsymbol{j}} \sum_{i \in \boldsymbol{k}} \delta_{ii'} p_i p_{i'}}{\sum_{i' \in \boldsymbol{j}} \sum_{i \in \boldsymbol{k}} p_i p_{i'}} \tag{4.35}$$

Tsirogiannis & Sandel (2016) describe a measure of *Community Distance* (CD), which they describe as being analogous to Webb et al.'s (2008) two-sample MPD. The Community Distance between $\boldsymbol{j}$ and $\boldsymbol{k}$ is calculated as the total branch length connecting each tip in $\boldsymbol{j}$ with each tip in $\boldsymbol{k}$, divided by the total number of paths, which is written,

$$CD = \frac{1}{S_j S_k} \sum_{u \in \boldsymbol{j}} \sum_{v \in \boldsymbol{k}} \ell_{u \to v} \tag{4.36}$$

where $\ell_{u \to v}$ is the sum of all branches connecting nodes u and v.

Software: Both weighted and unweighted versions of pairwise community MPD are provided in the `picante` package v1.7 (Kembel et al., 2010). Community Distance (CD) is provided in the `PhyloMeasures` package v2.1 (Tsirogiannis & Sandel, 2015).

4.3.3.2 Mean nearest taxon distance (MNTD)

Mean nearest taxon distance[1] (MNTD) is a measure of dissimilarity, which describes the average phylogenetic distance between each species and its nearest co-occurring neighbour, in a single community,

$$MNTD = \sum_{i=1}^{S} \min_{i' \neq i}(\delta_{ii'})p_i \tag{4.37}$$

This metric is typically considered a terminal metric of phylogenetic diversity Webb et al. (2002) and is used to determine whether neighbouring species are less closely related than expected by chance.

The corresponding beta diversity measure is the mean nearest taxon distance separating species in two communities and is calculated by measuring, for each species in sample $\boldsymbol{j}$, the nearest phylogenetic neighbour in sample $\boldsymbol{k}$, and determining the mean of these values Webb et al. (2008),

$$MNTD_{\boldsymbol{jk}} = \frac{1}{2}\left(\left(\frac{1}{S_j}\sum_{u \in \boldsymbol{j}} \min_{v \in \boldsymbol{k}} \delta_{u,v}\right) + \left(\frac{1}{S_k}\sum_{v \in \boldsymbol{k}} \min_{u \in \boldsymbol{j}} \delta_{u,v}\right)\right) \tag{4.38}$$

where $\delta_{u,v}$ is the distance between species u in subcommunity $\boldsymbol{j}$ and its phylogenetically nearest species v in subcommunity $\boldsymbol{k}$.

Tsirogiannis & Sandel (2015) describe a two-sample analogue of MNTD, the Community Distance-Nearest Taxon (CDNT) metric,

$$CDNT_{A,B} = \frac{1}{S_j}\sum_{u \in j} \min_{v \in k} \ell_{u,v} \tag{4.39}$$

Since $CDNT_{A,B}$ might not be the same as $CDNT_{B,A}$, Tsirogiannis & Sandel (2015) recommend taking the maximum (mCDNT) or the average (aCDNT) of these values:

$$aCDNT = \frac{\sum_{u \in j} \min_{v \in k} \ell_{u \to v} + \sum_{v \in k} \min_{u \in j} \ell_{u \to v}}{S_j + S_k} \tag{4.40}$$

$$mCDNT = \max\left(CDNT_{A,B}, CDNT_{B_A}\right) \tag{4.41}$$

Software: Weighted and unweighted versions of pairwise community $MNTD$ are provided in the `picante` package v1.7 (Kembel et al., 2010). Maximised Community Distance Nearest Taxon (mCDNT) and Community Distance Nearest Taxon (aCDNT) – beta diversity versions of MNTD – are provided in the `PhyloMeasures` package v2.1 (Tsirogiannis & Sandel, 2015).

[1] Also known as mean nearest neighbour distance (MNND)

4.3.3.3 Quadratic entropy

A well established and commonly used index that combines both species abundance and pairwise distance is Rao's (1982) quadratic entropy[1], Q. Consider an assemblage of species characterised by the relative abundance $\boldsymbol{p} = (p_1, ..., p_S)$. It is well known that Simpson's index of concentration describes the probability that two individuals sampled from the same community belong to the same species. Therefore, the complement of Simpson's index describes the probability of two individuals belonging to different species. This can be rewritten to incorporate phylogenetic relatedness as $D^P = \sum_i \sum_{i'} \delta_{ii'} p_{i.} p_{i'.}$, where $\delta_{ii'}$ is the distance between species i and i', such that $\delta_{ii'} = 0$ when species are identical ($i = i'$) and $\delta_{ii'} = 1$ when they are distinct ($i \neq i'$). Simpson's index is therefore equivalent to Rao's (1982) quadratic entropy in the special case where species are equally related to one another (*i.e.* in the case of a star phylogeny); this is defined as *species identity diversity* (D^I). Conversely, in a bifurcating tree, $\delta_{ii'}$ describes the phylogenetic distance, or divergence time, between species i and i', such that $0 \leqslant \delta_{ii'} \leqslant 1$. It follows then, that D^P measures the average time-since-divergence between two randomly sampled individuals, or equivalently, "the mean phylogenetic distance between distinct species" (Hardy & Senterre, 2007).

Hardy & Senterre (2007) then borrow from population genetics, using additive partitioning methods to define phylogenetic diversity in terms of alpha, beta, and gamma components; but see Villéger & Mouillot (2008), with a rebuttal by Hardy & Jost (2008). Consider a metacommunity comprising N subcommunities where p_{ik} is the relative abundance of species i in subcommunity k. The diversity of a subcommunity in isolation (alpha) is $D_S^P = \sum_i \sum_{i'} \delta_{ii'} f_{ik} f_{i'k}$ and the total diversity (gamma) of a metacommunity is $D_T^P = \sum_i \sum_{i'} \delta_{ii'} p_{i.} p_{i'.}$. The diversity amongst subcommunities (beta) is $D_T^P - D_S^P$, which is simply the additive partitioning of phylogenetic quadratic entropy into alpha and beta entropies; Chave et al. (2007) define the same formulation using different notation. This can be rewritten to describe the phylogenetic diversity of a subcommunity as a fraction of the phylogenetic diversity of the metacommunity as a whole, combining both *species and phylogenetic turnover*:

$$P_{ST} = \frac{D_T^P - D_S^P}{D_T^P} \tag{4.42}$$

where D_S^P and D_T^P denote the mean phylogenetic distance between individuals within a sample and across the metacommunity as a whole, respectively. This measure is equivalent to N_{ST} from classical population genetics. In the naïve case, comparing subcommunity species-identity diversity to the total diversity of the metacommunity yields a measure of *species turnover*[2],

[1] Independently discovered by Warwick & Clarke (1995) as *Taxonomic diversity*, Δ.

[2] Tuomisto (2010a) criticises interpreting this as a measure of either diversity or species turnover, explaining that it is actually a measure of variance

which is equivalent to F_{ST} (or G_{ST})[1]:

$$I_{ST} = \frac{D_T^I - D_S^I}{D_T^I} \tag{4.43}$$

where D_S^I and D_T^I denote the probability that two individuals belong to the same species, within a sample and across the metacommunity as a whole, respectively.

Given that these measures are based on Simpson's index, they intrinsically under-emphasise rare species compared to species richness based measures, and therefore Hardy & Senterre (2007) (following Clarke & Warwick, 1998) define an alternative measure of *phylogenetic distinctness* based on species incidence data:

$$\Pi_{ST} = \frac{\Delta_T^P - \Delta_S^P}{\Delta_T^P}, \tag{4.44}$$

where Δ_S^P and Δ_T^P denote the mean phylogenetic distance between species, averaged across samples, and metacommunity as a whole, respectively. An abundance-based measure of phylogenetic turnover is given in Hardy & Jost (2008):

$$B_{ST} = 1 - \frac{D_S^*}{D_T^*} \tag{4.45}$$

where D_S^* and D_T^* denote the mean quadratic entropy across samples and the metacommunity as a whole, respectively, including sample size correction.

Software: Rao's quadratic entropy is provided by the `picante` package v1.7 (Kembel et al., 2010), though can only be used with ultrametric trees. P_{ST}, I_{ST}, Π_{ST}, and B_{ST} are provided in the `spacodiR` package v0.13.0.0115 (Eastman et al., 2011).

New distance-based phylogenetic beta diversity

As with the new tree-based phylogenetic beta diversity measures, Reeve et al.'s (2016) framework of similarity-sensitive diversity measures can be extended to incorporate the distance-based phylogenetic relatedness between species in partitioned communities described above. However, since this framework requires similarities, these distances need to be transformed. In Chapter 3, I used two strategies for doing this, a linear transformation for taxonomic similarity and an exponential transformation for phenotypic similarity (whether a linear or exponential transformation is selected, determines how similarity is scaled relative to distance). For completeness, here I will investigate both.

Linearly transformed phylogenetic pairwise-distance based beta diversity (henceforth denoted PPD_l) is calculated from the relative abundance of species, with similarity calculated as

[1] A population genetic measure of allelic diversity (Wright, 1951, 1965; Nei, 1973) derived from the Gini-Simpson index and Rao's quadratic entropy

a linear transformation of the pairwise distance between terminal nodes (section 1.6). As with Shimatani's (2001) taxonomic diversity in Chapter 3, distances are transformed as an analogue of nucleotide diversity (Nei & Li, 1979),

$$Z_{ii'} = \begin{cases} 1 - {}^{d_{ii'}}/_{kd_{max}} & \text{if } d_{ii'} > 0 \\ 1 & \text{otherwise} \end{cases} \tag{4.46}$$

where $d_{ii'}$ is the distance between terminal nodes (the phylogenetic equivalent of the number of nucleotide differences per nucleotide site between sequences) and k is a constant that varies the distance required before species are considered completely distinct (see Figure 4.4). Using this method, a linear relationship is observed between distance and similarity, where the distance required for a species to be considered completely distinct increases with k, *i.e.* as k increases, distinct species share less evolutionary history.

Exponentially transformed phylogenetic pairwise-distance based beta diversity (henceforth denoted PPD_e) also calculates diversity from the relative abundance of terminal taxa. However, similarity is now calculated as an exponential transformation of the pairwise distance between terminal nodes:

$$Z_{ii'} = e^{-{}^{d_{ii'}}/_{\left(k \max d_{ii'}\right)}} \tag{4.47}$$

where terminal nodes separated by a particular $d_{ii'}$ are considered more distinct when k is increased.

These newly developed distance-based phylogenetic similarity measures are then used to calculate the standard metacommunity-level measures of beta diversity, and are written using the following notation (when phylogenetic distance is linearly transformed): ${}^{q}B^{\mathbf{Z}_{PPD_l}}$ (*average distinctiveness*), ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$ (*the effective number of subcommunities*), ${}^{q}R^{\mathbf{Z}_{PPD_l}}$ (*average redundancy*), and ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$ (*average representativeness*). Likewise, substituting $\mathbf{Z}_{PPD_l}$ for $\mathbf{Z}_{PPD_e}$ when phylogenetic distance is exponentially transformed.

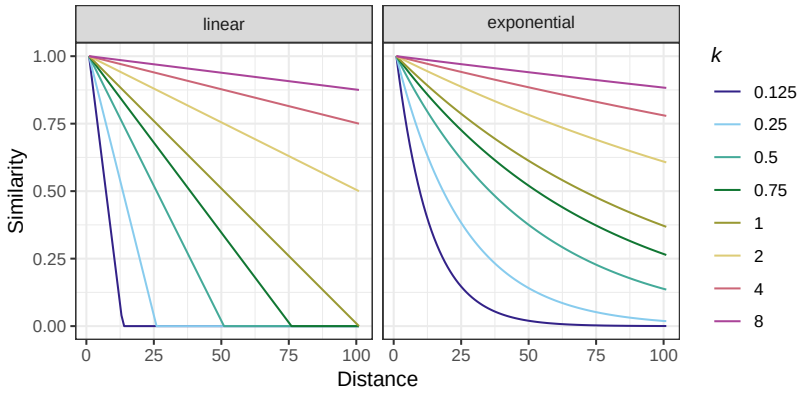


Figure 4.4: Transforming pairwise distances into similarities: How similarity changes with distance at different values of k.

Software: These new measures of phylogenetic diversity are provided by the `rdiversity` package v1.2.1, developed by myself.

4.3.4 Covariance matrix based measures of phylogenetic beta diversity

4.3.4.1 Phylogenetic community dissimilarity (PCD)

Phylogenetic community dissimilarity (PCD) highlights the pairwise differences between communities by considering, "how much of the variance among species in the values of a hypothetical nonselected trait in one community can be predicted by the known trait values of species in another community" Ives & Helmus (2010). This metric can be partitioned into phylogenetic and non-phylogenetic components, such that $PCD = PCD_c \times PCD_p$, where PCD_p is the evolutionary relatedness of unshared species and PCD_c is the proportion of shared species (analogous to Sørensen's index). PCD is calculated,

$$PCD = \frac{n_1 PSV_{1|2} + n_2 PSV_{2|1}}{n_1 PSV_1 + n_2 PSV_2} \tag{4.48}$$

where PSV is the Phylogenetic Species Variability metric, a variance-based metric, which is described in (Helmus et al., 2007).

Software: PCD is provided by the `picante` package v1.7 (Kembel et al., 2010).

4.4 Experimental methods

4.4.1 Generating phylogenies

Ultrametric trees (where terminal taxa are equidistant from the root) and non-ultrametric trees (where lineages vary in length) were randomly generated using the functions: `phytools::pbtree()` and `ape::rtree()`, from phytools v0.6-44 and ape v5.1, respectively (Paradis et al., 2004; Revell, 2012). Rather than modelling real world data, these functions are commonly used to generate simple phylogenetic trees (Chamberlain et al., 2014; Goberna & Verdú, 2016; Pavoine et al., 2017; Plazzotta & Colijn, 2017). The average distance between root and tip was normalised to the same value for ultrametric and non-ultrametric trees. Population structure was simulated by evolving traits (representing the subcommunity identity of terminal taxa) along each lineage of a phylogeny.

4.4.2 Generating community structure

In order to test the ability of phylogenetic beta diversity measures to detect phylogenetic signal in community structure, the presence of individual species in subcommunities within a metacommunity should be determined by some evolved trait or traits. The process by which an evolved trait is associated with a particular subcommunity is hereafter described as the determination of *subcommunity preference*. Where there's only one trait, subcommunity preference for one subcommunity over the other(s) will be either qualitative or quantitative, resulting in incidence or abundance (only in the case of a quantitative trait) data for each species and subcommunity. On the other hand, multiple traits may determine independent preferences for each subcommunity in turn (evolved independently), giving similar results, or multiple quantitative traits may act together to give inter-dependent incidence or abundance data for multiple subcommunities.

Two subcommunities

To generate community structure, traits (either qualitative or quantitative) were evolved along a phylogeny, either independently for each community or in a dependent fashion.

Qualitative traits (*i.e.* 'yes' or 'no' for a discrete character) were evolved along a phylogeny using the function, `phytools::sim.history()` (Revell, 2012). These discrete character traits are stochastically 'evolved' according to the following matrix:

$$\mathbf{Q} = \begin{array}{c} \begin{matrix} \text{a} & \text{b} \end{matrix} \\ \begin{bmatrix} -r_l & r_l \\ r_l & -r_l \end{bmatrix} \end{array} \begin{matrix} \text{a} \\ \text{b} \end{matrix} \tag{4.49}$$

where the entry Q_{ba} denotes the transition rate from states $a \rightarrow b$ (the preference for subcommunity a or b), which is fixed at r_l for all transitions. If the qualitative trait is required to evolve independently for each community, this process is carried out twice (or N times for N communities), and whether or not the state matches the appropriate community determines presence or absence of the species in that community. This results in a subcommunity location for each species, which is translated into a metacommunity matrix (see next section), $\mathbf{P}$, comprising elements $P_{ij} \in \{0, 1\}$, which denote the incidence of species i in subcommunity j.

Quantitative traits were evolved under Brownian motion using `phytools::fastBM()` (Revell, 2012), with a constant rate r_t. This process results in a vector of numeric values, $\boldsymbol{x} = (x_1, \ldots, x_S)$, comprising elements x_i, which correspond to an evolved value associated with species i. Each value was then transformed under an inverse logit (or logistic) function so that $0 < y_{j_i} < 1$ (allowing the value to be viewed as a probability) if the traits are evolved

independently for each community:

$$y_{j_i} = \text{logit}^{-1}(x_{j_i}) = \frac{1}{1 + e^{-x_{j_i}}} \tag{4.50}$$

If incidence or abundance of species in the two communities are evolved together (dependent), then equations for the two communities are as follows:

$$y_{a_i} = \text{logit}^{-1}(x_i) = \frac{1}{1 + e^{-x_i}} \tag{4.51}$$

$$y_{b_i} = \text{logit}^{-1}(-x_i) = 1 - \frac{1}{1 + e^{-x_i}} = \frac{1}{1 + e^{x_i}} \tag{4.52}$$

These quantitative traits were then used to generate either incidence data or abundance data (see next section) to fill the matrix $\mathbf{P}$.

Multiple subcommunities

Qualitative traits were evolved along each phylogeny using the same methods as above, with instantaneous rates of transition defined between four subcommunities:

$$\mathbf{Q} = \begin{array}{c} \begin{matrix} \text{a} & \text{b} & \text{c} & \text{d} \end{matrix} \\ \begin{bmatrix} -3r_l & r_l & r_l & r_l \\ r_l & -3r_l & r_l & r_l \\ r_l & r_l & -3r_l & r_l \\ r_l & r_l & r_l & -3r_l \end{bmatrix} \end{array} \begin{matrix} \text{a} \\ \text{b} \\ \text{c} \\ \text{d} \end{matrix} \tag{4.53}$$

where the diagonal is multiplied by -3 so that rows and columns sum to zero. Likewise, instantaneous rates of transition defined between eight subcommunities:

$$\mathbf{Q} = \begin{array}{c} \begin{matrix} \text{a} & \text{b} & \text{c} & \text{d} & \text{e} & \text{f} & \text{g} & \text{h} \end{matrix} \\ \begin{bmatrix} -7r_l & r_l & r_l & r_l & r_l & r_l & r_l & r_l \\ r_l & -7r_l & r_l & r_l & r_l & r_l & r_l & r_l \\ r_l & r_l & -7r_l & r_l & r_l & r_l & r_l & r_l \\ r_l & r_l & r_l & -7r_l & r_l & r_l & r_l & r_l \\ r_l & r_l & r_l & r_l & -7r_l & r_l & r_l & r_l \\ r_l & r_l & r_l & r_l & r_l & -7r_l & r_l & r_l \\ r_l & r_l & r_l & r_l & r_l & r_l & -7r_l & r_l \\ r_l & r_l & r_l & r_l & r_l & r_l & r_l & -7r_l \end{bmatrix} \end{array} \begin{matrix} \text{a} \\ \text{b} \\ \text{c} \\ \text{d} \\ \text{e} \\ \text{f} \\ \text{g} \\ \text{h} \end{matrix} \tag{4.54}$$

Quantitative traits were evolved using the same methods as above, which when the abundance of species in each subcommunity is defined as being independent:

$$y_{j_i} = \left(1 + e^{-x_{j_i}}\right)^{-1} \tag{4.55}$$

Alternatively, when the abundance of species in each subcommunity are dependent on each other, then N traits are evolved for 2^N subcommunities, which are transformed for four subcommunities:

$$\begin{aligned} y_{a_i} &= \left(1 + e^{-x_{(ab)_i} - x_{(ac)_i}}\right)^{-1} \\ y_{b_i} &= \left(1 + e^{-x_{(ab)_i} + x_{(ac)_i}}\right)^{-1} \\ y_{c_i} &= \left(1 + e^{x_{(ab)_i} - x_{(ac)_i}}\right)^{-1} \\ y_{d_i} &= \left(1 + e^{x_{(ab)_i} + x_{(ac)_i}}\right)^{-1} \end{aligned} \tag{4.56}$$

where $x_{(ab)_i}$ (the preference of species i for subcommunities a and b) and $x_{(ac)_i}$ (the preference of species i for subcommunities a and c) are independently evolved traits. Similarly, for eight subcommunities:

$$\begin{aligned} y_{a_i} &= \left(1 + e^{-x_{(abcd)_i} - x_{(abef)_i} - x_{(abeg)_i}}\right)^{-1} \\ y_{b_i} &= \left(1 + e^{-x_{(abcd)_i} - x_{(abef)_i} + x_{(abeg)_i}}\right)^{-1} \\ y_{c_i} &= \left(1 + e^{-x_{(abcd)_i} + x_{(abef)_i} - x_{(abeg)_i}}\right)^{-1} \\ y_{d_i} &= \left(1 + e^{-x_{(abcd)_i} + x_{(abef)_i} + x_{(abeg)_i}}\right)^{-1} \\ y_{e_i} &= \left(1 + e^{x_{(abcd)_i} - x_{(abef)_i} - x_{(abeg)_i}}\right)^{-1} \\ y_{f_i} &= \left(1 + e^{x_{(abcd)_i} - x_{(abef)_i} + x_{(abeg)_i}}\right)^{-1} \\ y_{g_i} &= \left(1 + e^{x_{(abcd)_i} + x_{(abef)_i} - x_{(abeg)_i}}\right)^{-1} \\ y_{h_i} &= \left(1 + e^{x_{(abcd)_i} + x_{(abef)_i} + x_{(abeg)_i}}\right)^{-1} \end{aligned} \tag{4.57}$$

where $x_{(abcd)_i}$, $x_{(abef)_i}$, and $x_{(abeg)_i}$ are independently evolved traits.

4.4.3 Experimental structure

To test the robustness of each measure, different experimental structures were designed for each phylogenetic simulation.

Experiment 1: Two subcommunities

In the simplest case, phylogenetic beta diversity metrics were compared across a metacommunity comprising two subcommunities and 100 species. For simplicity, the rate of evolution was assumed to be constant over the whole tree. The transition rate, r_l, and the evolutionary rate r_t, for the qualitative and quantitative traits were fixed to 0.02 and 0.2, respectively (different evolutionary rates are investigated in Experiment 5). Six distinct partitioning strategies were designed and these are described below (with examples shown in Figure 4.5).

1-1. Qualitative-dependent (qd): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ derived from the evolution of a single trait along a phylogeny (using the $\mathbf{Q}$ matrix in Equation 4.49). This process yields the presence or absence of species in each subcommunity, such that each species belongs to subcommunity A or B, but not both (Figure 4.5a).

1-2. Qualitative-independent (qi): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ derived from the evolution of two independent traits along a phylogeny (each using the $\mathbf{Q}$ matrix in Equation 4.49). This process yields the presence or absence of species in each subcommunity, such that each species belongs to either subcommunity A or B, both subcommunities A and B, or neither subcommunity A nor B (Figure 4.5b).

1-3. Quantitative-dependent-binary (qdb): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ derived from the evolution of a single trait, $\boldsymbol{x_i}$ (and transformed into probabilities using Equations 4.51 & 4.52). Population structure is generated as presence-absence data (Figure 4.5c):

$$p_{A_i} \sim \text{Bernoulli}(y_{a_i})$$
$$p_{B_i} \sim \text{Bernoulli}(y_{b_i})$$

where $y_{b_i} = 1 - y_{a_i}$ and y_{a_i} is *the probability of species i existing in subcommunity* A.

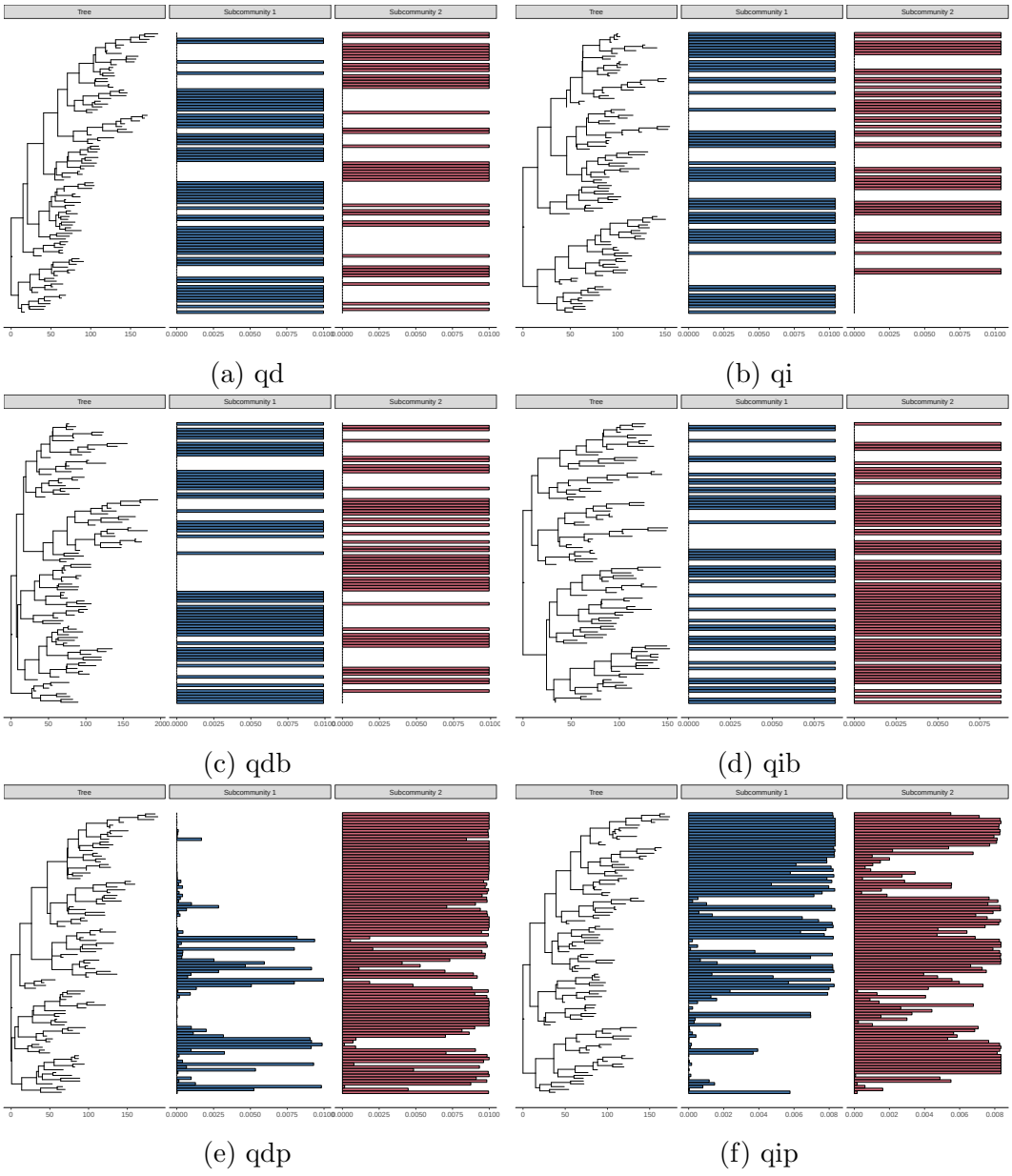
1-4. Quantitative-independent-binary (qib): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ derived from independently evolved traits, $\boldsymbol{x_a}$ and $\boldsymbol{x_b}$, respectively, using the same formulae as above, but y_{b_i} is *the probability of species i existing in subcommunity* B (from Equation 4.50). Population structure is generated as presence-absence data (Figure 4.5d).

1-5. Quantitative-dependent-proportional (qdp): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ dependent on a single trait, $\boldsymbol{x_i}$:

$$p_{A_i} \sim \text{Binomial}(1000, y_{a_i})$$
$$p_{B_i} \sim \text{Binomial}(1000, y_{b_i})$$

where $y_{b_i} = 1 - y_{a_i}$, y_{a_i} is *the probability of an individual of species i belonging to subcommunity* A, and metacommunity structure is generated as species-abundance data (Figure 4.5e).

1-6. Quantitative-independent-proportional (qip): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ derived from independently evolved traits, $\boldsymbol{x_a}$ and $\boldsymbol{x_b}$, respectively, using the same formulae as above, but y_{b_i} is defined as *the probability of an individual of species i belonging to subcommunity* B. Population structure is generated as species-abundance data (Figure 4.5f).



(a) qd (b) qi

(c) qdb (d) qib

(e) qdp (f) qip

Figure 4.5: Examples of abundance distributions generated from each partitioning strategy in Experiment 1: (a) dependent and (b) independently evolved qualitative traits; (c) dependent and (d) independently evolved quantitative traits transformed as the probability of species i being present in subcommunity A; and (e) dependent and (f) independently evolved quantitative traits transformed as the probability of an individual of species i being present in subcommunity A.

Experiment 2: Nested subcommunities

The effect of nestedness was examined using independently evolved traits. Again, each metacommunity comprised two subcommunities and 100 species, where subcommunity structure was defined in six distinct partitioning strategies:

2-1. Qualitative-independent (qi): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ derived from the evolution of two independent traits along a phylogeny (each using the $\mathbf{Q}$ matrix in Equation 4.49).

$$p_{A_i} \sim \text{the incidence of species } i \text{ in subcommunity A evolved from trait } \boldsymbol{x_a}$$
$$p_{B_i} \sim \text{the incidence of species } i \text{ in subcommunity A evolved from trait } \boldsymbol{x_a}$$
$$\times \text{ the incidence of species } i \text{ in subcommunity B evolved from trait } \boldsymbol{x_b}$$

This process yields the presence or absence of species in each subcommunity, such that each species belongs to subcommunity A or B, or both

2-2. Quantitative-independent-binary (qib): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ derived from independently evolved traits, $\boldsymbol{x_a}$ and $\boldsymbol{x_b}$, respectively:

$$p_{A_i} \sim \text{Binomial}(N, y_{a_i})$$
$$p_{B_i} \sim \text{Binomial}(N, y_{a_i} \times y_{b_i})$$

where y_{a_i} and y_{a_i} are defined as in Experiment 1-4. Population structure is generated as presence-absence data.

2-3. Quantitative-independent-proportional (qip): $\boldsymbol{p_A}$ and $\boldsymbol{p_B}$ derived from independently evolved traits, $\boldsymbol{x_a}$ and $\boldsymbol{x_b}$, respectively:

$$p_{A_i} \sim \text{Binomial}(N, y_{a_i})$$
$$p_{B_i} \sim \text{Binomial}(N, y_{a_i} \times y_{b_i})$$

where y_{a_i} and y_{a_i} are defined as in Experiment 1-6. Population structure is generated as species-abundance data.

Experiment 3: Varying tree size

The effect of scale was assessed by generating phylogenetic trees with varying species counts. Preliminary investigation determined that 100 tips was enough to give almost 100% power, therefore experiments were carried out for 10, 25, 50, 75, and 100 tips. Phylogenetic beta diversity metrics were assessed using metacommunities comprising two subcommunities, where six distinct partitioning strategies were used to generate community structure: strategies 3-1 to 3-6 are defined as in Experiments 1-1 to 1-6.

Experiment 4: Multiple subcommunities
The effect of multiple subcommunities was examined by repeating these experiments with four and eight subcommunities. The six partitioning strategies were adapted in the following way:

4-1. Qualitative-dependent (qd): the presence or absence of each species in each subcommunity is determined by the evolution of a single trait, using an appropriate **Q**-matrix (Equations 4.53 & 4.54).

4-2. Qualitative-independent (qi): the presence or absence of each species in subcommunity j is determined by the evolution of a single trait, using an appropriate **Q**-matrix (Equations 4.53 & 4.54). This is repeated for each subcommunity.

4-3. Quantitative-dependent-binary (qdb): the presence or absence of species i in subcommunity j is sampled from $p_{j_i} \sim \text{Bernoulli}(y_{j_i})$, where for 2^N subcommunities y_{j_i} is the logit transform of N independenly evolved traits (Equations 4.56 & 4.57).

4-4. Quantitative-independent-binary (qib): the presence or absence of species i in subcommunity j is sampled from $p_{j_i} \sim \text{Bernoulli}(y_{j_i})$, where y_{j_i} is the logit transform of x_{j_i} (Equation 4.55). This is repeated for each subcommunity.

4-5. Quantitative-dependent-proportional (qdp): the abundance of species i in subcommunity j is sampled from $p_{j_i} \sim \text{Binomial}(1000, y_{j_i})$, where for 2^N subcommunities y_{j_i} is the logit transform of N independenly evolved traits (Equations 4.56 & 4.57).

4-6. Quantitative-independent-proportional (qip): the abundance of species i in subcommunity j is sampled from $p_{j_i} \sim \text{Binomial}(1000, y_{j_i})$, where y_{j_i} is the logit transform of x_{j_i} (Equation 4.55). This is repeated for each subcommunity.

Experiment 5: Varying qualitative evolutionary rates
The effect of evolutionary rates was investigated in qualitative traits. Varying the qualitative transition rate, $r_l \in \{1, 0.2, 0.1, 0.02, 0.01\}$:

5a-1. Qualitative-dependent (qd): defined in Experiment 1-1.

5a-2. Qualitative-independent (qi): defined in Experiment 1-2.

Varying the quantitative rate, $r_t \in \{0.002, 0.02, 0.2, 2\}$:

5b-1. Quantitative-dependent-binary (qdb): defined in Experiment 1-3.

5b-2. Quantitative-independent-binary (qib): defined in Experiment 1-4.

5b-3. Quantitative-dependent-proportional (qdp): defined in Experiment 1-5.

5b-4. Quantitative-independent-proportional (qip): defined in Experiment 1-6.

Examples of kinds of subcommunity structure generate using each of these parameters is shown in Appendix B.1.

4.4.4 Calculating diversity

The aim of this chapter was to compare the new measures of phylogenetic beta diversity (developed in this work) to those commonly used in the literature, in their ability to detect phylogenetic signal in community structure. Commonly used measures for which R packages (R Core Team, 2016) are available are listed in Table 4.1 and described in Section 4.3.

The new measures of phylogenetic beta diversity were calculated for all integer values of $0 \leq q \leq \infty$, as well as non-integer values when $0 \leq q \leq 1$. Tree-based measures (described in Section 4.3.2.6) were calculated for all $T \in \{0, 0.25, 0.5, 0.75, 1, 2\}$. The entire phylogeny is preserved when $T = 1$ and removed when $T = 0$. When $0 < T < 1$, only part of the phylogeny is preserved, at a proportionate distance from the most recent tip. For non-ultrametric trees, this may result in the removal of early terminal taxa. Finally, when $T > 1$ a root is extended proportionately. The following measures were calculated: *average distinctiveness* (${}^{q}B^{\mathbf{Z}_{tree}}$), *the effective number of subcommunities* (${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$), *average redundancy* (${}^{q}R^{\mathbf{Z}_{tree}}$), and *average representativeness* (${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$).

Distance-based measures (described in Section 4.3.3.3), PPD_l and PPD_e, were calculated for the same values of q, for all $k \in \{0.125, 0.25, 0.5, 0.75, 1, 2, 4, 8\}$. The following measures were calculated: *average distinctiveness* (${}^{q}B^{\mathbf{Z}_{PPD_l}}$), *the effective number of subcommunities* (${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$), *average redundancy* (${}^{q}R^{\mathbf{Z}_{PPD_l}}$), and *average representativeness* (${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$). Likewise, substituting $\mathbf{Z}_{PPD_l}$ for $\mathbf{Z}_{PPD_e}$ when phylogenetic distance is exponentially transformed.

4.4.5 Statistical analysis

The purpose of this study was to extend Reeve et al.'s (2016) framework to handle phylogenetic beta diversity and compare these new measures to those commonly used in the literature (discussed in Section 4.3). Phylogenetic beta diversity was calculated using the packages described in Section 4.3 (in R) and permutation tests were used to determine whether measures were capable of detecting phylogenetic signals in community structure.

Permutation test

Here I test the hypothesis that no phylogenetic signal can be detected in the subcommunity structure (the null hypothesis). To simulate a null model of phylogenetic structure, species (and their associated incidences/abundances in the subcommunities) were randomly assigned to tips on the tree, obscuring any association between species identity and phylogenetic structure.

This was done 999 times and phylogenetic diversity was recalculated from the resultant distributions, resulting in 1000 diversity values. The resultant p-value is the probability of observing *an effect* at least as extreme as that observed from measured results of the true tree. That is, assuming subcommunity structure is not related to the phylogeny, how likely is it that measured diversity would be at least as large as that observed? These values are calculated by determining the proportion of all observations that were at least as extreme as the true value of diversity. For each experiment and each partitioning strategy, p-values were calculated individually for 128 randomly generated phylogenies. Then to summarise results, an average was taken.

Power analysis

In hypothesis testing, the significance level (α) is an arbitrarily selected level of acceptable false positive detection (type I error; incorrect rejection of the null hypothesis). For each phylogeny, a p-value less than $\alpha = 0.05$ is considered significant, the null hypothesis is rejected[1], and a phylogenetic signal is considered to have been detected. That is, when $\alpha = 0.05$, there is a 1 in 20 chance of detecting a false positive[2].

Power (*sensitivity*) describes the probability that a true positive has been detected (correctly rejecting the null hypothesis when it is false). This is calculated as the proportion of p-values less than or equal to $\alpha = 0.05$. That is, the proportion of each group of 128 repeats in which *the effect* can be successfully distinguished from the null model. Assuming the null hypothesis really is false, and here we have set up the experiment so that it is, a power analysis determines the quality of the test, in this case how well the measured diversity can identify phylogenetic signals in community structure.

4.5 Results and discussion

For the first experiment I show the results in detail. However, this requires an extraordinary level of often superfluous detail, so I develop a summary that extracts the key information from all subsequent experiments.

[1] A p-value greater than 0.05 requires that the null hypothesis (that community structure is random) is not rejected. However, not rejecting the null hypothesis is not the same as saying the null hypothesis is true.

[2] Increasing α results in a greater probability that the null hypothesis will be rejected, increasing the power of the test, but also increasing the probability that a false positive will be detected. Decreasing α reduces the probability that a false positive will be detected (*e.g.* setting $\alpha = 0.000001$ gives a 1 in 1,000,000 chance of detecting a false positive), but increases the chance of detecting a false negative (type II error; incorrect support of the null hypothesis).

4.5.1 Experiment 1: Two subcommunities

4.5.1.1 Tree-based measures of phylogenetic beta diversity

Permutation tests

Figure 4.6 plots p-values for all values of q, for each of the new tree-based phylogenetic beta diversity measures, where blue lines denote results obtained from individual tests (128 in total) and the red line highlights the mean. Here, subcommunity structure is generated using a qd (qualitative dependent) partitioning strategy. Results show that these tree-based measures are slightly better able to pick out subcommunity structure when phylogenies are ultrametric, where mean p-values are consistently near or less than 0.05 for all values of T (the proportion of the tree preserved) and $q \leqslant 100$ (Figure 4.6a). When phylogenies are non-ultrametric, the calculated p-values are much more variable, perhaps because there is less information contained within the branches of a non-ultrametric tree. Indeed, ${}^{q}G^{\mathbf{Z}_{tree}}$ is slightly higher for ultrametric trees than non-ultrametric trees with the same number of tips. Mean p-values are near to or less than 0.05 for all measures when $1 \leq q \leq 2$ and $T \in \{0.75, 1, 2\}$ (*i.e.* when most of the phylogeny is included), but as T decreases, the average p-values increase, often dramatically (Figure 4.6b). This shows that the phylogenetic signal in the community structure is harder to detect with only part of the tree. Interestingly, at higher values of q ($q > 2$ when $T = 0.25$, increasing to $q > 10$ for $T = 2$, when the root is extended), p-values increase rapidly for normalised measures (${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$ and ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$), whereas raw measures (${}^{q}R^{\mathbf{Z}_{tree}}$ and ${}^{q}B^{\mathbf{Z}_{tree}}$) remain relatively constant for all values of q. Whilst these figures are useful to visualise the results, there would be at least 200 of them for our measures alone and so it was necessary to devise summary figures.

In Figure 4.7, mean p-values (the red lines from the previous figure) are plotted against q for all partitioning strategies and for all values of T. Again, results indicate that tree-based measures of phylogenetic beta diversity are better able to identify community structure from ultrametric trees (Figure 4.7a). Specifically, average p-values are found to be near to or less than 0.05 when $0 < q \leqslant 100$ for abundance-based data (qdp and qip) evolved across ultrametric trees for all values of T. For incidence-based data (qd, qi, qdb, and qib), similar results are found for the same parameters, as well as $q = 0$. For non-ultrametric trees when $T \in \{0.75, 1, 2\}$, for all partitioning strategies, normalised measures (${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, and ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$) result in average p-values near to or less than 0.05 when $0 \leqslant q \leqslant 2$ (Figure 4.7b). Whereas raw measures (${}^{q}R^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$) have similar results when $1 \leqslant q \leqslant 100$. As T is decreased below 0.75 (*i.e.* less of the phylogeny is used to calculate the similarity between historical species), the average p-values increase for all measures.

Table 4.1: Functions to calculate phylogenetic beta diversity: List of measures and the notation used in this chapter to represent them. Also listed are the R packages used to calculate each measure with associated references and citation count (as of January 2018, Web of Science).

	Metric	Notation	R package
Tree-based	Weighted Unifrac	$w\_UniFrac$	`phyloseq (v1.22.3)` (McMurdie & Holmes, 2013), 677
	Phylogenetic Sørensen index	$PhyloSor$ / SOR	`betapart (v1.5.0)` (Baselga & Orme, 2012), 217
	Turnover component of PhyloSor	sim / SIM	
	Nestedness component of PhyloSor	sne / SNE	
	Unique fraction distance	$UniFrac$ / JAC	
	Turnover component of Unifrac	jtu / JTU	
	Nestedness component of Unifrac	jne / JNE	
	Generalised Unifrac	$gUniFrac$	`GUniFrac (v1.1)` (Chen et al., 2012), 130
	Weighted generalised Unifrac	$w\_gUniFrac$	
	VAW generalised Unifrac	$vaw\_gUniFrac$	
	Tree-based phylogenetic beta diversities ($tree$)	${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$ ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, ${}^{q}B^{\mathbf{Z}_{tree}}$	`rdiversity (v1.2.1)` This thesis
Distance-based	Mean nearest taxonomic distance	$MNTD$	`picante (v1.7)` (Kembel et al., 2010), 1158
	Weighted MNTD	w_MNTD	
	Mean pairwise distance	MPD	
	Weighted MPD	w_MPD	
	Rao's quadratic entropy[1]	Rao	
	Phylogenetic Turnover[1]	F_{ST}	
	Phylogenetic turnover	P_{ST}	`spacodiR (v0.13.0115)` (Eastman et al., 2011), 16
	Abundance-based phylogenetic turnover	B_{ST}	
	Phylogenetic distinctness	PI_{ST}	
	Community distance[2]	CD	`PhyloMeasures (v2.1)` (Tsirogiannis & Sandel, 2015), 15
	Common Branch Length[3]	CBL	
	Community Distance Nearest Taxon[4]	$aCDNT$	
	Maximised CDNT	$mCDNT$	
	Linearly transformed phylogenetic distance-based beta diversities (PPD_l)	${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$ ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}B^{\mathbf{Z}_{PPD_l}}$	`rdiversity (v1.2.1)` This thesis
	Exponentially transformed phylogenetic distance-based beta diversities (PPD_e)	${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$ ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}B^{\mathbf{Z}_{PPD_e}}$	
Covariance-based	Phylogenetic community dissimilarity	PCD	`picante (v1.7)` (Kembel et al., 2010), 1158

[1] Only calculated for ultrametric trees;

[2] Beta diversity version of MPD (calculated for more than 2 subcommunities);

[3] Beta diversity version of PD (calculated for more than 2 subcommunities);

[4] Beta diversity version of $MNTD$ (calculated for more than 2 subcommunities).

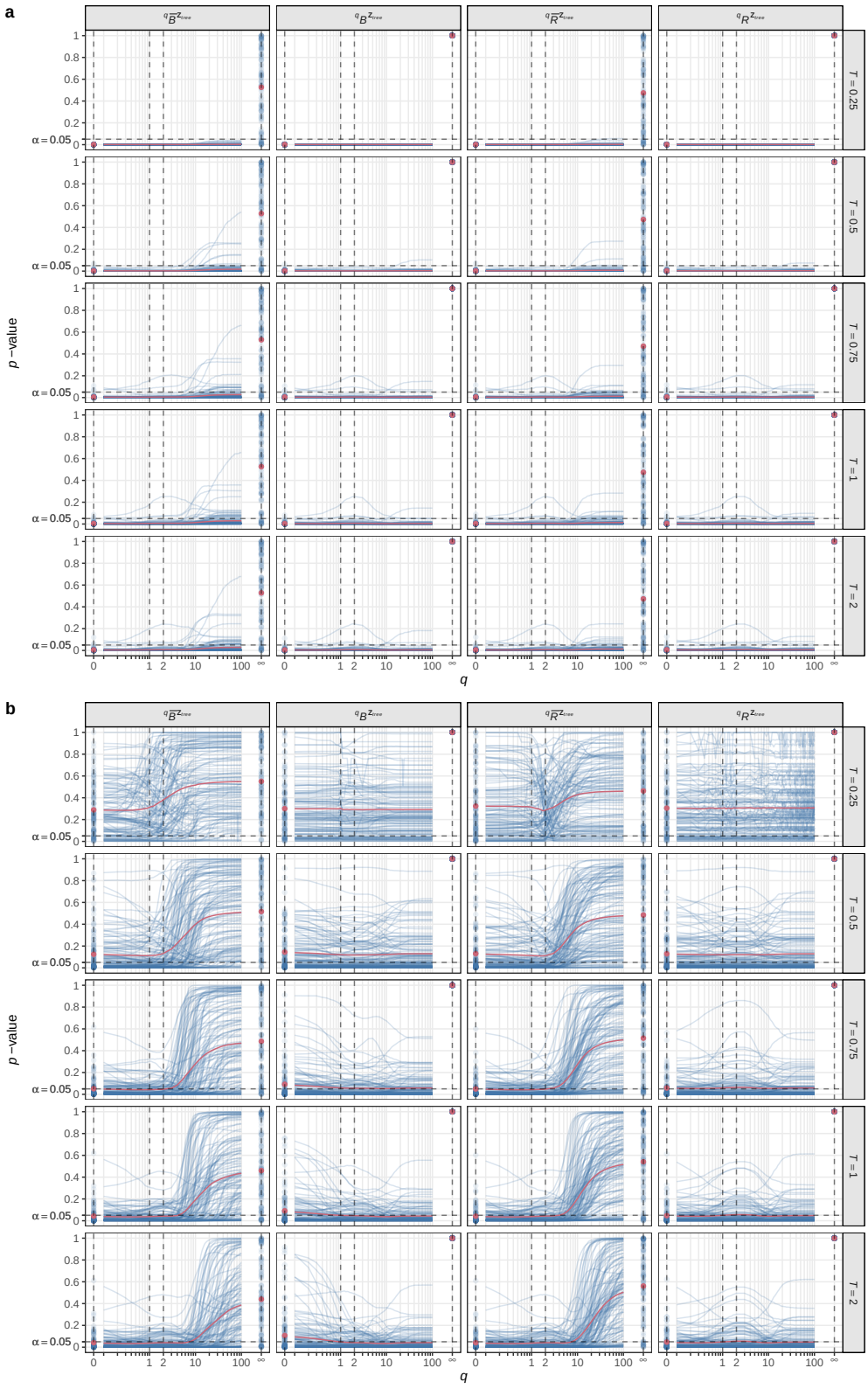


Figure 4.6: Plots of p-values against q, calculated from tree-based phylogenetic beta diversity measures, with community structure generated using a qd (qualitative-dependent) partitioning strategy: Plots show ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$ across different lengths of evolutionary history (different values of T) for (a) ultrametric and (b) non-ultrametric phylogenetic trees. Blue lines correspond to result obtained from different phylogenies ($N = 128$) and red lines highlight average p-values.

Overall, p-values are near 0.05 for all measures and partitioning strategies evolved across ultrametric or non-ultrametric trees, when $T \in \{0.75, 1, 2\}$ and $q \in \{1, 2\}$. However, although these figures allow much more data to be presented, and enable comparisons between different partitioning strategies, averaging the p-values does not allow us to see how often signals can be detected. By selecting a specificity, we can determine the power of each measure to identify phylogenetic signal in community structure.

Power

Figure 4.8 shows the power of each tree-based measure of phylogenetic beta diversity to detect subcommunity structure, based on a chosen specificity ($\alpha = 0.05$). For ultrametric trees, for all values of T, and all partitioning strategies, power is between 0.938 and 1 for all measures when $1 \leq q \leq 2$ (Figure 4.8a). Likewise, for non-ultrametric trees when $T = 1$, measures have a power of between 0.891 and 1 when partitioning strategies are quantitative (qdb, qib, qdp, and qip, Figure 4.8b). Generally for non-ultrametric trees, results are poor for $T \leq 0.5$. This is not surprising, since the tree is being cut in such a way that many evolutionarily older tips are removed, so species, as well as evolutionary history data are being lost. When partitioning strategies are qualitative (qd and qi), power varies between 0.641 and 0.758 over the same parameters. However, these results may be due to different evolutionary rates between qualitative and quantitative partitioning strategies, since the parameters r_l and r_t are not directly comparable.

Note 1: After preliminary testing, the values of r_l and r_t were set to 0.02 and 0.2, respectively, for most of the experiments in this chapter. The effect of changing these values, to vary the rate of evolutionary change in community preference is investigated in Section 4.5.5.

Note 2: Power is sometimes marginally higher when $T = 2$ (*i.e.* when the root is extended to the same length as the tree itself) for non-ultrametric trees, and when T is low for ultrametric trees. However, in general $T = 1$ (*i.e.* the whole tree) is both a natural choice for tree-based phylogenetic diversity measures and is generally close to or actually the most powerful value of T to use in these analyses, and so this is the value I use from now on.

In the following sections, results are further summarised by averaging p-values across partitioning strategies within the four categories: (1) incidence data evolved over ultrametric trees, (2) incidence data evolved over non-ultrametric trees, (3) abundance-based data evolved over ultrametric trees, and (4) abundance data evolved over non-ultrametric trees. Categories 1 and 2, therefore include the qd, qi, qdb, and qib partition schemes, whereas categories 3 and 4 include the qdp and qip strategies. An example of this is shown in Figure 4.9, for tree-based phylogenetic beta diversity measures. These plots show that for partitioning strategies based on incidence data, power is high for all measures when $q \leq 8$, and for higher values of q, raw diversity measures (${}^{q}R^{\mathbf{Z}_{tree}}$ and ${}^{q}B^{\mathbf{Z}_{tree}}$) perform best overall. Likewise for non-ultrametric abundance-based strategies, though power is lower for all measures when

$q \leq 0.2$. On the other hand, for ultrametric abundance-based strategies, power is high for all measures when $0.2 \leq q \leq 8$, but at higher values of q (when power starts to drop off), ${}^{q}R^{\mathbf{Z}_{tree}}$ and ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$ yield the best results. Power is lower when attempting to detect structure from metacommunities generated by non-ultrametric incidence-based partitioning strategies. Again, this may be caused by the qualitative results in Figure 4.8a, since there is no direct correspondence between values of r_l and r_t. However, since all diversity measures are tested on exactly the same trees and community structures, averaging across these partitioning strategies provides a fair comparison. Critically however, across all partitioning schemes, the power of each measure is high when $1 \leq q \leq 2$, and so when measures are compared later in the chapter, $q = 1$ is used for simplicity (to reduce the number of comparisons made).

4.5.1.2 Distance-based measures of phylogenetic beta diversity

Measures of linearly transformed distance-based phylogenetic beta diversity (${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$) are best able to detect community structure (for 100-tip phylogenies and two subcommunities) when $k = 0.25$ (Appendix B.5). When results are summarised in the same way as above, at this value of k, power is found to be high across all measures for $1 \leq q \leq 2$ (Appendix B.8). Indeed for these values of q, power is generally as high or higher than the tree-based method above, especially for non-ultrametric incidence-based partitioning strategies, where the tree-based measure was weakest.

Measures of exponentially transformed distance-based phylogenetic beta diversity (${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$) are best able to detect community structure when $k = 0.125$ (Appendix B.6). When results are summarised (for 100-tip phylogenies and a 2-subcommunity assemblage), power is as high as or slightly higher than the linearly-transformed distance-based measures, particularly when $1 \leq q \leq 2$ (Appendix B.9).

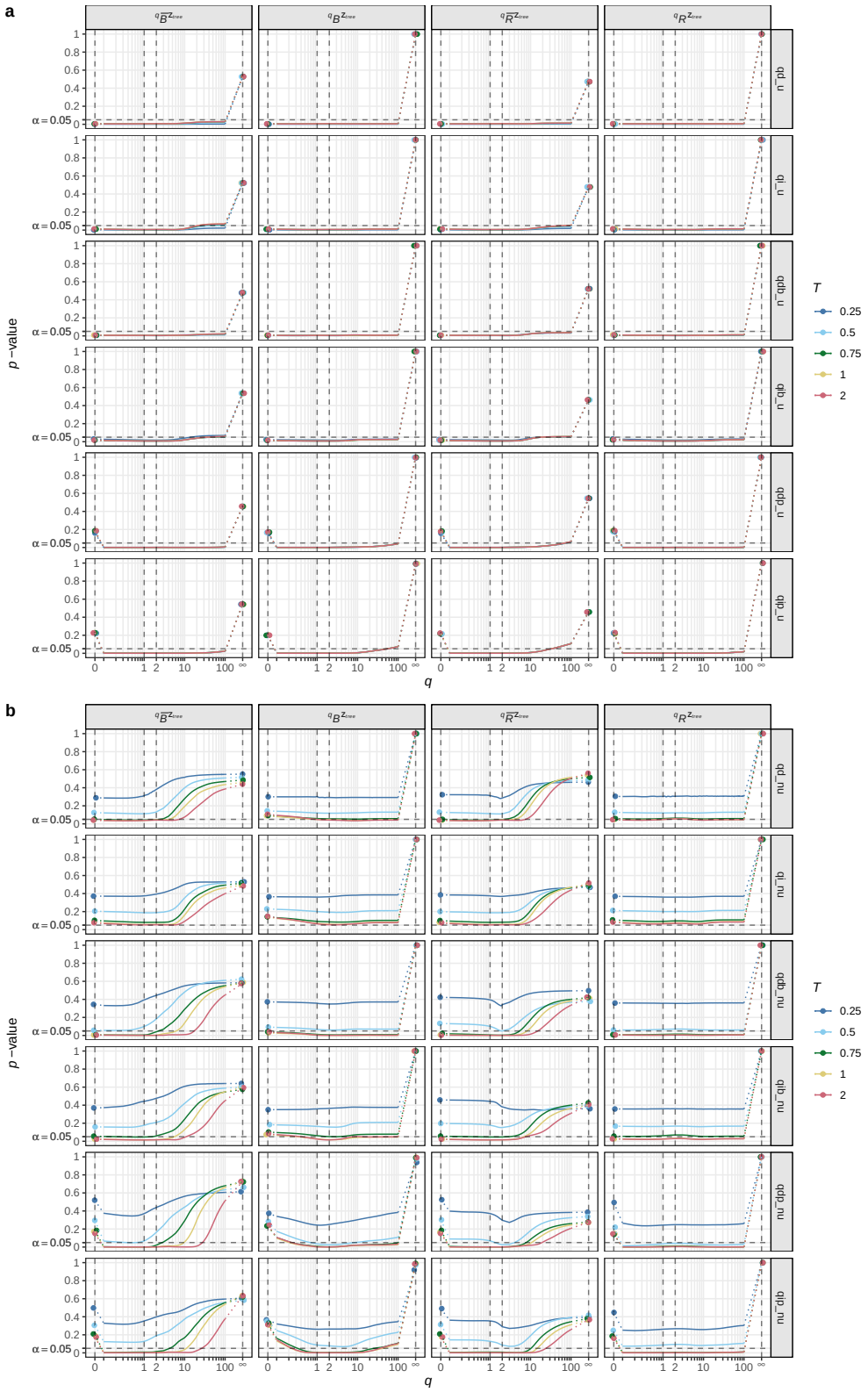


Figure 4.7: Plots of *average* p-values against q, calculated from tree-based phylogenetic beta diversity measures, with community structure generated for *all* partitioning strategies: Plots show ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$ across different proportions of evolutionary history (different values of T) for (a) ultrametric and (b) non-ultrametric phylogenetic trees.

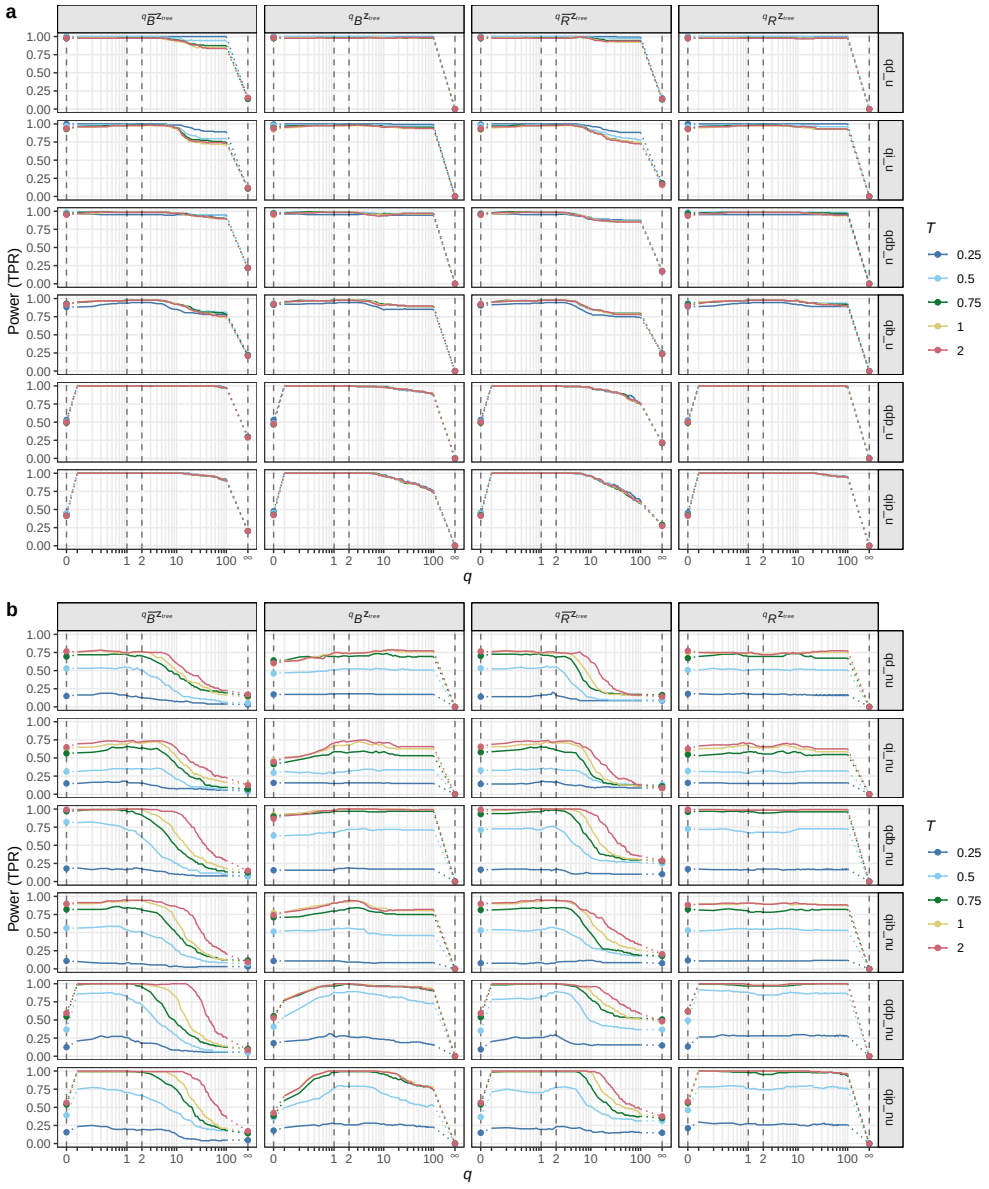


Figure 4.8: Plots of power against q, calculated from tree-based phylogenetic beta diversity measures, with community structure generated for all partitioning strategies: Plots show ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$ across different lengths of evolutionary history (different values of T) for (a) ultrametric and (b) non-ultrametric phylogenetic trees.

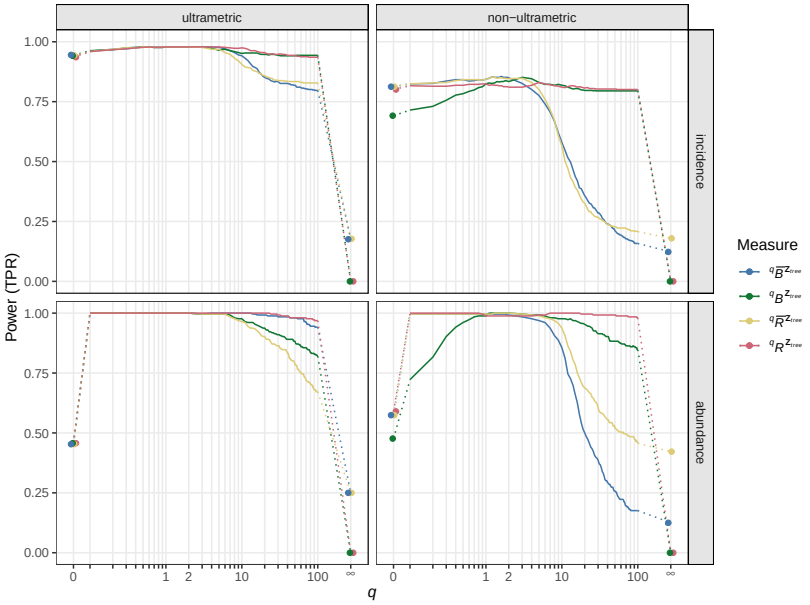


Figure 4.9: Summary of power against q, calculated from tree-based phylogenetic beta diversity measures: Power is calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, and averaged across all subcommunity partitioning strategies within each categorical group.

4.5.1.3 Comparison of phylogenetic beta diversity measures

Now, results are compared to the commonly used tree-based and distance-based measures presented in Section 4.3. In general, results show that for every measure, the power to detect community structure is high when $q = 1$, and so this value is selected for all comparisons throughout the rest of this chapter. Table B.1 lists these values for the new tree-based ($tree$) and distance-based (PPD_l and PPD_e) phylogenetic beta diversity measures, for $T = 1$, $k = 0.25$, and $k = 0.125$, respectively. These results show that across all partitioning strategies, power is greater than 0.66, 0.84, and 0.86 for all $tree$, PPD_l, and PPD_e measures, respectively. All power values for the remainder of this chapter are shown in Appendix B.2.

Figure 4.10 shows a comparison between the power of the phylogenetic extensions of Reeve et al. (2016) measures derived in this chapter and those commonly used in the literature. Note that $q = 1$ is a special case where ${}^{1}\bar{B}^{\mathbf{Z}} = ({}^{1}\bar{R}^{\mathbf{Z}})^{-1}$, so calculated power is identical for each of these measures. These are therefore plotted together and referred to simply as normalised $tree$, PPD_l, and PPD_e. Likewise, ${}^{1}B^{\mathbf{Z}} = ({}^{1}R^{\mathbf{Z}})^{-1}$ are plotted together and referred to as raw. Overall, the PPD_e (red) measures have the highest power. However, PPD_l (yellow) is almost as good. Amongst the traditional measures, $w\_Unifrac$ (weighted UniFrac), B_{st}, and P_{st} also have high power for abundance-based data, whereas PCD (phylogenetic community dissimilarity), w_MNTD (weighted mean nearest taxon distance), $MNTD$ (mean nearest taxon distance), and $gUnifrac(a0.5)$ and $gUnifrac(a0)$ (generalised UniFrac) have high power for incidence-based data. For the remaining experiments, this summarised structure is used to compare the diversity measures, with more detailed figures shown in Appendix B.3.

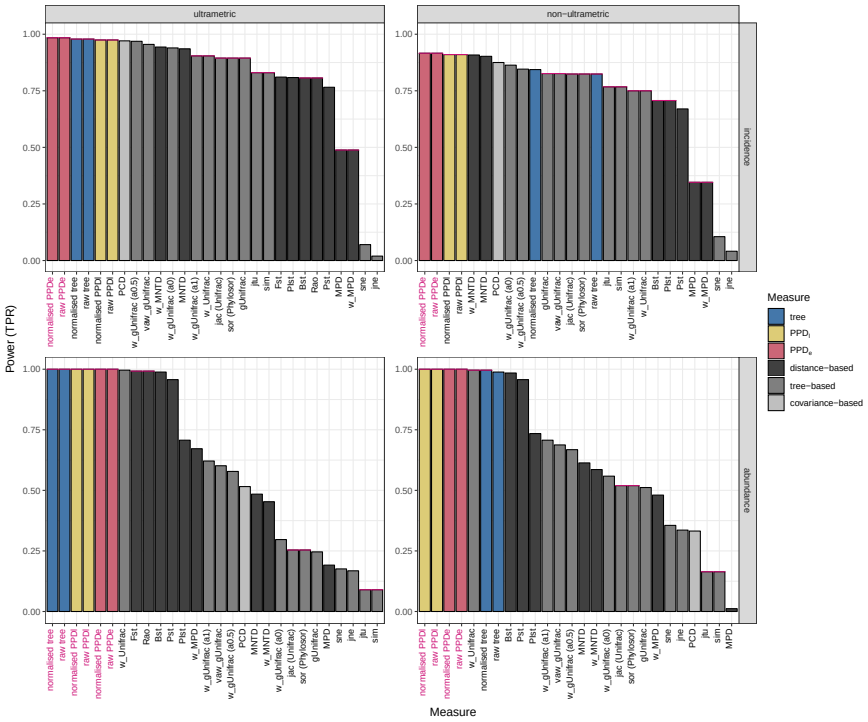


Figure 4.10: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure for *two* subcommunities (Experiment 1): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

4.5.2 Experiment 2: Nested subcommunities

This experiment investigates the effect of nestedness on how well phylogenetic signal can be detected from community structure. The effect of nestedness was examined using independently evolved traits in a 2-subcommunity assemblage. As with the previous experiment, Reeve et al.'s (2016) measures have a very high power to detect community structure within nested subcommunities. For abundance-based data, PPD_e, normalised *tree*, and B_{st} have power greater than 85% (Figure 4.11). Whilst the remaining measures are close to 50% or less. All measures perform poorly for incidence-based data, though PPD_e, B_{st}, and PI_{st} (and for ultrametric trees, *tree*) have the highest power.

As described in Section 4.3.2.5, measures *jne* and *sne* were specifically derived to measure nestedness (Baselga, 2010), but have a power of less than 30% when tested. This is because the underlying *jac* and *sor* measures, from which they are derived, are themselves this low in power. However, *jtu* and *sim*, the turnover components, do 'successfully' achieve a power of 0%.

Summary: power is highest for abundance-based data, where PPD_e, normalised *tree*, and B_{st} are best able to detect phylogenetic signal from nested communities.

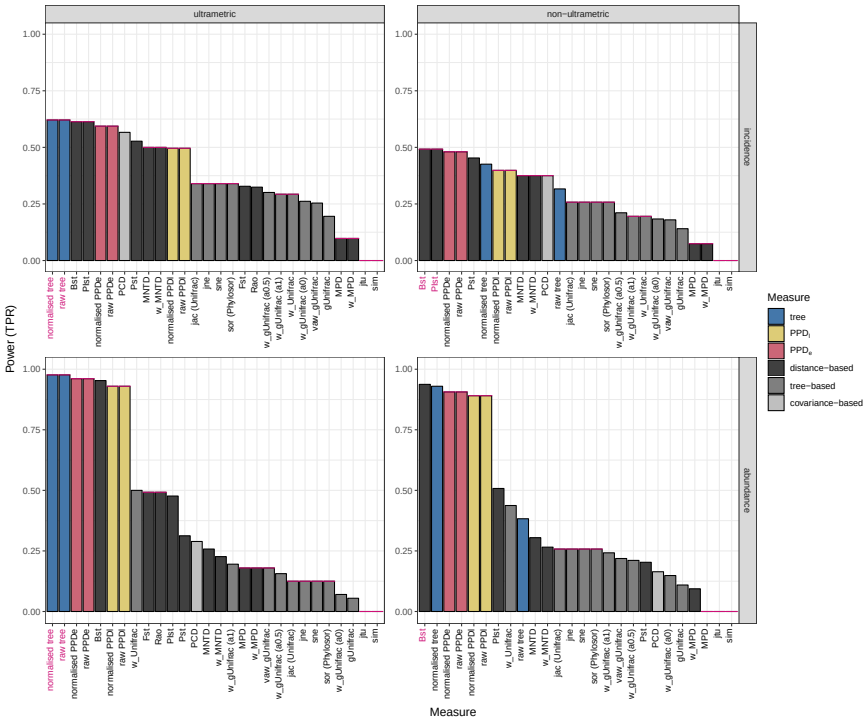


Figure 4.11: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure for *nested* subcommunities (Experiment 2): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

4.5.3 Experiment 3: Varying tree size

In this experiment, the number of tips are varied to determine how well each measure is able to detect community structure from different sized phylogenies. A 10-tip phylogeny corresponds to a metacommunity comprising 10 species, which are again distributed across two subcommunities. As expected, the power of all measures is considerably lower than for previous experiments (Figure 4.12). Nevertheless, the new measures continue to provide good results, particularly PPD_e measures, which are almost always the best at detecting community structure at this small scale. In fact, these are the only measures to have a power greater than 50%, though this is only for abundance-based data evolved across ultrametric trees.

For a 25-tip phylogeny, PPD_e measures always have the highest power overall, with PPD_l and *tree* also yielding good results (Figure 4.13). This suggests that the non-ultrametric incidence result for 10-tip phylogenies, where $vaw\_gUnifrac$ was very slightly better, may have been due to chance. Traditional measures perform less well, though some do stand out. For abundance-based data evolved across non-ultrametric trees, $w\_Unifrac$, B_{st}, and P_{st} have a power greater than 70%. Likewise, the same measures, in addition to *Rao* and F_{st} (which are only calculated for ultrametric trees), perform well for abundance-based data evolved across ultrametric trees, with power greater than 65%.

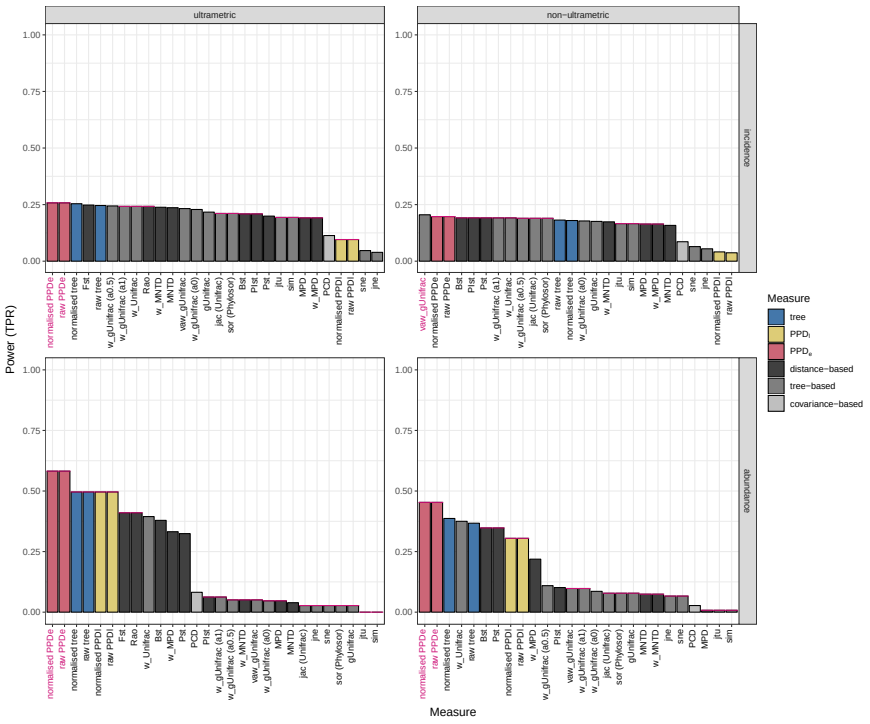


Figure 4.12: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from *10-tip* phylogenies (Experiment 3): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

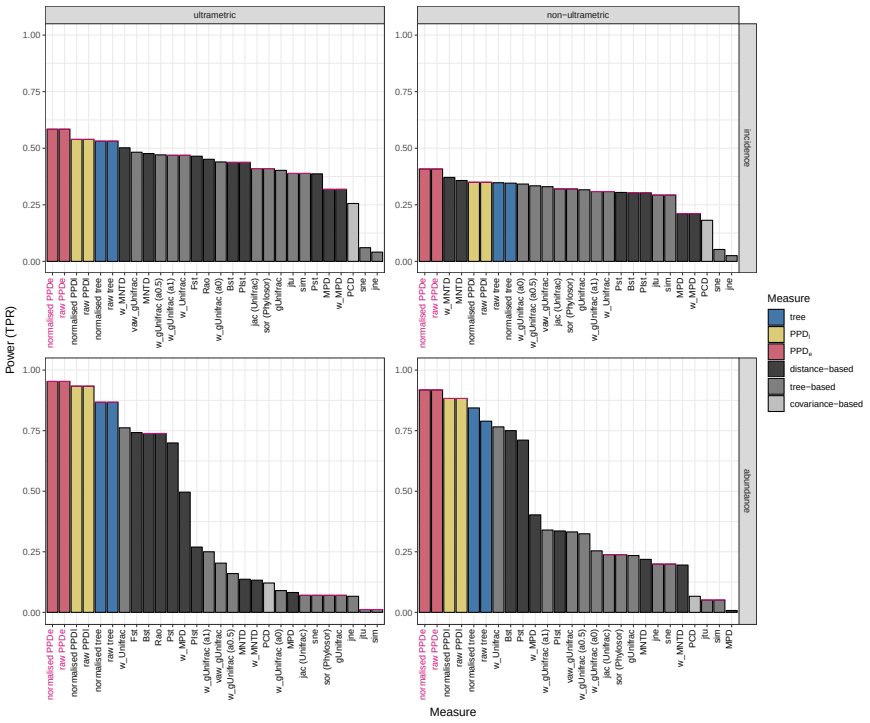


Figure 4.13: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from *25-tip* phylogenies (Experiment 3): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

For 50-tip phylogenies, again PPD_e measures perform best overall, with PPD_l, and *tree* close behind (Figure 4.14). These measures are now equal or close to 100% power. And again, $w\_Unifrac$, B_{st}, P_{st}, *Rao* and F_{st} also yield good results. Increasing to 75-tips, and then 100-tips, for abundance-based data, PPD_e and PPD_l have now plateaued at 100% power, with *tree* close behind (Figures 4.15 & 4.16).

Summary: power is highest for abundance-based data, where PPD_e is best able to detect phylogenetic signal from community structure, even when trees are very small.

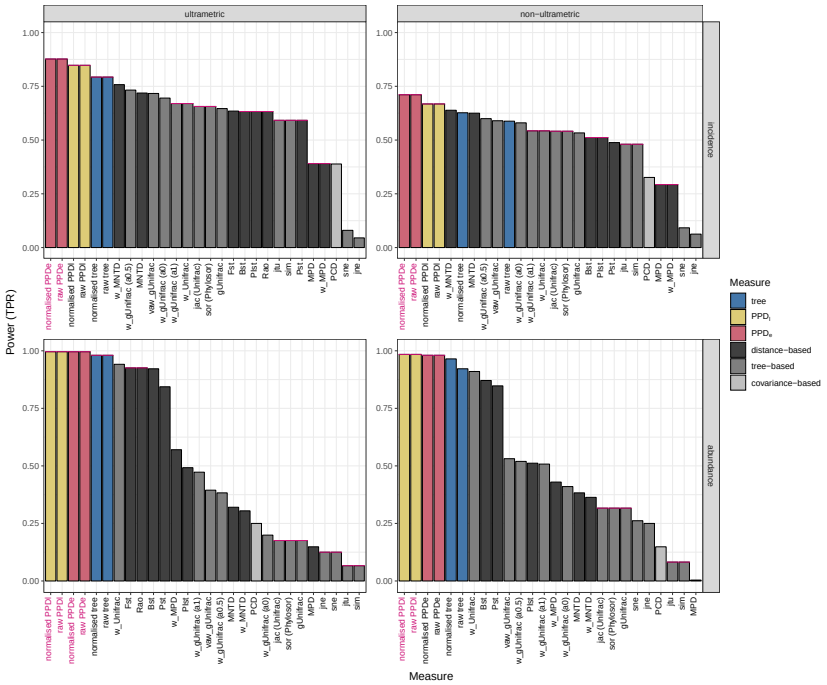


Figure 4.14: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from *50-tip* phylogenies (Experiment 3): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

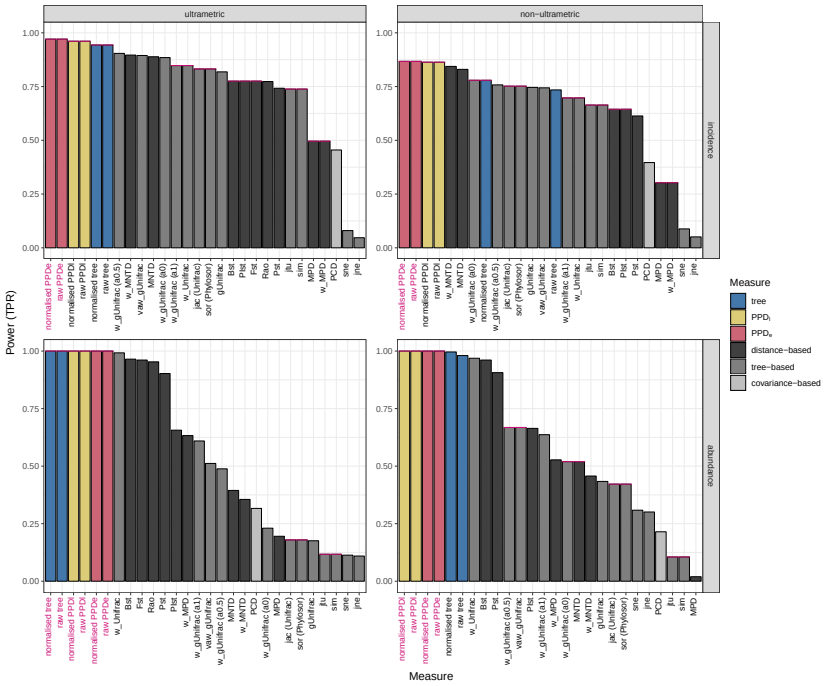


Figure 4.15: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from *75-tip* phylogenies (Experiment 3): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

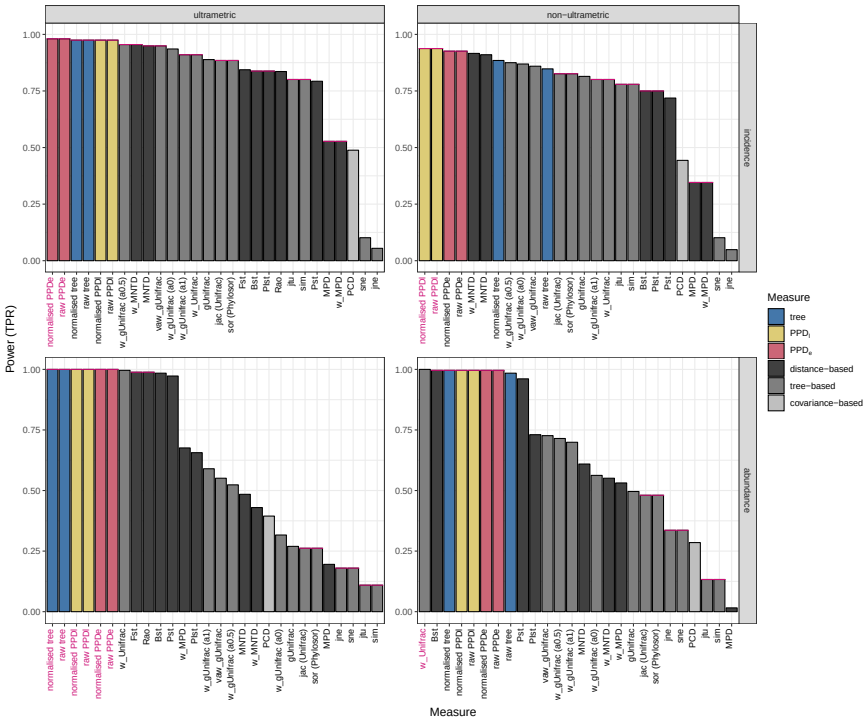


Figure 4.16: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from *100-tip* phylogenies (Experiment 3): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

4.5.4 Experiment 4: Multiple subcommunities

In this experiment, I investigate how well each measure is able to identify community structure when a metacommunity is divided into 2, 4, and 8 subcommunities. Figure 4.17 shows how each measure of diversity performs in the two subcommunity case. This is the default for all other experiments, since many traditional measures are designed for this purpose. As observed in Section 4.5.1, the new phylogenetic beta diversity measures derived in this chapter, particularly PPD_l and PPD_e have the highest power, though PCD, $MNTD$, and w_MNTD for incidence-based data and B_{st} and P_{st} for abundance-based data also have high power.

Increasing to four subcommunities, for non-ultrametric data, and abundance-based ultrametric data, JAC, JTU, SOR, and SIM, the multiple (more than two) subcommunity versions of *jac*, *jtu*, *sor*, and *sim* increase in power. For abundance-based data and ultrametric incidence-based data, *tree*, PPD_l, and PPD_e, as well as B_{st}, P_{st}, and PI_{st}, either increase or reach a maximum power of 100%. On the other hand, CD and $aCDNT$, which quantify MPD and $MNTD$ for more than two subcommunities, are comparable in power for abundance-based data, but drop considerably in power for incidence-based data.

When the number of subcommunities is increased to 8, similar patterns are observed for abundance-based ultrametric data. The new measures, *tree*, PPD_l, and PPD_e, as well as B_{st}, P_{st}, and PI_{st} have 100% power. For incidence-based data, there is a general decrease in power for all measures, though *tree*, PPD_l, and PPD_e continue to have the best power.

Summary: for large numbers of subcommunities and abundance-based data, all of the new phylogenetic beta diversity measures, *tree*, PPD_l, and PPD_e, as well as B_{st}, P_{st}, and PI_{st} are best able to detect phylogenetic signal from community structure; whereas for incidence-based data, PPD_e performs best.

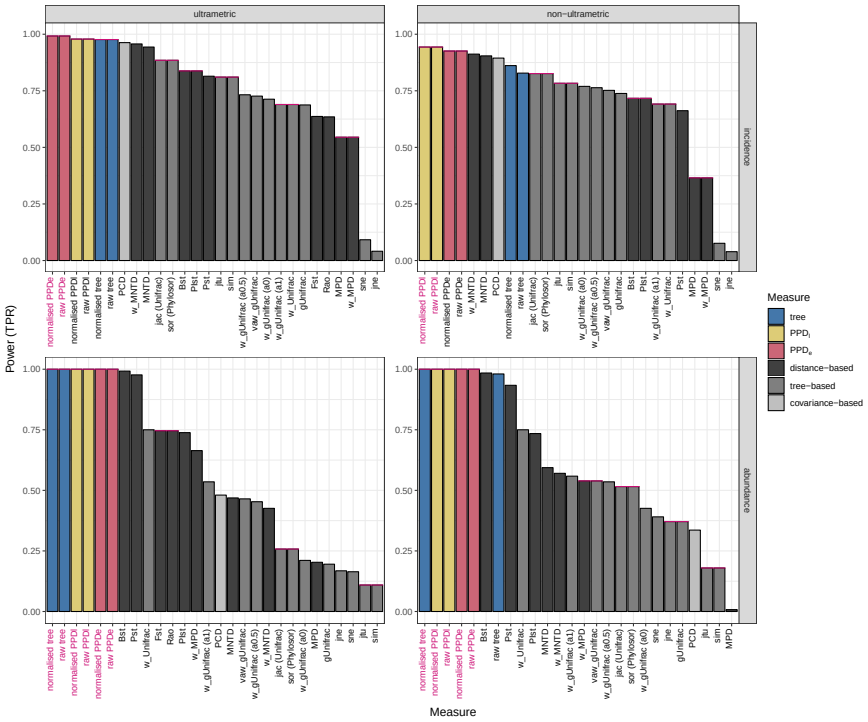


Figure 4.17: Bar chart comparing the power of measures of phylogenetic beta diversity to detect phylogenetic signal from a *2-subcommunity* assemblage (Experiment 4): Power is calculated for all subcommunity partitioning strategies.

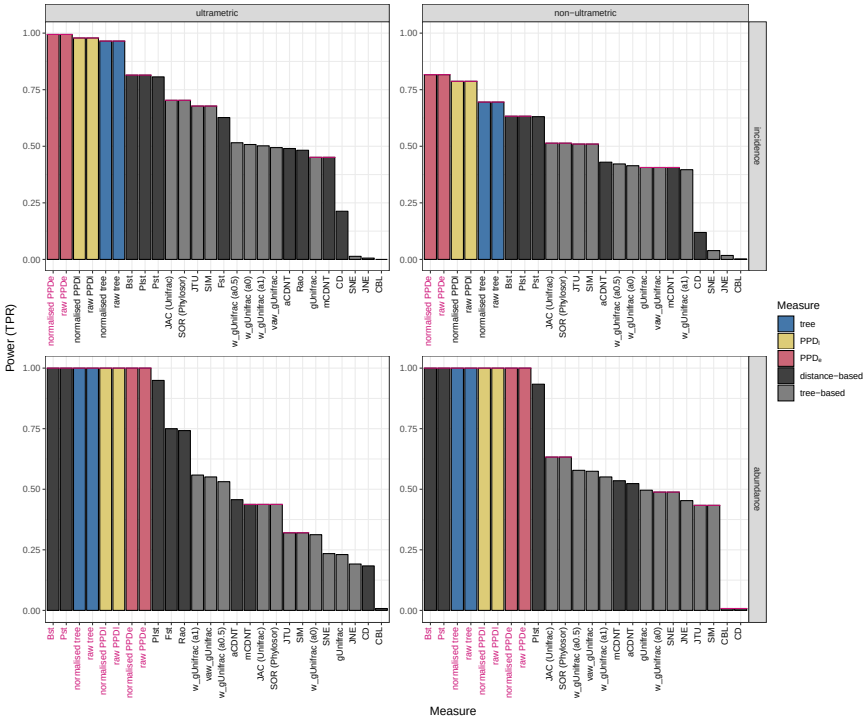


Figure 4.18: Bar chart comparing the power of measures of phylogenetic beta diversity to detect phylogenetic signal from a *4-subcommunity* assemblage (Experiment 4): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

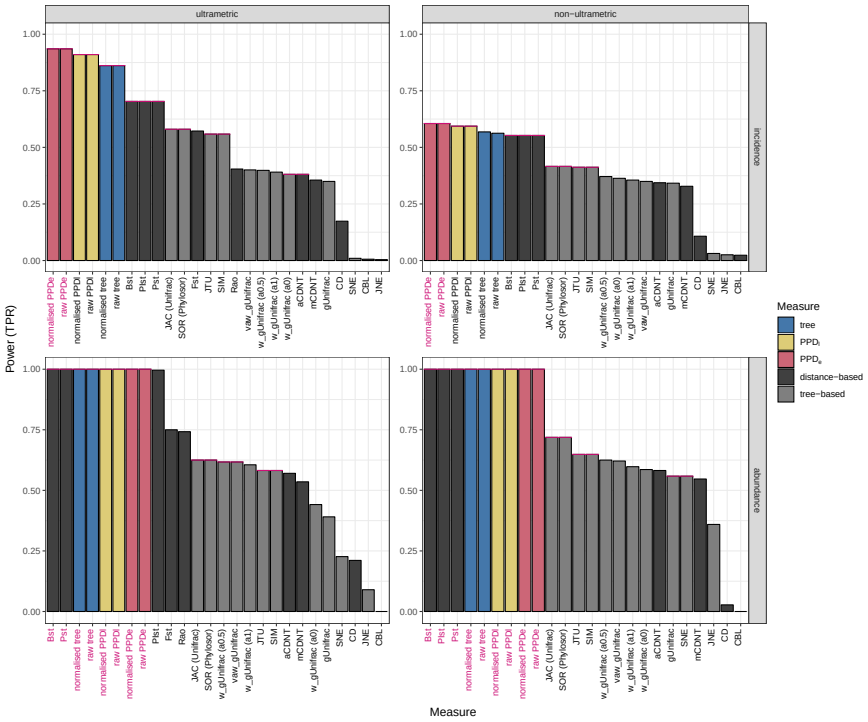


Figure 4.19: Bar chart comparing the power of measures of phylogenetic beta diversity to detect phylogenetic signal from an *8-subcommunity* assemblage (Experiment 4): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

4.5.5 Experiment 5: Varying evolutionary rates

This experiment examines how well each measure of phylogenetic beta diversity is able to detect community structure derived from different evolutionary processes. Since the qualitative and quantitative models use different model parameters, these are analysed separately.

4.5.5.1 Qualitative data

Qualitative partitioning schemes (qd_u, qd_nu, qi_u, and qi_nu) are derived from discrete models based on predefined transition rates between traits (subcommunities a and b). As a default, in all other experiments, qualitative transition rates are set to 0.02. In this experiment, the rate (r_l) is decreased through 1, 0.2, 0.1, 0.02, and 0.01, to investigate how traditional measures compare with those derived from Reeve et al.'s (2016) framework in their ability to detect phylogenetic signal in community structure. Figures B.1 and B.2 show examples of community structure derived along ultrametric trees for all transition rates, for qualitative-dependent (qd_u) and qualitative-independent (qi_u) partition strategies, respectively.

At the highest transition rate ($r_l = 1$), all measures have a power of around 10% or less for all partitioning strategies (Figure 4.20), which is barely above chance (the false positive rate, α, which is set at 0.05). This is because the switching rate is so high that the resultant metacommunity is well mixed between subcommunities and therefore shuffling the tips has minimal effect on these results.

When the transition rate drops to 0.2 (Figure 4.21), the power for the ultrametric tests rises above chance for the new measures and a few others, though it is still less than 25%. At $r_l = 0.1$, PPD_l, PPD_e, $gUnifrac(a0)$, and PCD have power greater than 25% for ultrametric trees, but all measures continue to perform poorly when attempting to detect structure from non-ultrametric trees (Figure 4.22).

However, when the transition rate is increased to 0.02, a considerable jump in power is observed across all measures (Figure 4.23), except for *sne* and *jne*, which is desirable since there is no nestedness in this experiment. It is clear that PPD_l and PPD_e are best at detecting community structure, particularly for non-ultrametric trees. Amongst the remaining measures, PCD, $MNTD$, and w_MNTD also perform well. This is confirmed when $r_l = 0.01$, as these measures begin to plateau at 100% power for both ultrametric and non-ultrametric cases (Figure 4.24).

Summary: when power is high enough that a measure is able to detect phylogenetic signal from community structure, PPD_l and PPD_e are best able to detect phylogenetic signal from community structure.

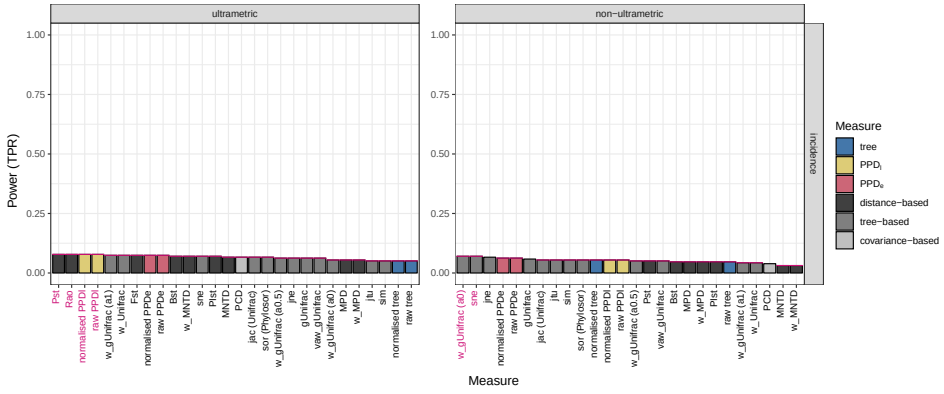


Figure 4.20: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from qualitative data with a *transition rate of 1* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

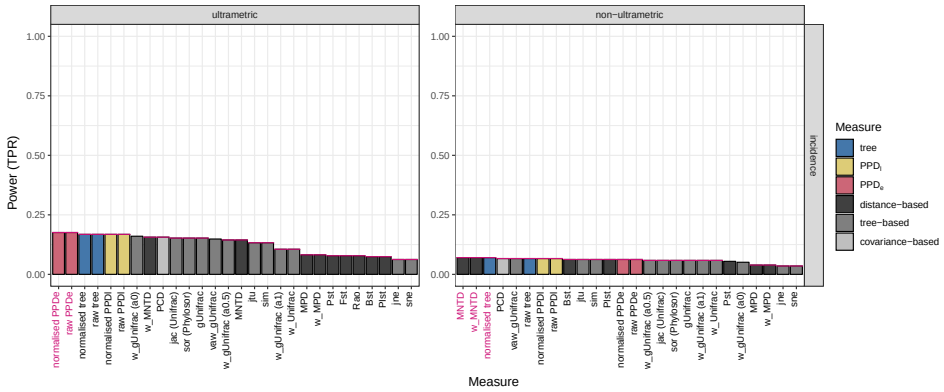


Figure 4.21: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from qualitative data with a *transition rate of 0.2* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

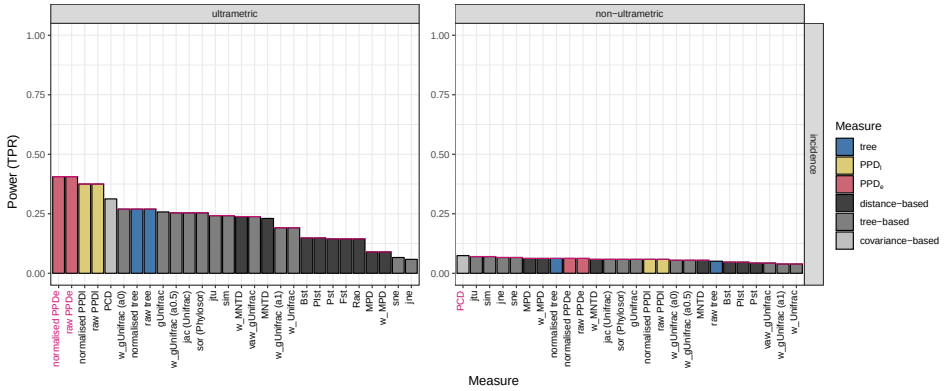


Figure 4.22: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from qualitative data with a *transition rate of 0.1* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

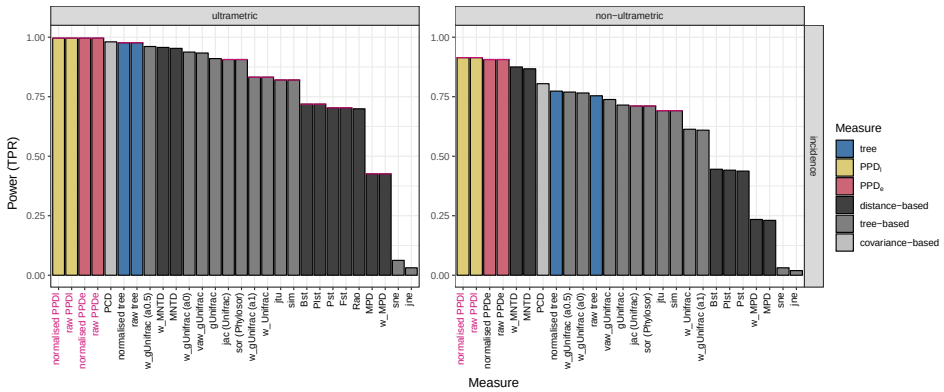


Figure 4.23: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from qualitative data with a *transition rate of 0.02* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

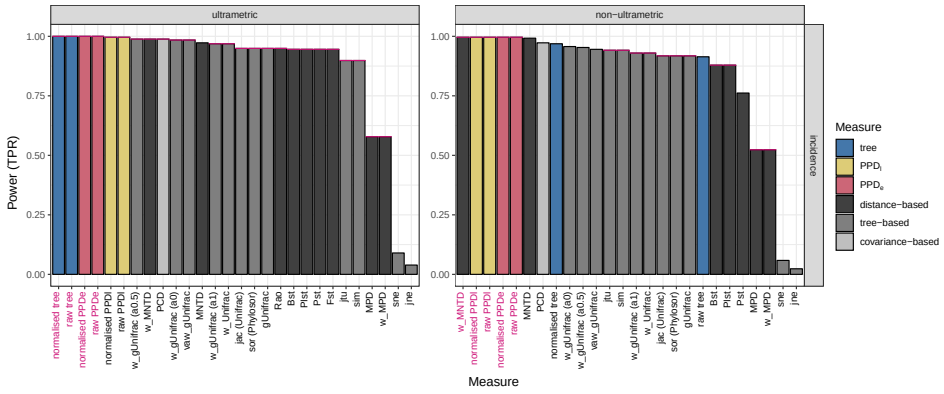


Figure 4.24: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from qualitative data with a *transition rate of 0.01* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

4.5.5.2 Quantitative data

Quantitative partition schemes (both ultrametric: qdb_u, qib_u, qdp_u, and qip_u; and non-ultrametric: qdb_nu, qib_nu, qdp_nu, and qip_nu) are derived using a Brownian motion model, with an evolutionary scaling factor, where high values result in a low evolutionary rate. In all other experiments, this scaling parameter is set to 0.2; however in this experiment, it is decreased through 0.002, 0.02, 0.2, and 2. Example metacommunities are shown in Figures B.3 and B.4 for quantitative traits evolved along ultrametric trees for all evolutionary scaling factors, for quantitative-dependent proportional (qdp_u) and quantitative-independent proportional (qip_u) partitioning strategies, respectively.

At the highest rate of evolution for incidence-based data, power is generally poor for all measures. However, $tree$, PPD_l, PPD_e, P_{st}, B_{st}, and $w\_Unifrac$ all have a power of greater than 50%, as do Rao and F_{st} for ultrametric trees (Figure 4.25). As before, for qualitative partition schemes, measures are better able to detect community structure at lower evolutionary rates. Therefore, decreasing the evolutionary scaling factor to 0.02 results in an increase in the power of all measures, except jne and sne, which are expected to remain low (Figure 4.26). However, for abundance-based data, $tree$, PPD_l, PPD_e, as well as $w\_Unifrac$, B_{st}, and P_{st}, and for ultrametric trees, Rao and F_{st}, all have very high power.

As the evolutionary rate continues to decrease through 0.2 and 2, for incidence-based data, the power of the best measures (including the new measures) fluctuates slightly, whilst the power of the remaining measures systematically increases, but remains low (Figures 4.27 & 4.28). For abundance-based data the high powered measures remain high powered, whilst the weaker measures begin to improve.

Summary: power is highest for abundance-based data, where across all evolutionary rates, all of the new phylogenetic beta diversity measures, $tree$, PPD_l, and PPD_e are best able to detect phylogenetic signal from community structure.

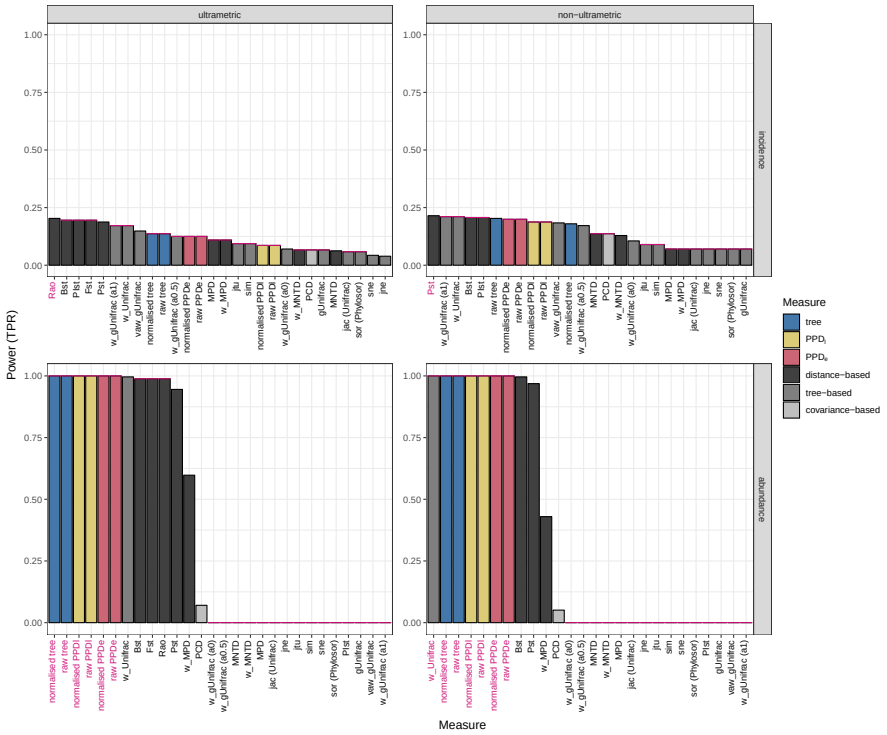


Figure 4.25: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from quantitative data with an *evolutionary scaling factor of 0.002* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

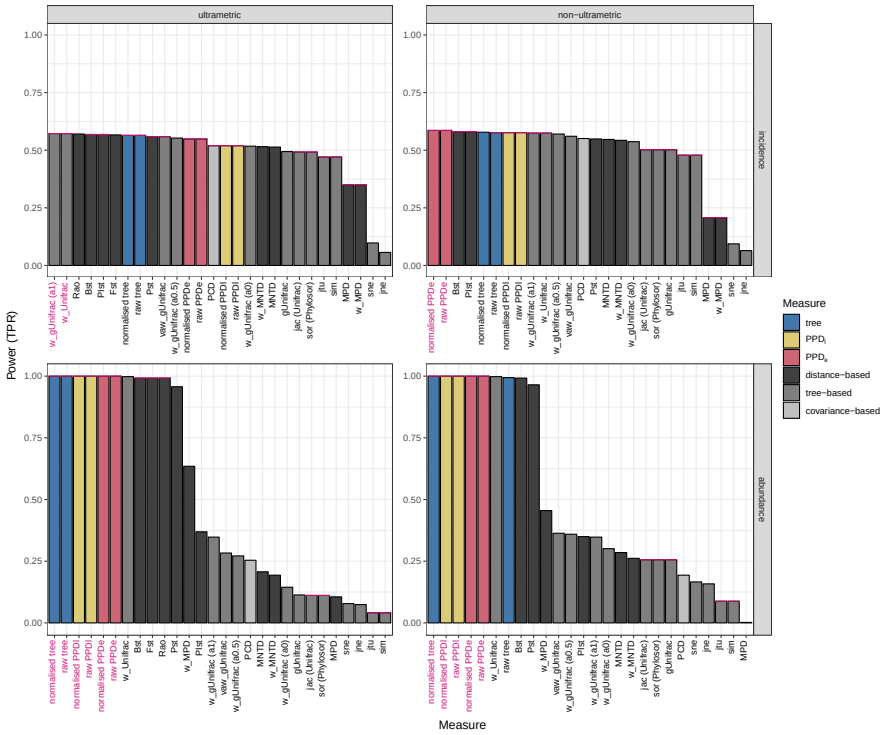


Figure 4.27: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from quantitative data with an *evolutionary scaling factor of 0.2* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

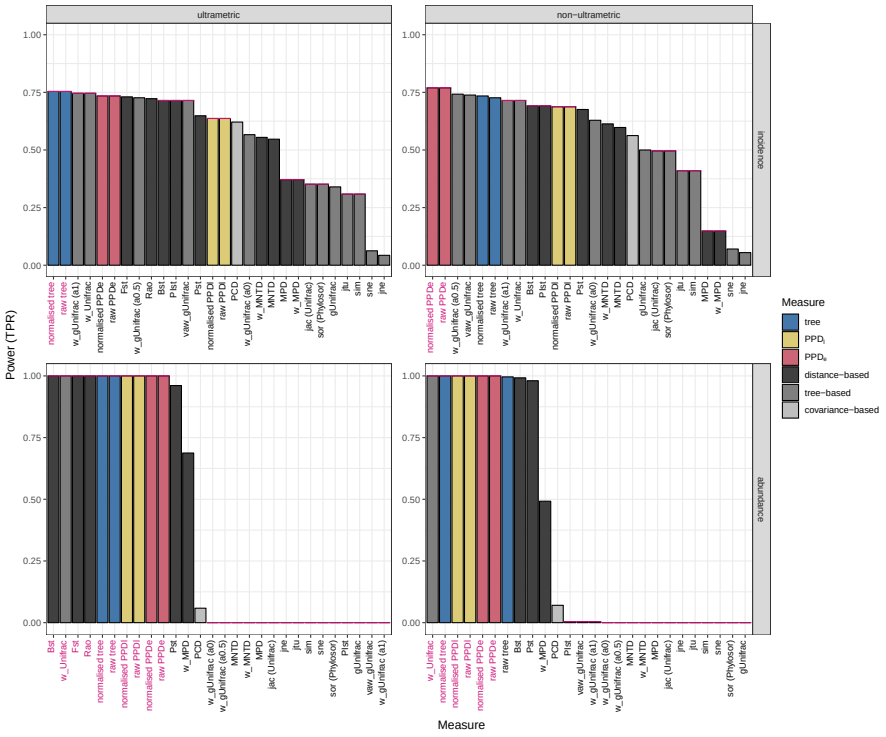


Figure 4.26: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from quantitative data with an *evolutionary scaling factor of 0.02* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

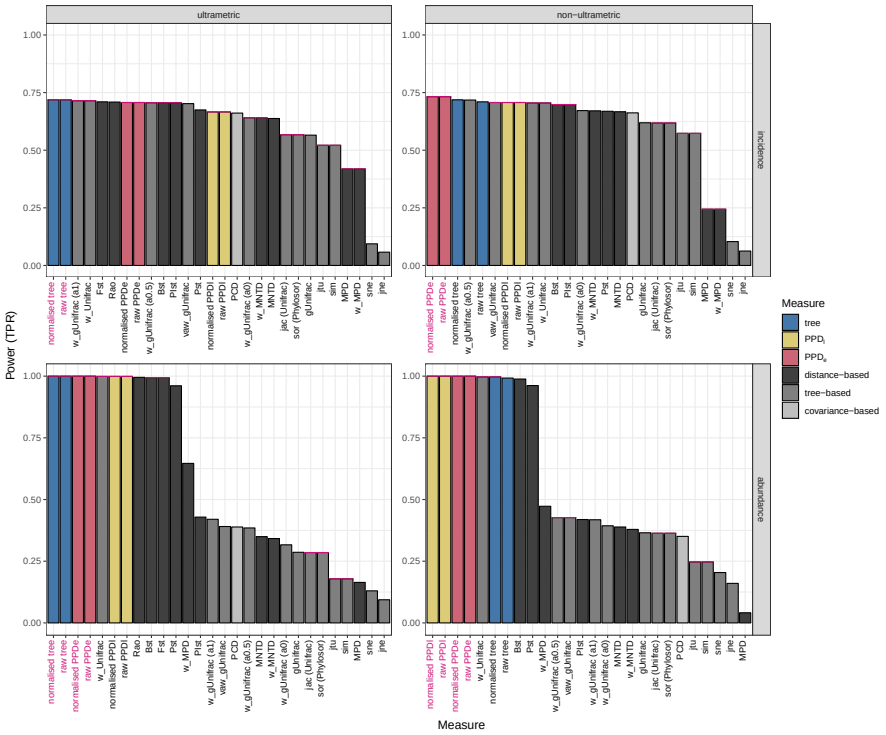


Figure 4.28: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure from quantitative data with an *evolutionary scaling factor of 2* (Experiment 5): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

4.5.6 Summary

The power of each measure is summarised by averaging across all experiments. These results are shown in Figure 4.29. It is clear that the new measures, extending Reeve et al. (2016) (*tree*, PPD_l and PPD_e) are better able to detect phylogenetic signal in community structure than than those commonly used in the literature, particularly PPD_e, which has the highest power in all categories.

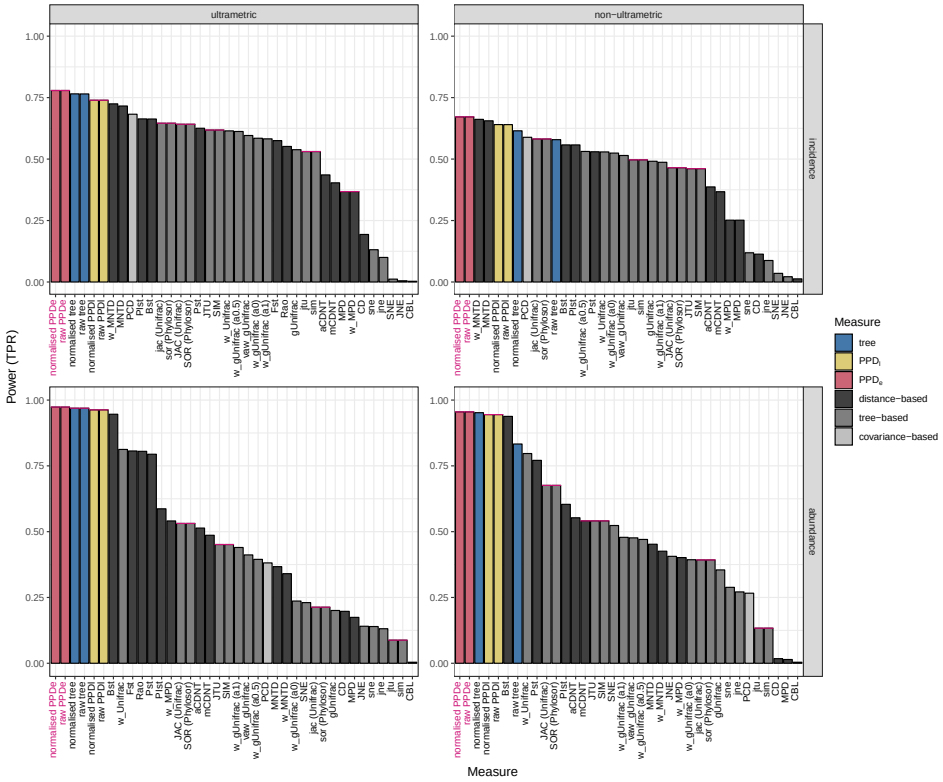


Figure 4.29: Bar chart comparing the power of measures of phylogenetic beta diversity to detect subcommunity structure averaged across all experiments: Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

4.6 Case study: Phylogenetic diversity of AMR in sympatric host populations

4.6.1 Introduction

Following from the previous section, this case study examines whole sequence data from a representative sample of sympatric populations of human and animal hosts in Scotland. Following Mather et al. (2014), I evaluate the interspecies transmission of AMR in a representative sample of *Salmonella* DT104 isolates to determine whether or not extensive transmission of epidemic strains occurred across host populations. Data used in this study were Scottish samples, collected between 1990-2011, from sympatric and contemporaneous populations of human and animal hosts.

Mather et al. (2014) used Bayesian phylogenetic diffusion models to reconstruct the host population of each branch in the phylogenetic tree, whilst estimating the number of branch-to-branch transitions using Markov jumps (Figure 4.30). Their results showed a low number of between-host transitions (human-to-animal and animal-to-human), accompanied by a high number of within-host transitions (animal-to-animal and human-to-human); suggesting that either (1) DT104 was circulating separately within each host population, or that (2) human and animal hosts were each sinks for a different and separate source of infection, with a small degree of spill-over in both directions. To validate these results, new diversity measures are used to examine the degree of mixing between host phylogenies and investigate whether or not there was extensive transmission of DT104 between human and animal populations during the epidemics.

Acknowledgements: The phylogenetic antimicrobial resistance data was kindly provided by Alison Mather.

4.6.2 Methods

The existing phylogeny (Figure 4.30A) was derived from whole genome sequence data comprising 135 human- and 113 animal-origin isolates (Mather et al., 2014). The isolates were divided into two subcommunities, of human and animal origin respectively. In order to observe the degree of mixing between the resultant human and animal phylogenies, the effective number of distinct phylogenies, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, was calculated for a range of q values, between 1968 and 2001 (see Sections 2.2.6.3 & Section 4.3.2.6). The reason why *tree* measures are used rather than one of the PPD_e measures – which performed best in the previous study – is that these measures allow the pre-epidemic period to be excluded using T. The null distribution was simulated at a 95 % confidence interval (grey bar) by fixing the phylogenetic tree, randomly relabelling each tip 10000 times, and calculating ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$ for all values of q.

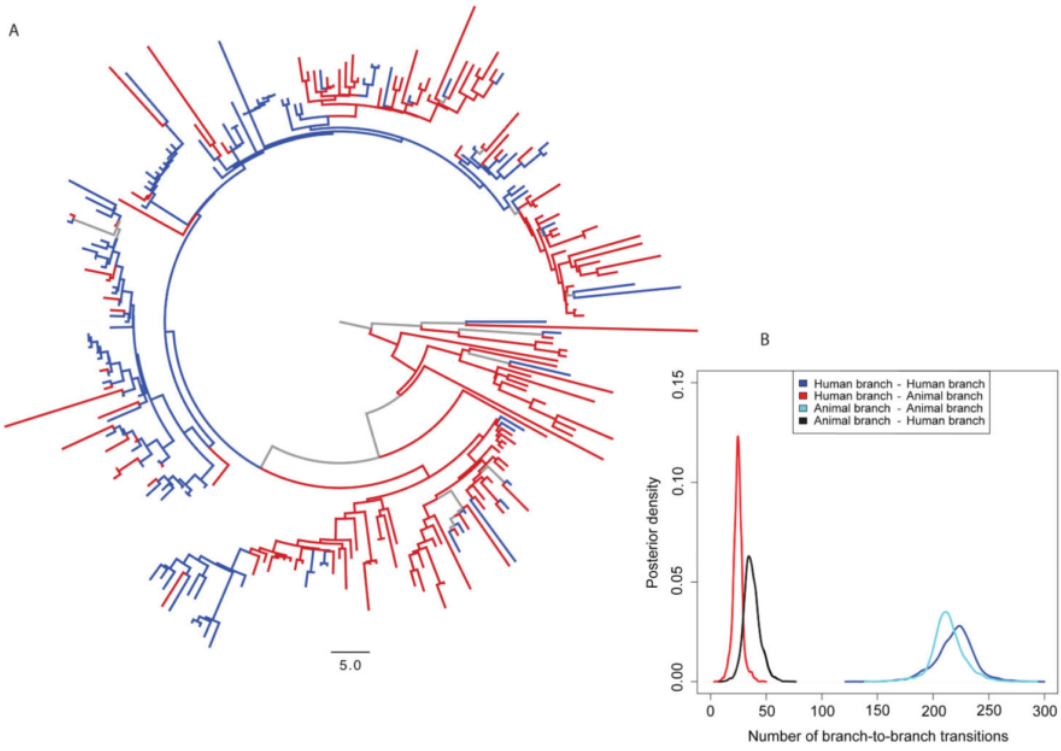


Figure 4.30: Phylogenetic analysis of *Salmonella Typhimurium* DT104 in Scotland, taken from Mather et al. (2014): (a) Bayesian phylogenetic diffusion models were used to reconstruct the host population of each branch; animal isolates are coloured blue, and human isolates are coloured red. (b) Markov jumps estimate the number of human-to-animal and animal-to-human transitions within the phylogeny.

4.6.3 Results and discussion

Results show that human and animal phylogenies are not well mixed at almost all scales, for $q \leqslant 25$ (${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, Figure 4.31a). These findings agree with previous conclusions, that DT104 epidemics are broadly distinguishable within each host population, whilst not requiring inference of the host of putative ancestral isolates.

Fixing the phylogeny and 'cutting the tree' (varying T in Section 4.3.2.6) allows snapshots of the evolutionary history to be extracted. In this way, information regarding the transmission of DT104 between host populations may be recovered. When an interval is defined (from 1991 – 2011) to exclude evolutionary history prior to the epidemic (Figure 4.32), the effective number of phylogenies increases from 1.084 to 1.402, suggesting that the host phylogenies are distinctly not well mixed, where $p < 0.05$ for $q < 16$ (Figure 4.31b).

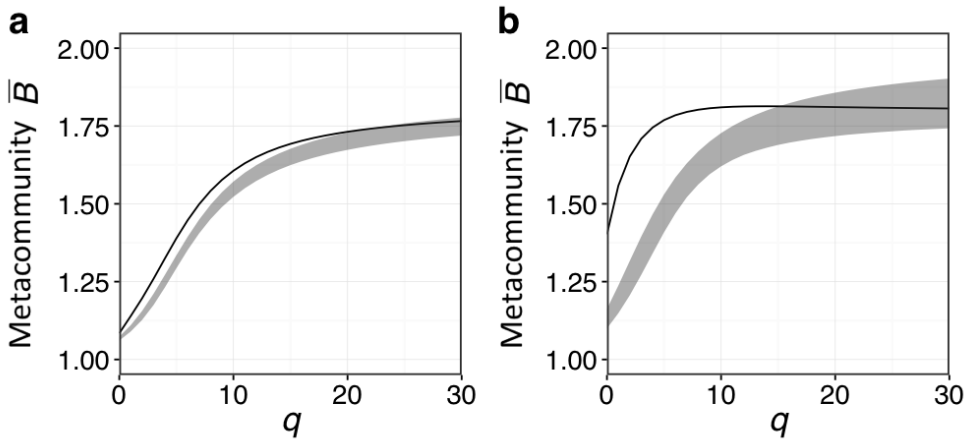


Figure 4.31: Phylogenetic diversity of a representative sample of *Salmonella* DT104: Metacommunity $\bar{B}$ is calculated for $0 \leqslant q \leqslant 30$, yielding the effective number of phylogenies between (a) 1968–2011 and (b) 1991–2011. Shaded ribbons simulate the null distribution within a 95 % confidence interval, where $p < 0.05$ for $q < 25$ and $q < 16$, respectively.

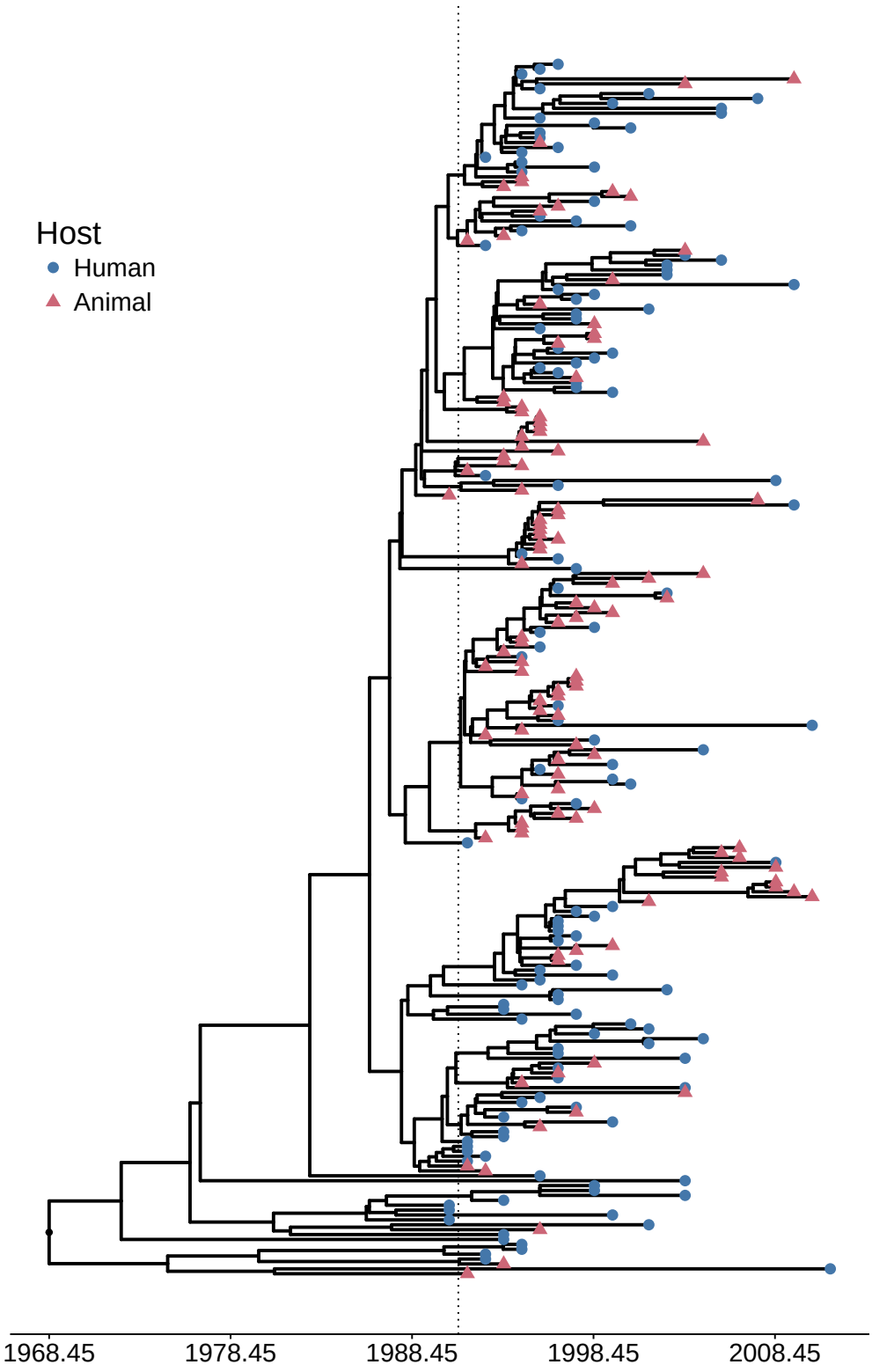


Figure 4.32: Phylogeny of a representative sample of *Salmonella* DT104: Taken from sympatric human (blue) and animal (red) isolates, between 1968 – 2011. The dotted line denotes 1991, the start of the epidemic.

4.7 Conclusion

Traditional measures of phylogenetic diversity can be broadly split into two categories: tree-based and distance-based measures. In this chapter, I developed new methods – extensions of Reeve et al.'s (2016) framework – to quantify phylogenetic diversity based on both of these perspectives. I then investigated the robustness of these measures by comparing them to those commonly used in the literature. This analysis was conducted by simulating phylogenies, evolving traits representing subcommunity preferences along them, and using these traits to define subcommunity structure. A power analysis was then used to determine how well each measure was able to identify phylogenetic signals in community structure.

Overall, results indicate that measures are best able to detect phylogenetic signal from community structure when data is abundance-based rather than incidence-based. Particularly when subcommunities are nested, phylogenies are small, or the evolutionary transition rate is low. Results showed that the measures of phylogenetic beta diversity based on Reeve et al.'s (2016) framework outperformed those commonly used in the literature under nearly all circumstances. In future work, it should be possible to examine the mathematical properties of these measures in order to determine what aspect of data they each focus on, and how this relates to the differences in these results. In particular, exponentially transformed phylogenetic distance-based beta diversity measures (PPD_e) performed best overall (detecting phylogenetic structure from community structure for nestedness vs. turnover, different tree sizes, different numbers of subcommunities, and different evolutionary rates) in all four identified categories (incidence-based vs. abundance-based data, and ultrametric vs. non-ultrametric trees). It seems sensible, therefore, to suggest using a permutation test on ${}^{1}\bar{B}^{\mathbf{Z}}{}_{PPD_e}$ to investigate the presence of a phylogenetic signal in this kind of dataset, since the best results were generally obtained for $1 \leq q \leq 2$ and determining the effective number of distinct communities (${}^{q}\bar{B}^{\mathbf{Z}}$) most closely aligns with the question of whether a phylogenetic signal exists in the subcommunity structure (and all PPD_e beta diversity measures were broadly equivalent). Note, however, that although these measures are better at identifying nestedness and turnover than specific measures used for this purpose, they cannot distinguish between them. So, if the research question requires these two types of change to be distinguished from one another, I would suggest identifying the presence of change using our measures, and then attempting to distinguish between them using the (much less powerful) specific tools for this purpose.

Transmission of antimicrobial resistance was then re-investigated in terms of shared evolutionary history. By fixing the phylogeny and 'cutting the tree', snapshots of evolutionary history are extracted to recover information surrounding the transmission of DT104 between host populations. Evolutionary relationships are explored by quantifying the phylogenetic diversity (*tree*) of a representative sample of animal-origin and human-origin isolates. Information surrounding the transmission dynamics during only the epidemic period are revealed by reducing the amount of evolutionary history in this analysis. These results validate the conclusions of

the previous study and suggest that epidemic strains of DT104 and its resistance genes were maintained separately in human and animal populations.

In conclusion then, the robustness of these new measures has been demonstrated against a variety of different phylogenies, with different numbers of tips, different evolutionary rates, different numbers of subcommunities, different tree structures (ultrametric vs. non-ultrametric), and nestedness vs. turnover. Furthermore, their utility has been showcased by validating against known results of AMR transmission in sympatric populations.

This chapter illustrated that with reduced tree sizes (*i.e.* lower amounts of data), the ability to detect a signal dropped dramatically for all diversity measures, and although our measures performed as well as any others. Nonetheless it raises the question of how well this family of diversity measures performs under reduced sampling intensity. In the next chapter, I investigate this question in detail with a fully sampled dataset, investigating how well Reeve et al.'s (2016) beta diversity measures are conserved under repeatedly subsampling.

Chapter 5

Sampling properties of the framework

> 'Researchers measuring beta diversity have rarely concerned themselves with the problems of how complete the species lists of studied communities are, and of how the varying degrees of completeness can actually change estimates of beta diversity. [...] a situation which is more common than usually recognised.'
>
> — Cardoso et al. (2009)

5.1 Abstract

Measures of beta diversity or community dissimilarity (compositional heterogeneity, differentiation between subcommunities, turnover, nestedness, and distance) are commonly used to describe changes in species composition between subcommunities. The utility and interpretation of these measures is important across applications ranging from conservation management to the study of viral systems. However, even when a suitable metric is selected, incomplete data resulting from low sampling effort is a common problem. Without a complete inventory of the assemblage being studied, perfect comparisons between subcommunities are impossible. As a consequence of this, various estimators have been developed (*e.g.* Colwell & Coddington (1994) for gamma diversity and Chao et al. (2005) for beta diversity), but these rely on certain assumptions being met and are not always accurate, especially for extreme undersampling (Chao et al., 2005). In general, moreover, estimators do not exist for beta diversity measures as they do for gamma diversity. Therefore, it is important to understand how the performance of the measures themselves is affected by subsampling.

The aim of this chapter is to investigate the performance of Reeve et al.'s (2016) measures of diversity under decreased sampling effort. The Barro Colorado Island (BCI) Forest dynamics plot is a fully sampled 50 ha study site (Condit, 1998; Hubbell et al., 1999) and is therefore ideal for this purpose – that is, to consider *how well diversity measurement is conserved during incomplete sampling*. Questions that arise are whether it is better to partially sample every subcommunity or fully sample only a few subcommunities, and whether some measures are more robust under subsampling than others. It was found that reducing sampling effort at every site was a worse strategy than reducing the number of fully sampled sites, and in the latter case, beta diversity measures were generally well conserved.

5.2 Introduction

Sampling methods

Species richness (described in Section 1.2.1) is one of the most popular measures of biological diversity, being mathematically intuitive and easily interpretable. Unfortunately, being equally sensitive to rare species as it is to common ones, it is strongly affected by sampling issues. As well as this, the observed species richness of a sampled assemblage depends on many factors, including: the number of individuals in the area, how easily these individuals can be observed, the size of the region being studied, and its heterogeneity.

Typically, sampling methods are either: individual-based, where a number of individuals are randomly sampled from within each quadrat; or sampled-based, where for a set of sampling units (*e.g.* traps, quadrats, nets, or a period of time) within the study site, every individual is recorded. Sample-based methods preserve the spatial (or temporal) structure of the data, which may reflect the aggregation or segregation[1] of species among samples. Consequently, fewer species will be represented by these individuals than by an equal number of individuals sampled randomly within the same habitat (Gotelli & Colwell, 2001). Though examining the benefits of each sampling method is beyond the scope of this chapter, I assess the robustness of Reeve et al.'s (2016) framework of similarity-sensitive diversity measures under decreased sampling effort using both of theses methods.

Beta diversity

Diversity indices, particularly measures of beta diversity or community dissimilarity (and their similarity complements), are numerous in the literature (Tuomisto, 2010a). Some of the most widely used are those that describe compositional similarity, such as the Jaccard (1901) and Sørensen (1948) indices, and those derived from the additive or multiplicative partitioning of alpha and gamma diversities, such as Jost's (2007) 'true beta' (discussed in Section 1.6). Though appropriate choice of metric is of fundamental importance, data quality issues such as undersampling may affect the measurement and interpretation of results.

Consider the following example – two communities that both share a number of rare species. When both communities are fully sampled, beta diversity (or some measure of community dissimilarity) is accurately measured. However, when sampling effort is low, if each rare shared species is sampled in only one of the subcommunities, this beta diversity may be artificially inflated (communities appear more dissimilar). Conversely, if each subcommunity contains some common shared species, but many rare species that are unique to each community, then beta diversity may be underestimated (communities appear more similar) when sampling effort is low, as the rare unique species are likely to be missed. Though both scenarios are possible, Chao et al. (2005) and others note that beta diversity typically increases with decreased sampling effort, since "rarity (either in nature or because of small sample size) increases the

[1] Known as patchiness, heterogeneity, or autocorrelation

chance that a species will be spuriously absent from one sample but not from the other, thus negatively biasing similarity indices".

Despite this assertion, surprisingly few studies have considered the robustness of these beta diversity measurements to subsampling, something Cardoso et al. (2009) suggests might be "a consequence of the general lack of agreement about which index of beta diversity should be used in the first place". In response to this, Cardoso et al. (2009) test the effect of subsampling on the performance of 15 incidence-based dis(similarity) measures – which they describe as being "equivalent to beta diversity" – between pairs of subcommunities. They conclude that no index is able to perform without bias in all circumstances[1] (though indices that are insensitive to changes in species richness between communities are most robust). Beck et al. (2013) advance this work, examining the effect of undersampling on 14 measures of beta diversity and compositional dis(similarity). They confirm that decreased sampling effort more often resulted in overestimates than underestimates, whilst also noting a reduction in precision; Likewise, Plotkin & Muller-Landau (2002) show that local clustering of conspecifics reduces the similarity between sampled communities and increases the variance of similarity indices. Furthermore, they observed that beta diversity measures that are more sensitive to species of greater abundance are more robust to incomplete sampling. See also Morisita (1959); Wolda (1981); Ricklefs & Michael (1980) for early work examining the effect of reduced sampling effort on indices of beta diversity and Tuomisto (2010b) for a discussion of how different diversity components are affected by incomplete sampling.

The main focus of this chapter is to determine the robustness of Reeve et al.'s (2016) framework of diversity measures under reduced sampling effort. This is done using different sampling strategies (sampling individuals versus sampling entire subcommunities) to investigate whether existing problems with the estimation of beta diversity from incomplete data can be mitigated using this framework. I also investigate, whether these new beta diversity measures typically increase with decreased sampling effort when individuals are sampled from a community (*sensu* Chao et al., 2005) and whether those that are more sensitive to species of higher abundance are more robust to reduced sampling effort (*sensu* Beck et al., 2013).

5.3 Methods

Dataset

The Barro Colorado Island (BCI) Forest dynamics plot (Condit et al., 2012b, 2017a) is a fully sampled, 50 ha study site, which makes it ideal for this purpose. The BCI dataset is described in Section 3.2. Data was taken from the first inventory survey (recorded during 1981/82), and the 50 ha study site was subdivided into 1250 20 m × 20 m quadrats. The mean species

[1] Perhaps this is not surprising, since each of the indices tested quantify very different phenomenon, as noted by Tuomisto (2010b)

abundance per quadrat is shown in Figure 5.1 and the total incidence of species across all quadrats is shown in Figure 5.2.

Subsampling and calculating diversity

The dataset was subsampled increasingly sparsely (at 100%, 50%, 25%, 12.5%, 6.25%, and 3.125%), using two different approaches: sampling by individual, where a proportion of individuals were sampled from within each subcommunity; and sampling by subcommunity, where entire subcommunities were sampled from the BCI study site as a whole. These methods simulate *individual-based* and *sample-based* data, respectively, as discussed in Gotelli & Colwell (2001). See Tables 5.1 and 5.2 for summaries of sampling structure.

Subcommunity diversity was calculated in the naïve-type case (ignoring species similarity) for $q \in \{0, 1, 2, \infty\}$, for each subsampled (as well as the fully sampled) metacommunity. The measures calculated were: *representativeness* (${}^q\bar{\rho}_j^{\mathbf{I}}$), *redundancy* (${}^q\rho_j^{\mathbf{I}}$), an *estimate of the effective number of distinct subcommunities* (${}^q\bar{\beta}_j^{\mathbf{I}}$), *distinctiveness* (${}^q\beta_j^{\mathbf{I}}$), the *contribution per tree toward metacommunity diversity* (${}^q\gamma_j^{\mathbf{I}}$), the *effective number of tree species* (${}^q\bar{\alpha}_j^{\mathbf{I}}$), and an *estimate of naïve-community metacommunity diversity* (${}^q\alpha_j^{\mathbf{I}}$). These were also calculated at the metacommunity levels: ${}^q\bar{R}^{\mathbf{I}}$, ${}^qR^{\mathbf{I}}$, ${}^q\bar{B}^{\mathbf{I}}$, ${}^qB^{\mathbf{I}}$, ${}^qG^{\mathbf{I}}$, ${}^q\bar{A}^{\mathbf{I}}$, and ${}^qA^{\mathbf{I}}$, respectively. Each measure was assessed a number of different ways, to investigate:

- How well is measured diversity preserved under subsampling?
- How good is the correlation between fully sampled and subsampled diversity measures?

And since these measures calculate beta diversity individually for each subcommunity:

- How well is order preserved between subcommunities? That is, are the most important communities identified?

Each of the measures were calculated (a) for the complete census and (b) after subsampling; these two values were then compared to establish how well each measure was conserved under subsampling. An understanding of subsampling properties is valuable to understand how accurate these measures can be expected to be when sampling is incomplete, as typically occurs in real ecological applications. Intuitively, because subcommunity alpha (${}^q\bar{\alpha}_j^{\mathbf{I}}$ and ${}^q\alpha_j^{\mathbf{I}}$) relies only on information from the subcommunity of interest, it is only affected by reductions that affect that subcommunity directly. In this case, when sampling by subcommunity, subcommunity alpha should be maintained for all retained communities; when sampling by individual, it should only be affected by subsampling that targets the subcommunity of interest. On the other hand, subcommunity gamma (${}^q\gamma_j^{\mathbf{I}}$) relies on information taken from the metacommunity as a whole. In this case, when sampling by subcommunity, subcommunity gamma should be well conserved even when sampling effort is low (since, when only 3% of individuals are subsampled 217 of the total 307 species still remain, and 214 species remain

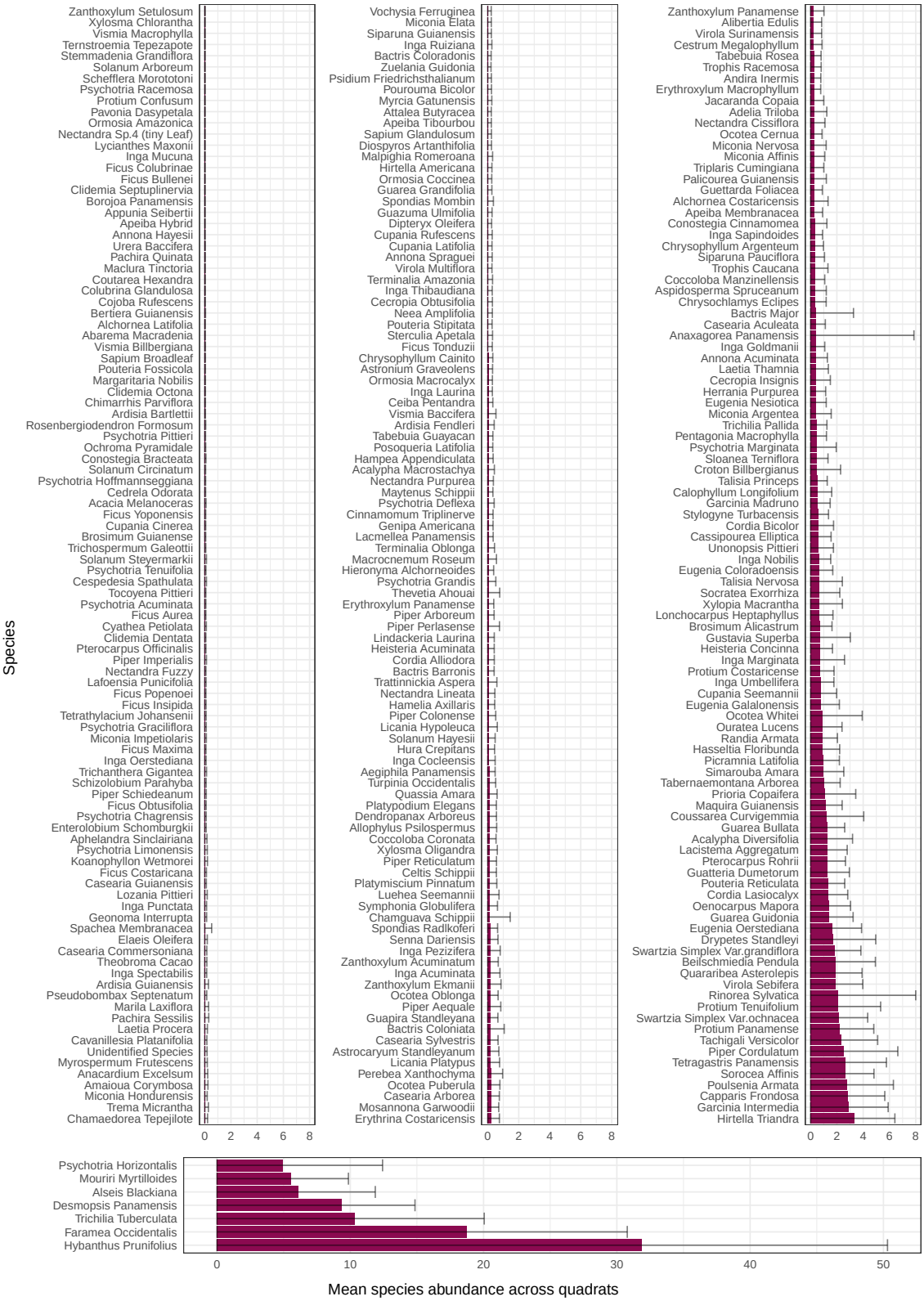


Figure 5.1: Mean species abundance per quadrat from the Barro Colorado Island Forest dynamics plot 1981/82 census: where species are ordered by total abundance and error bars denote standard deviations from the mean.

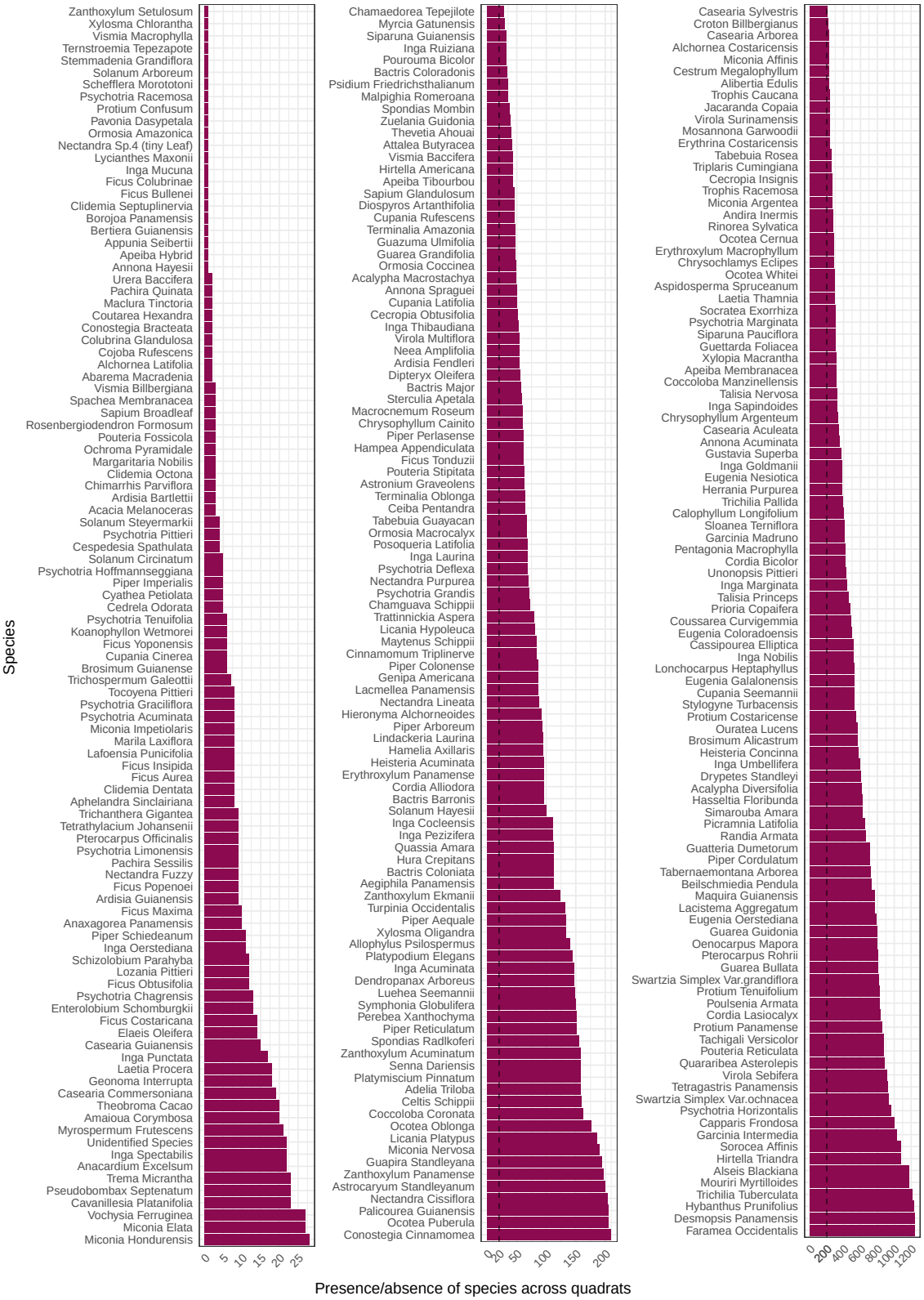


Figure 5.2: Presence/absence of species across 1250 quadrats in the Barro Colorado Island Forest dynamics plot 1981/82 census: where species are ordered by prevalence and the dashed line in the second and third plots show the maxima of the plot to the left.

when subcommunities are subsampled). When sampling by individual, subcommunity gamma is also affected by information taken from the subcommunity, since trees that are rare in the subcommunity may or may not be rare across the metacommunity as a whole. Similarly, measures of subcommunity beta (${}^{q}\bar{\rho}_j^{\mathbf{I}}$, ${}^{q}\rho_j^{\mathbf{I}}$, ${}^{q}\bar{\beta}_j^{\mathbf{I}}$, and ${}^{q}\beta_j^{\mathbf{I}}$) are based on comparisons between the subcommunity and the metacommunity, and therefore rely on information at both levels. Subsampling therefore affects measures of subcommunity beta and gamma via influences on the structure of the subcommunity of interest, and those that affect the metacommunity. In turn, this means that measures of subcommunity beta and gamma should be affected not only by sampling effects in the subcommunity of interest, but also by effects in all other subcommunities, via their effect on the metacommunity.

At the metacommunity level, as might be expected, alpha, beta, and gamma metacommunity diversity measures (${}^{q}\bar{A}^{\mathbf{I}}$, ${}^{q}A^{\mathbf{I}}$, ${}^{q}\bar{R}^{\mathbf{I}}$, ${}^{q}R^{\mathbf{I}}$, ${}^{q}\bar{B}^{\mathbf{I}}$, ${}^{q}B^{\mathbf{I}}$, and ${}^{q}G^{\mathbf{I}}$) should be affected by sampling across all subcommunities, because they are (power mean) averages of the subcommunity measures. When sampling by subcommunity, they are computed across fewer subcommunities; when sampling by individual, they are affected by sampling in each of the constituent subcommunities.

Statistical analysis

To determine how well measured diversity was preserved at low sampling effort, subcommunity diversity values were compared between the subsampled and fully sampled datasets. For robustness, experiments were repeated 100 times and data were pooled. Subsampled diversity was plotted against the true diversity of the study site. Note that here, the term 'true diversity' is used to describe the real, or actual diversity of the fully sampled dataset. Linear regression was used to determine whether any positive or negative sampling bias was present and fit was assessed using R^2 values. For metacommunity-level measures, 'Diversity conserved (%)' was plotted against 'Sampling effort'. For clarity, when diversity is more than 100% conserved, fully sampled values have been over-estimated, and when diversities are less than 100% conserved, fully sampled values have been under-estimated.

To determine how well the ranking of each subcommunity is preserved under reduced sampling effort, subcommunities were ranked by value (of a particular measure of diversity), and the top 5% of values (in the fully sampled dataset) were identified. The extent to which subsampling the data preserves these rankings was then assessed. Receiver operating characteristic (ROC) curves were used to determine how well each measure is able to detect the highest ranking subcommunities under reduced sampling effort. That is, are the top 5% of subcommunities in the subsampled dataset also observed in the top 5% of subcommunities in the fully sampled dataset? ROC curves represent the full trade-off between sensitivity (the true positive rate, or TPR) and 1− specificity (the false positive rate, or FPR) for all possible thresholds of a given test, where the threshold is used as a binary classifier to determine whether results are

positive or negative (example in Figure 5.3). Here, a positive result was recorded when a subcommunity was observed in the top x% of subsampled diversities, where x ranges from 1% to 100%. False positives were detected when a subcommunity was ranked within the top x% of subsampled diversities but not within the top 5% of fully diversities. True positives were detected when a subcommunity was ranked within the top x% of subsampled diversities and the top 5% of fully sampled diversities. ROC curves were generated by plotting TPR against FPR, where a purely random test would result in a diagonal line (TPR = FPR), and perfect discrimination (100% sensitivity and 100% specificity) would result in the ROC curve reaching the upper left corner (where TPR = 1 and FPR = 0).

Area under the curve (AUC) values were calculated where appropriate to determine the overall quality of the test, where 1 is a perfect result and 0.5 is equivalent to chance. Values below 0.5 indicate that the test is doing the reverse of its intended function (*i.e.* identifying 'incorrect' subcommunities), whereas values falling below 0.85 are generally considered to be a poor test. The following measures of beta diversity were investigated: ${}^{q}\rho_j^{\mathbf{I}}$, ${}^{q}\bar{\rho}_j^{\mathbf{I}}$, ${}^{q}\beta_j^{\mathbf{I}}$, and ${}^{q}\bar{\beta}_j^{\mathbf{I}}$; along with ${}^{q}\gamma_j^{\mathbf{I}}$, a novel measure of gamma diversity, and the alpha diversities ${}^{q}\alpha_j^{\mathbf{I}}$ and ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ for completeness. All measures were investigated at both the subcommunity and metacommunity levels.

Results were summarised by plotting the relative sampling accuracy against sampling effort for each value of q, where relative sampling accuracy was calculated as the ratio of subsampled against fully sampled diversities.

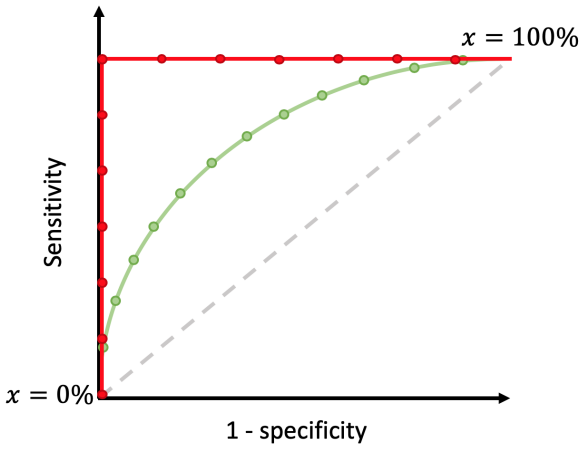


Figure 5.3: Illustration of a ROC curve: where thresholds are marked for an ideal (red) and subsampled (green) scenario.

Table 5.1: Summary of subcommunity structure when subsampling by individual: For each sample fraction the minimum, maximum, median, mean, and standard deviation of the number of individuals per quadrat is tabulated, alongside the total number of individuals in the metacommunity as a whole. Likewise for species count data. All of these values are averaged across 100 repeats.

Sample fraction	Individuals						Species					
	Min	Max	Median	Mean	SD	Total	Min	Max	Median	Mean	SD	Total
1	36	315	187	188.244	40.876	235305	19	85	54	54.051	9.63	307
1/2	18	157	93	93.87	20.435	117338	11.1	61.26	36.92	37.033	7.413	292.77
1/4	9	78	46	46.694	10.224	58367	6.25	41.99	23.945	23.979	5.246	278.24
1/8	4	39	23	23.095	5.115	28869	3.25	26.97	14.67	14.724	3.454	261.32
1/16	2	19	11	11.305	2.57	14131	1.72	16.11	8.47	8.571	2.142	241.38
1/32	1	9	5	5.398	1.301	6747	1	8.76	5	4.661	1.251	217.46

Table 5.2: Summary of subcommunity structure when subsampling by subcommunity: For each sample fraction and total number of subcommunities (N), as well as the minimum, maximum, median, mean, and standard deviation of the number of individuals per quadrat is tabulated, alongside the total number of individuals in the metacommunity as a whole. Likewise for species count data. All of these values are averaged across 100 repeats.

Sample fraction	N	Individuals						Species					
		Min	Max	Median	Mean	SD	Total	Min	Max	Median	Mean	SD	Total
1	1250	36	315	187	188.244	40.876	235305	19	85	54	54.051	9.63	307
1/2	625	41.17	310.09	186.65	188.253	40.853	117657.82	21.58	82.73	53.99	54.028	9.605	291.88
1/4	312	48.87	304.52	186.995	188.518	40.642	58817.66	25.02	80.81	53.91	54.08	9.602	274.82
1/8	156	60.65	296.27	186.985	188.337	40.261	29380.6	27.92	78.88	53.87	54.087	9.615	257.18
1/16	78	74.16	288.51	187.03	188.123	40.421	14673.59	30.35	76.78	53.95	54.146	9.677	237.52
1/32	39	87.74	276.84	186.29	187.266	40.752	7303.39	32.26	74.56	53.64	53.899	9.736	214.44

5.4 Results and discussion

In this section, I investigate each beta diversity measure in turn, followed by gamma and alpha diversities for completeness. These are calculated at both the subcommunity (single quadrat) and metacommunity (full study site) levels under decreasing sampling effort (both sampling by individual and subcommunity).

Sampling by subcommunity: For each diversity measure, I first investigated the effect of sampling a random fraction of subcommunities, both on the calculated diversity of those subcommunities and the overall metacommunity diversity. The first potential problem was that when only a fraction of subcommunities are sampled (for this study I went down to $\frac{1}{32}$, or about 3%), the metacommunity diversity is estimated from only a small fraction of the subcommunities. However, most existing (traditional) beta diversity measures are pairwise comparisons between assemblages and as such, 3% of the subcommunities is equivalent to only 0.1% of comparisons, which is potentially a much worse problem. The second problem (which does not apply to existing beta diversity measures) was that Reeve et al.'s (2016) measures are based on subcommunity to metacommunity comparisons, and these metacommunity distributions are estimates based on the subcommunities selected. However, in a dataset this large, even 3% of subcommunities should produce a reliable estimate of metacommunity diversity.

Sampling by individual: As an alternative, individuals were sampled from within every subcommunity with decreasing sampling effort. This was done to investigate the effect of a different sampling strategy (sampling evenly over the whole study site rather than sampling intensively at a subset of locations), whilst maintaining the same sampling effort. In this case, the problem was that every subcommunity distribution is an estimate based on the individuals selected, as is the metacommunity distribution. For existing (traditional) pairwise beta diversity measures, this problem may potentially, again, be worse because the comparison is between two poorly sampled subcommunities.

5.4.1 Representativeness

Subcommunity ${}^{q}\bar{\rho}_{j}^{\mathbf{I}}$ calculates the representativeness of each subcommunity by comparing the species composition within subcommunity j to that of the metacommunity (described in Section 2.2.6.1). Consider the extreme case, where subcommunity representativeness equals one (its maximum value). Here, the species present in the subcommunity exactly match those present in the metacommunity and, for $q > 0$, the distribution of species must exactly match that of the metacommunity as well. When sampling by subcommunity, the subcommunity diversity measures should be accurate so long as the metacommunity species distribution is, since the subcommunity species distributions are preserved. However, the accuracy of the

metacommunity diversity will also depend on the representativeness of the subcommunities selected. When sampling effort within the subcommunity is decreased (*i.e.* sampling by individual), some rare species are likely to be missed, and so the representativeness of this subcommunity should be underestimated.

5.4.1.1 Sampling by subcommunity

Subcommunity representativeness, ${}^{q}\bar{\rho}_j^{\mathbf{I}}$, is almost perfectly conserved for $q \leq 2$ (though it is more variable at $q = 2$) at all levels of sampling effort (Figure 5.4a). For $q \leq 1$, R^2 is greater than 0.95, even at only 3.125% sampling effort (see Table C.1 for all R^2 values in this chapter), and linear regression (solid black lines on the plots) shows very little systematic bias where R^2 remains high. At higher values of q, where $\bar{\rho}$ focuses more on the least representative species, R^2 drops to 0.79 at $q = 2$ and 0.46 at $q = \infty$ for 3.125% sampling effort.

The ROC curves improve on these results, showing that although subsampling has the least impact on the ability to identify the top 5% of the most representative subcommunities at $q = 0$, all values of q and levels of subsampling accurately identify the most representative subcommunities (Figure 5.4b). Even at 3.125% sampling effort, AUC values are calculated as 0.992, 0.977, 0.972, and 0.897 for $q \in \{0, 1, 2, \infty\}$, respectively (see Table C.2 for all AUC values in this chapter).

At the metacommunity level, ${}^{q}\bar{R}^{\mathbf{I}}$ (the weighted average of the ${}^{q}\bar{\rho}_j^{\mathbf{I}}$s) is almost perfectly conserved for $q < 2$, but for $q = \infty$, ${}^{\infty}\bar{R}^{\mathbf{I}}$ is increasingly overestimated with decreased sampling effort (Figure 5.5). At $q = \infty$, ${}^{\infty}\bar{R}^{\mathbf{I}}$ describes the representativeness of the least representative species in the least representative subcommunity, which is increasingly overestimated at low sampling effort. This is because: (1) Subcommunity representativeness is low when the relative abundance of species in a particular subcommunity is high relative to the abundance of those same species across the metacommunity as a whole (which can happen when subcommunities contain relatively few species), and the variation in this value depends on which subcommunities have been sampled; and (2) representativeness increases with the relative size of the subcommunity (which is relatively larger given that the size of the metacommunity is smaller when fewer subcommunities are sampled).

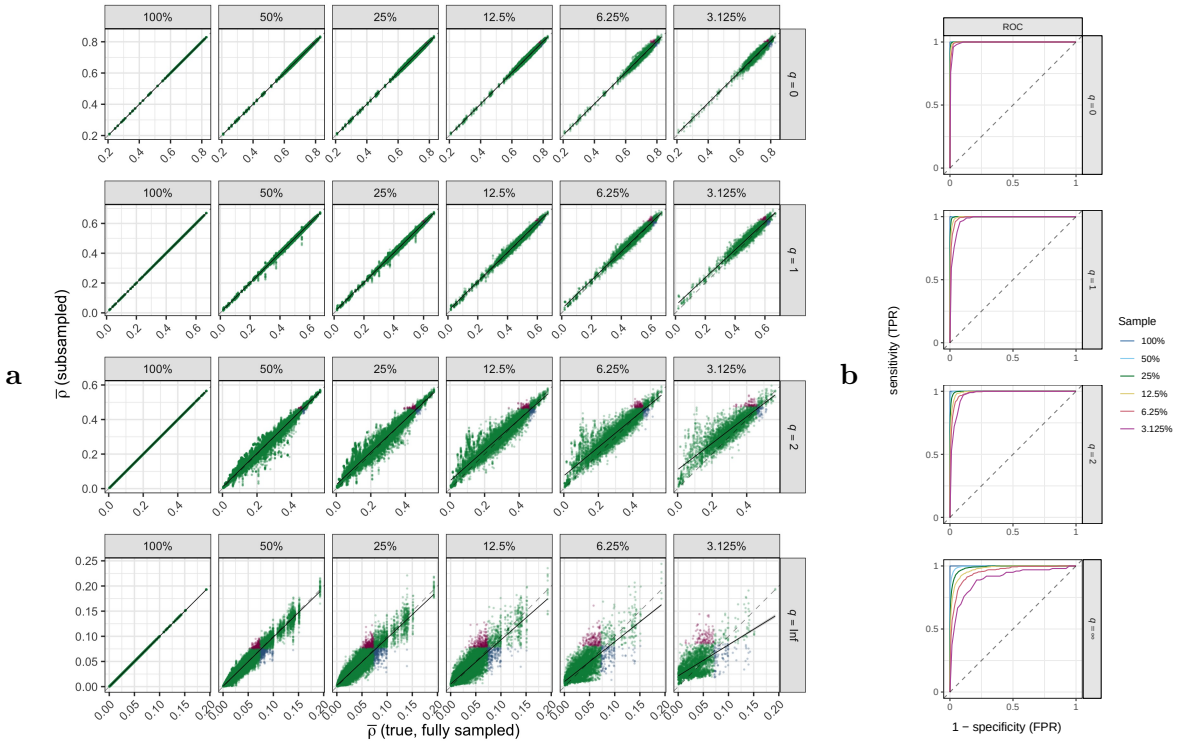


Figure 5.4: The effect of subsampling *subcommunities* on values of *subcommunity representativeness* ($^q\bar{\rho}_j^{\mathbf{I}}$) at different values of q and sampling effort: (a) The true representativeness of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity representativeness is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

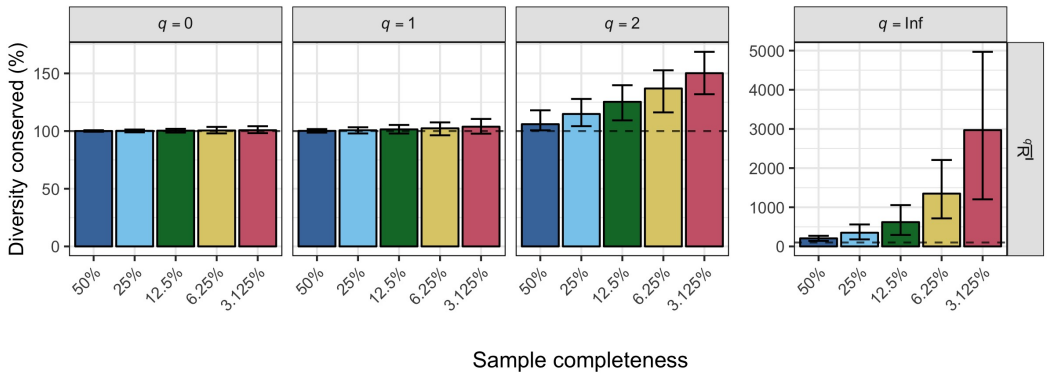


Figure 5.5: The effect of subsampling *subcommunities* on values of *metacommunity representativeness* ($^q\bar{\rho}_j^{\mathbf{I}}$) at different values of q and sampling effort: The bar plot shows the relative value of $^q\bar{R}^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line, whereas the lower dashed lines scale by fraction sampled.

5.4.1.2 Sampling by individual

Unlike sampling by subcommunity, when subsampling by individual, subcommunity representativeness, ${}^{q}\bar{\rho}_j^{\mathbf{I}}$, is very poorly conserved for almost any value of q or level of sampling effort (Figure 5.6a). As well as this, estimates are extremely variable, where R^2 values are lower at 50% sampling effort under this sampling strategy, compared to 3.125% when sampling by subcommunity, for all values of q (Table C.1).

As anticipated in Section 5.4.1, at $q = 0$, where representativeness directly measures the proportion of individuals in the metacommunity with a species representative (a conspecific) in a given subcommunity, representativeness is clearly underestimated, particularly for lower levels of sampling effort. However, the extent of this underestimation depends on the species composition of the subcommunity, with huge variability even for subcommunities with initially high representativeness (when fully sampled). This underestimation is also observed for all values of q and all levels of sampling effort for the most representative subcommunities.

As q increases, the variance in estimates increases substantially as ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ focuses more and more on the species that have low representativeness, those that are relatively common in each subcommunity compared to the metacommunity (the least redundant species); for instance, as q increases from 0 to ∞ at 25% sampling effort, the R^2 decreases from 0.50 to 0.18. These species might be rare within the whole metacommunity, but they might not be. If they are rare, then as sampling effort is decreased, they are likely to be missed in one or many subcommunities and the representativeness of the subcommunity could change dramatically depending on which samples are missed. Conversely, if these species have a high metacommunity-abundance then the representativeness of the subcommunity should be generally conserved. Linear regression shows that ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ is best conserved at $q = 1$. However, ROC curves show that for all values of q, with sampling effort less than 50%, ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ is unable to accurately identify the top 5% of the most representative subcommunities. At 25% sampling effort, AUC values are 0.827, 0.875, 0.862, and 0.836 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.2), which are worse than subsampling subcommunities at 3.125% sampling effort.

At the metacommunity level, ${}^{q}\bar{R}^{\mathbf{I}}$ is again very poorly conserved compared to subsampling by subcommunity. On average, these values are best conserved at $q = \infty$, though results are variable (${}^{q}\bar{R}^{\mathbf{I}}$, Figure 5.7). When 50% of individuals are sampled from each subcommunity, more than 75% of metacommunity ${}^{q}\bar{R}^{\mathbf{I}}$ is conserved for all values of q, however for $q \leq 2$, ${}^{q}\bar{R}^{\mathbf{I}}$ is increasingly underestimated as sampling effort is decreased. This is not a systematic scaling by fraction sampled, but looks regular enough that a method may be developed in the future to compensate for this effect.

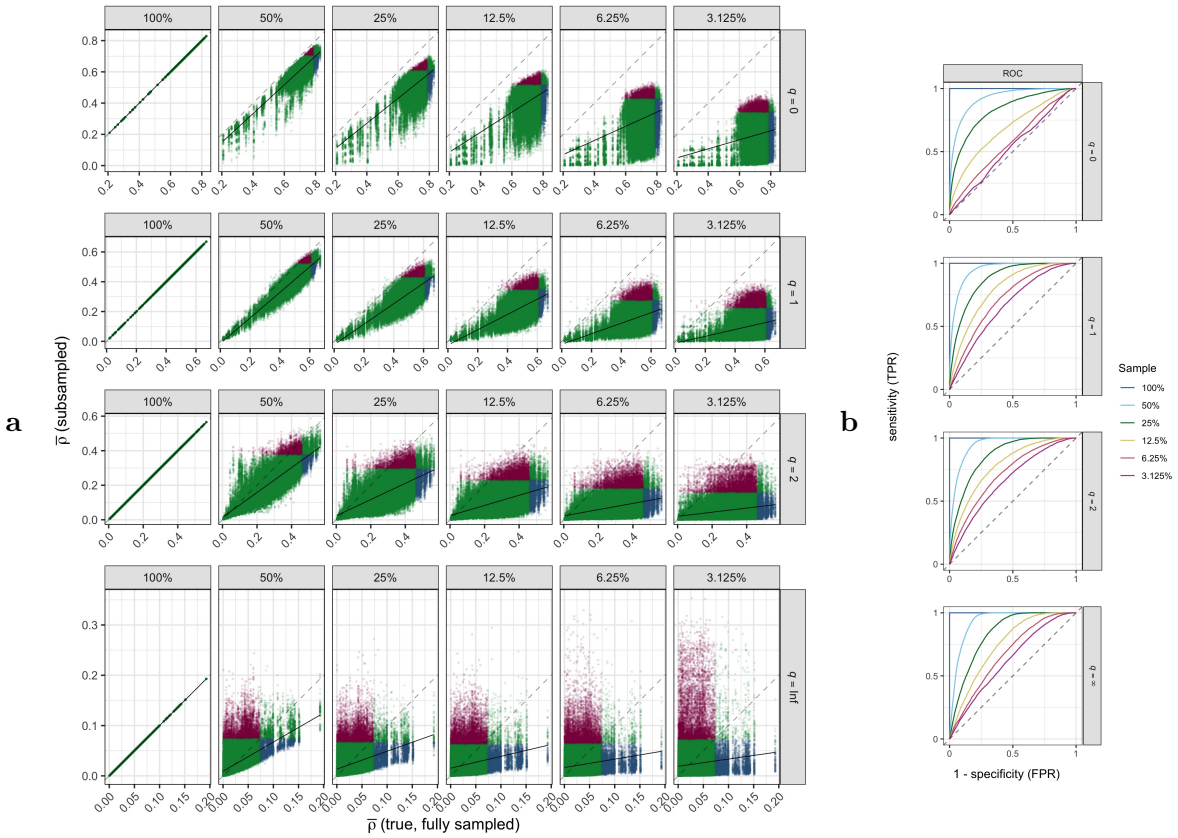


Figure 5.6: The effect of subsampling *individuals* on values of *subcommunity representativeness* ($^q\bar{\rho}_j^{\mathbf{I}}$) at different values of q and sampling effort: (a) The true representativeness of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity representativeness is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

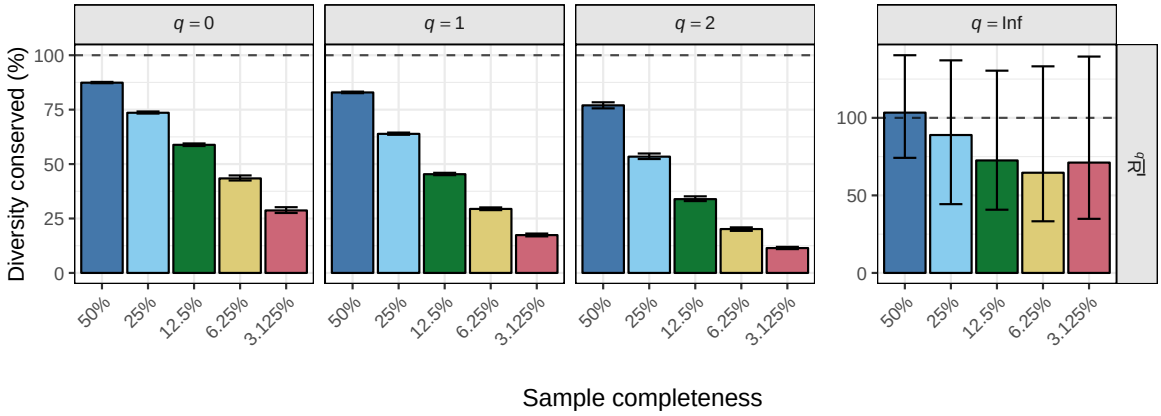


Figure 5.7: The effect of subsampling *individuals* on values of *metacommunity representativeness* ($^{q}\bar{R}^{\mathbf{I}}$) at different values of q and sampling effort: The bar plot shows the relative value of $^{q}\bar{R}^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

5.4.2 Redundancy

Subcommunity $^{q}\rho_j^{\mathbf{I}}$ quantifies the redundancy of species within each subcommunity (described in Section 2.2.6.2). For subsampling by subcommunity, w_j for the sampled subcommunities should be scaled by sampling effort, where f is the fraction sampled, and so redundancy should be scaled by the same fraction. As a result, values are expected to be well conserved, though scaled when sampling subcommunities, especially for lower values of q. Since $^{q}\rho_j^{\mathbf{I}} = {}^{q}\bar{\rho}_j^{\mathbf{I}}/w_j$, where w_j denotes the weight of subcommunity j (which is constant when subsampling by individual, since the same proportion of individuals are sampled from each subcommunity), then results for subsampling by individual should be comparable to those observed for subcommunity representativeness, that is, poorly conserved.

5.4.2.1 Sampling by subcommunity

Figure 5.8a shows how subcommunity redundancy, $^{q}\rho_j^{\mathbf{I}}$, is affected by reduced sampling effort when sampling by subcommunity. It is clear that subcommunity redundancy is indeed well conserved, following the red dashed line (true diversity scaled by $\frac{1}{f}$), even when sampling effort is extremely low. For $q \leq 1$, the redundancy of fully sampled subcommunities is accurately predicted for all levels of sampling effort tested. As q increases, $^{q}\rho_j^{\mathbf{I}}$ increasingly focuses on the least redundant (most distinct) subcommunities (those that contain the highest abundances of rare species). Here, values are less well conserved, but R^2 still only drops to 0.80 at $q = 2$, and 0.47 at $q = \infty$, for the lowest sampling effort (Table C.1). Furthermore, the ROC curves demonstrate that this does not affect the ability of the measure to identify highly redundant subcommunities, which are accurately identified for all values of q and all levels of sampling

effort (Figure 5.8b). For the lowest level of sampling effort, AUC are calculated as 0.997, 0.995, 0.985, and 0.897 for $q \in \{0, 1, 2, \infty\}$ respectively (Table C.2).

At the metacommunity level, subsampled ${}^{q}R^{\mathbf{I}}$, like ${}^{q}\rho_j^{\mathbf{I}}$, is scaled by the fraction sample for $q \leq 2$ (coloured dashed lines, Figure 5.9). This isn't surprising since redundancy is measuring the number of other subcommunities that contain the same species distribution, which naturally decreases as fewer subcommunities are sampled. Interestingly however, at $q = \infty$, redundancy is perfectly conserved for all values of sampling effort, without the scaling that is found for lower values of q. This is because there are many species that are only found in one subcommunity and even under heavy sampling pressure, at least one of these species remains, and therefore ${}^{\infty}R^{\mathbf{I}} = 1$, since at $q = \infty$ only the lowest value of redundancy in the whole metacommunity is considered. This also explains why subsampled ${}^{\infty}\bar{R}^{\mathbf{I}}$ is such a poor estimate of the true value (Figure 5.5). It is in fact, just a redundancy of 1 scaled by $\frac{1}{f}$, where f is the sample fraction.

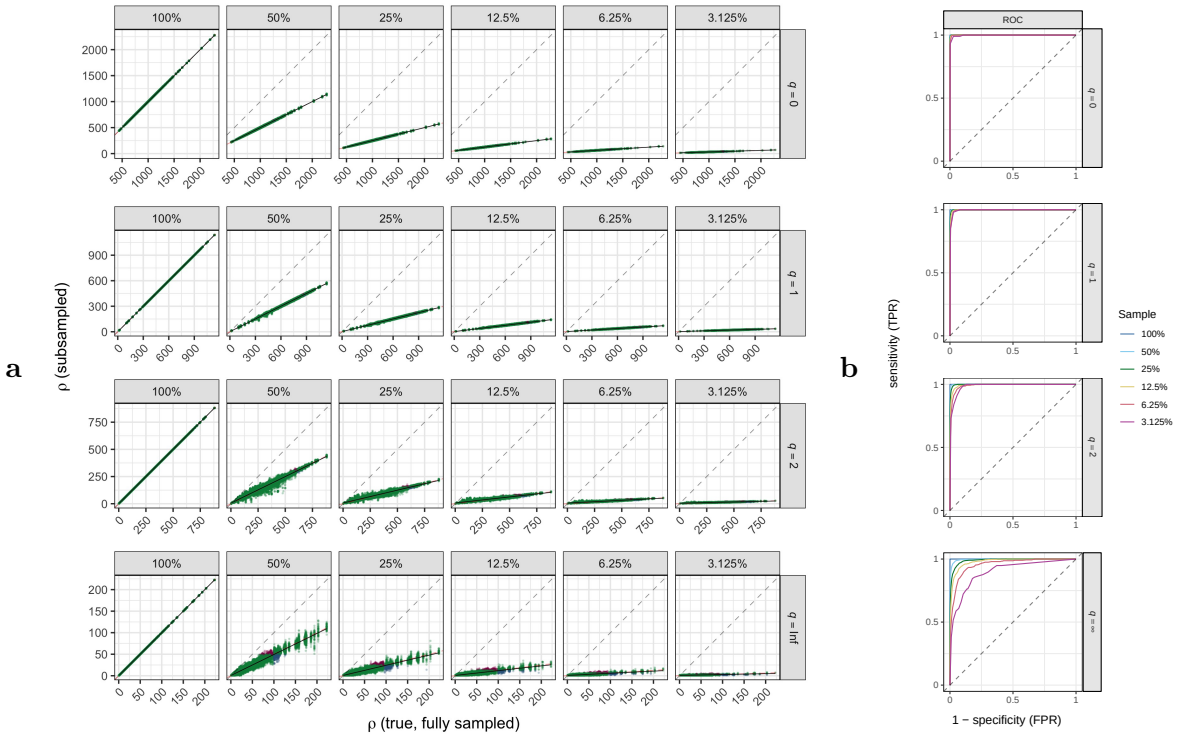


Figure 5.8: The effect of subsampling *subcommunities* on values of *subcommunity redundancy* (${}^{q}\rho_j^{\mathbf{I}}$) at different values of q and sampling effort: (a) The true redundancy of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, subcommunity redundancy is fully conserved when points follow the dashed grey line, whereas the dashed red line scales by the fraction sampled. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

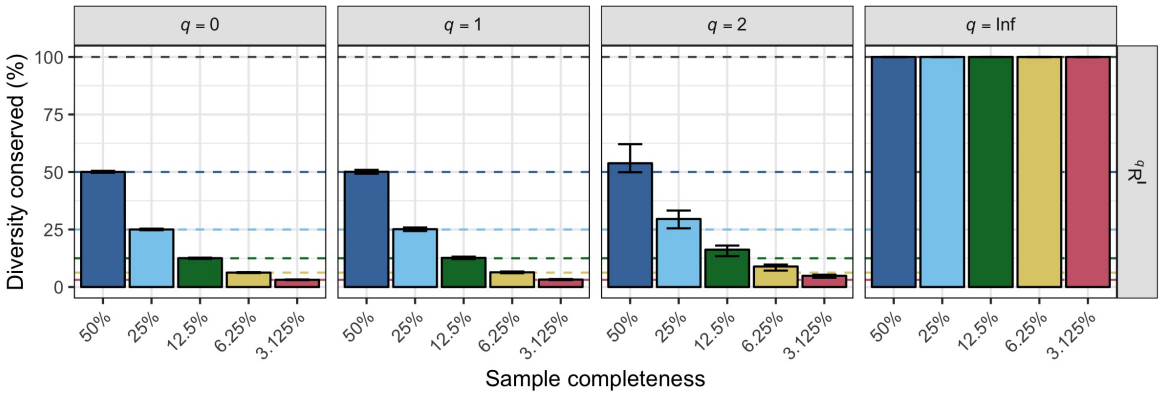


Figure 5.9: The effect of subsampling *subcommunities* on values of *metacommunity redundancy* ($^qR^{\mathbf{I}}$) at different values of q and sampling effort: The bar plot shows the relative value of $^qR^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line, whereas coloured dashed lines at 50%, 25%, 12.5%, 6.25%, and 3.125% are scaled by fraction sampled.

5.4.2.2 Sampling by individual

As with representativeness, redundancy is poorly conserved for all levels of q and sampling effort ($^q\rho_j^{\mathbf{I}}$, Figure 5.10a). The R^2 values are less than 65% for $q \leq 1$ and less than 45% for $q = 2$, even at 25% sampling effort (Table C.1). The ROC curves also show that for all values of q and sampling effort of less than 50%, these measures are unable to accurately identify the most redundant subcommunities (Figure 5.10b), demonstrating that redundancy is strongly affected by sample completeness when sampling by individuals, since the composition of the subsample no longer reflects that of the subcommunity. AUC values are 0.839, 0.881, 0.866, and 0.844, at 25% sampling effort, for $q = \in \{0, 1, 2, \infty\}$, respectively (Table C.2), worse than subsampling by subcommunity, even at 3.125% sampling effort.

At the metacommunity level, $^qR^{\mathbf{I}}$ calculates the average redundancy across subcommunities. At $q = \infty$, these values are remarkably well conserved (as when sampling by subcommunity), again because at least one species is found to be unique to a subcommunity across the metacommunity ($^qR^{\mathbf{I}}$, Figure 5.11). On the other hand, for lower values of q, $^qR^{\mathbf{I}}$, like representativeness, is increasingly underestimated as sampling effort is decreased. Unlike sampling by subcommunity, this is not a systematic scaling by fraction sampled, but again looks regular.

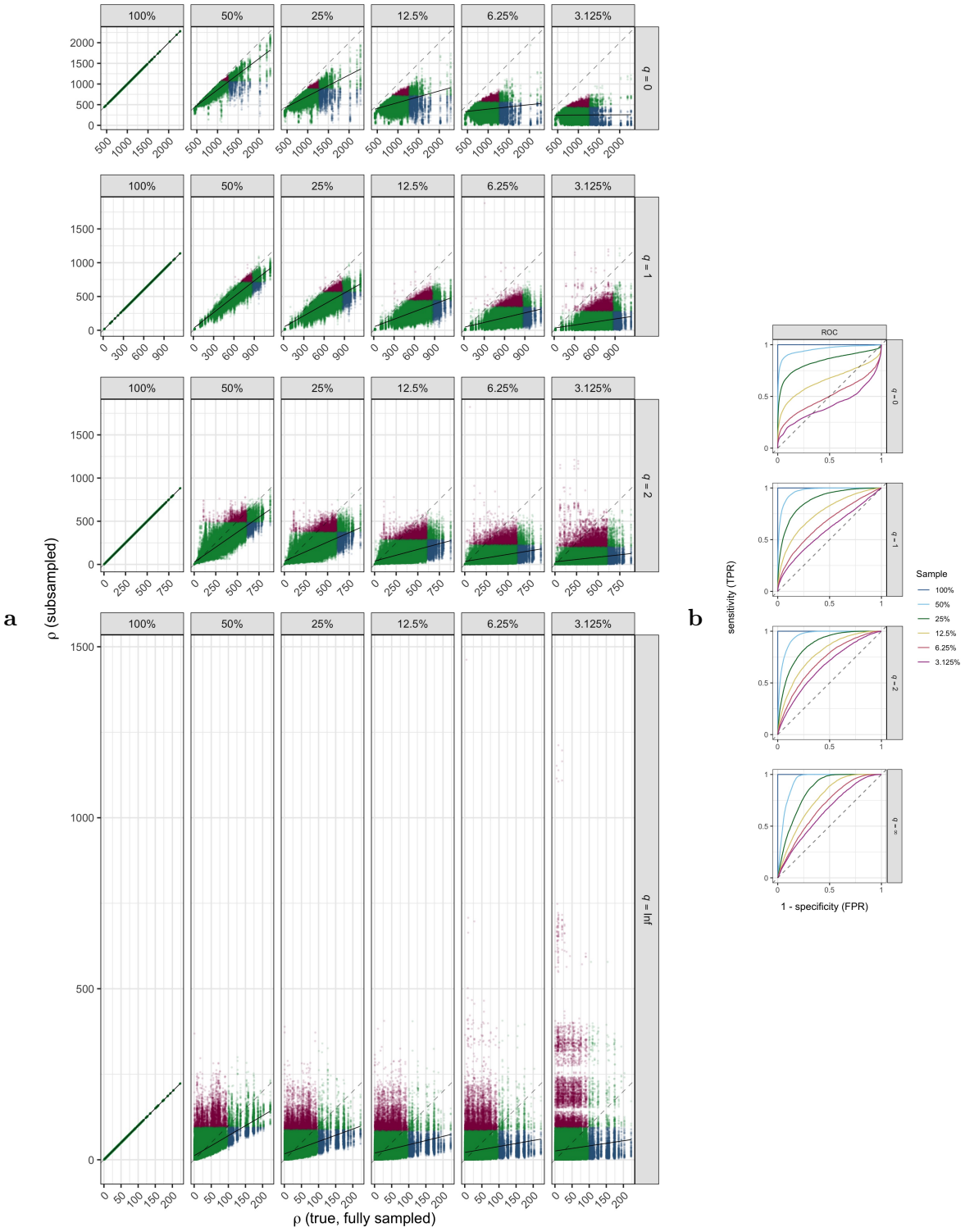


Figure 5.10: The effect of subsampling *individuals* on values of *subcommunity redundancy* ($^q\rho_j^{\mathbf{I}}$) at different values of q and sampling effort: (a) The true redundancy of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity redundancy is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

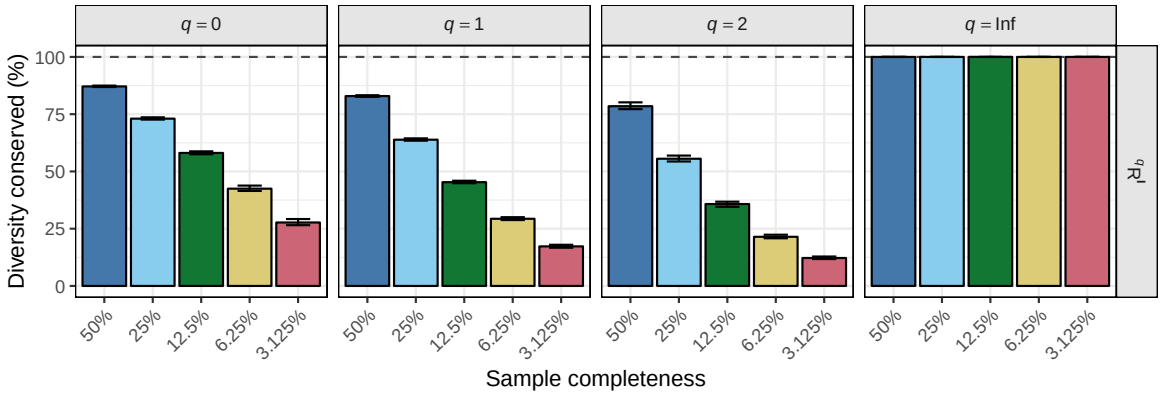


Figure 5.11: The effect of subsampling *individuals* on values of *metacommunity redundancy* ($^{q}R^{\mathbf{I}}$) at different values of q and sampling effort: The bar plot shows the relative value of $^{q}R^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

5.4.3 Effective number of distinct subcommunities

Subcommunity $^{q}\bar{\beta}_j^{\mathbf{I}}$ is an estimate of the effective number of distinct subcommunities in a metacommunity (described in Section 2.2.6.3). In the naïve-type case, this measure compares the abundance of species in subcommunity j to the abundance of those same species across the metacommunity. Therefore, when sampling by subcommunity, like representativeness, the subcommunity diversity measures should be accurate so long as the overall distribution of species in the metacommunity is well conserved.

When species are completely distinct from the rest of the metacommunity, $^{q}\bar{\beta}_j^{\mathbf{I}}$ takes a maximum value of $1/w_j$, and decreases to 1 when the distribution of species in subcommunity j matches that of the metacommunity – or equivalently, when representativeness equals one (its maximum value). When sampling by individual, some rare species are likely to be missed in individual subcommunities, and the estimate of the effective number of subcommunities should therefore be overestimated.

5.4.3.1 Sampling by subcommunities

Like representativeness, sampled estimates of $^{q}\bar{\beta}_j^{\mathbf{I}}$ are very accurate for $q = 0$, for all levels of sampling effort (Figure 5.12a), with $R^2 = 0.98$ at 3.125% sampling effort (Table C.1). However, at higher values of q, R^2 drops much faster than it does for representativeness, even though $^{q}\bar{\beta}_j^{\mathbf{I}} = 1/{}^{q}\bar{\rho}_j^{\mathbf{I}}$. Furthermore, these estimates are systematically too low, with the bias increasing with decreasing sampling effort and higher values of q. This is because the representativeness of subcommunities with very small fully sampled values of representativeness (*e.g.* low $^{q}\bar{\rho}_j^{\mathbf{I}}$ at 100% sampling effort Figure 5.4a) are slightly overestimated when subsampled

and have large errors relative to their size, although these errors are small in absolute terms for representativeness. These subcommunities have large ${}^{q}\bar{\beta}_j^{\mathbf{I}}$, which when subsampled are consequently greatly underestimated with large uncertainty. This is particularly apparent at $q = \infty$, where species that are unique to individual subcommunities (where ${}^{\infty}\rho_j^{\mathbf{I}} = 1$, which caused the systematic overestimation of ${}^{\infty}\bar{R}^{\mathbf{I}}$) can be seen to form a line with slope f (the sample fraction). This effect could not be seen in the subcommunity plots for representativeness as the values were too close to zero. As q increases and ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ increasingly focuses more on the least representative species, R^2 values drop to 0.768, 0.627, and 0.144 for $q \in \{1, 2, \infty\}$, respectively, at the lowest sampling effort (Table C.1). Despite this, the ROC curves show that for $q \leq 2$, order is sufficiently well conserved that the top 5% of subcommunities with the highest estimates of the effective number of subcommunities are accurately identified (Figure 5.12b). Even at the lowest level of sampling effort, AUC values are 0.995, 0.997, 0.959, and 0.785 for $q \in \{0, 1, 2, \infty\}$ respectively (Table C.2).

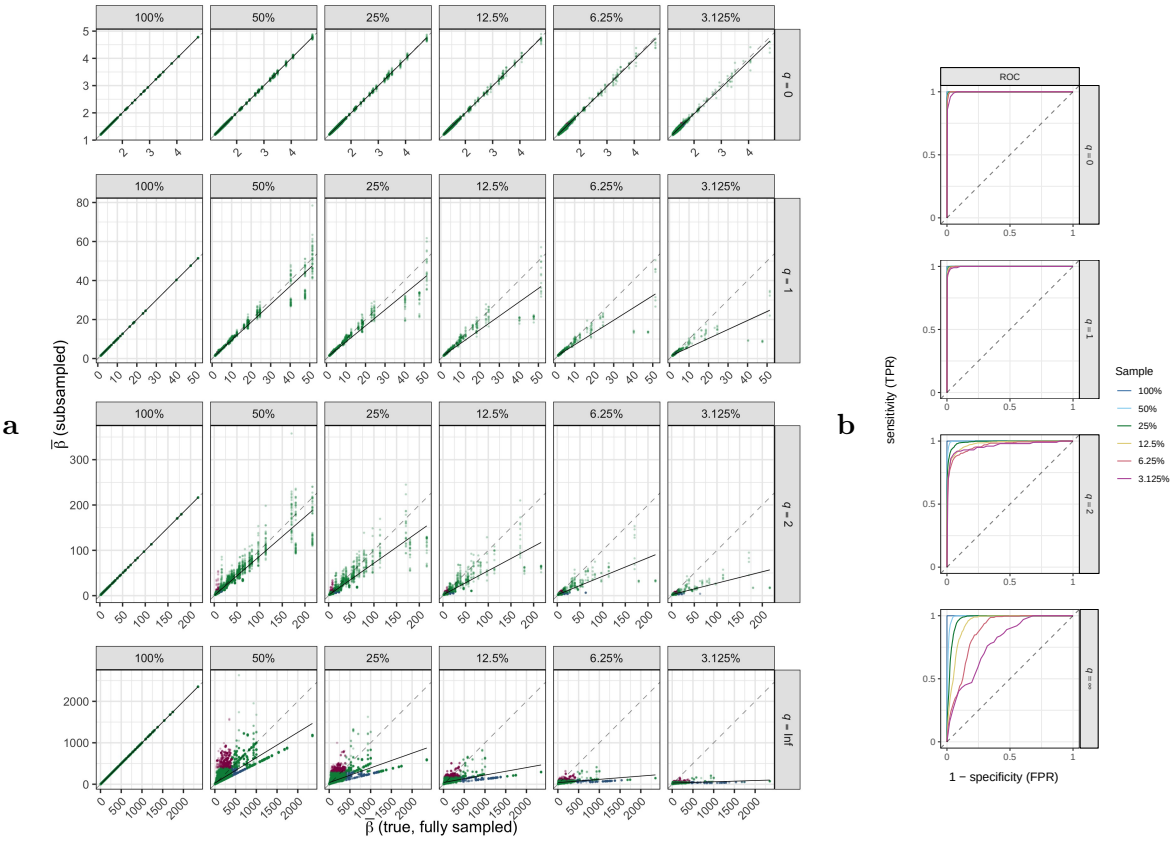


Figure 5.12: The effect of subsampling *subcommunities* on values of *normalised subcommunity beta* (${}^{q}\bar{\beta}_j^{\mathbf{I}}$) diversity at different values of q and sampling effort: (a) The true subcommunity ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ diversity of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ diversity is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

At the metacommunity level, as sampling effort is decreased, ${}^{q}\bar{B}^{\mathbf{I}}$ is well conserved for $q \leq 2$, though it is over-estimated at $q = \infty$ (Figure 5.13). This is in contrast to ${}^{q}\bar{R}^{\mathbf{I}}$, which is overestimated at $q = 2$ and more severely over-estimated at $q = \infty$, even though ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ has higher R^2 values at these values of q. The seeming contradiction that both ${}^{\infty}\bar{R}^{\mathbf{I}}$ and ${}^{\infty}\bar{B}^{\mathbf{I}}$ are both overestimated, despite ${}^{q}\bar{\beta}_j^{\mathbf{I}} = {}^{1}/_{{}^{q}\bar{\rho}_j^{\mathbf{I}}}$, is explained by the observation that ${}^{\infty}\bar{B}^{\mathbf{I}}$ identifies the lowest value of ${}^{\infty}\bar{\beta}_j^{\mathbf{I}}$, whereas ${}^{\infty}\bar{R}^{\mathbf{I}}$ identifies the lowest value of ${}^{\infty}\bar{\rho}_j^{\mathbf{I}}$. These are different and so can both be overestimates.

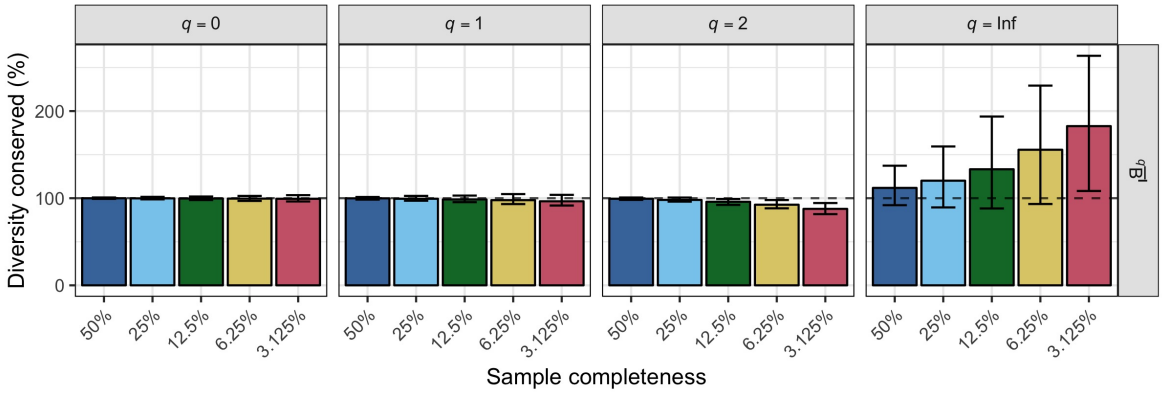


Figure 5.13: The effect of subsampling *subcommunities* on values of *normalised metacommunity beta* (${}^{q}\bar{B}^{\mathbf{I}}$) diversity at different values of q and sampling effort: The bar plot shows the relative value of ${}^{q}\bar{B}^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

5.4.3.2 Sampling by individual

As with representativeness, ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ is poorly conserved for all values of q and sampling effort (Figure 5.14a). The R^2 values are 0.357, 0.730, 0.582, and 0.214 for $q \in \{0, 1, 2, \infty\}$, respectively, even at 25% sampling effort (Table C.1), worse than subsampling by subcommunity at 3.125% sampling effort. Furthermore, ROC curves show that for all values of q and levels of sampling effort, subsampling by individual (Figure 5.14b) is worse than subsampling by subcommunity (Figure 5.12b) at identifying subcommunities with the highest estimates of the effective number of subcommunities. Results show that $q \in \{1, 2\}$ are best able to identify the most interesting subcommunities (in this case the most distinct subcommunities). At 25% sampling effort, AUC values are 0.898, 0.987, 0.945, and 0.721 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.2).

At the metacommunity level, for $q \leq 2$, ${}^{q}\bar{B}^{\mathbf{I}}$ is increasingly overestimated as sampling effort is decreased (Figure 5.15). Like ${}^{q}R^{\mathbf{I}}$ and ${}^{q}\bar{R}^{\mathbf{I}}$, this overestimation is somewhat regular but does not scale with sampling effort in an obvious fashion. On average, ${}^{q}\bar{B}^{\mathbf{I}}$ is best conserved at $q = \infty$, though results are variable. Subcommunity ${}^{\infty}\bar{\beta}_j^{\mathbf{I}}$ considers the most distinct species in a subcommunity and metacommunity ${}^{\infty}\bar{B}^{\mathbf{I}}$ considers only the least distinct of these subcommunities by that measure. As a result, ${}^{\infty}\bar{B}^{\mathbf{I}}$ selects a subcommunity with very few rare species, which are less likely to be missed under subsampling.

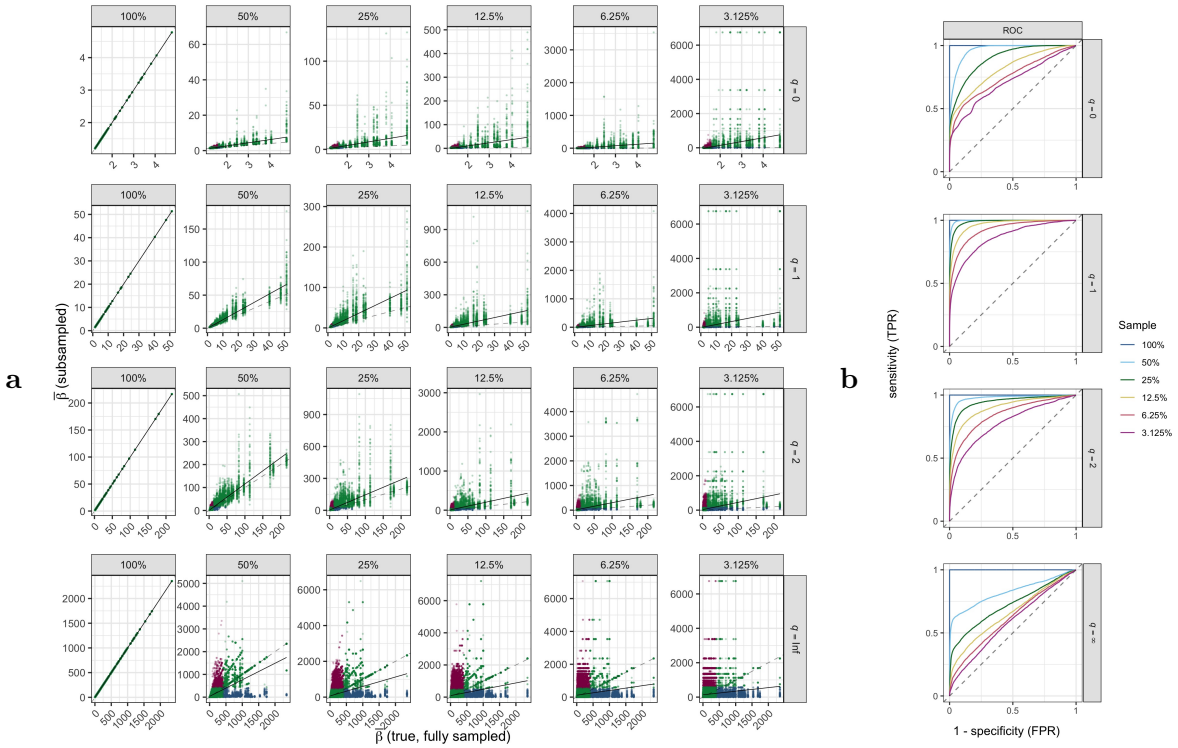


Figure 5.14: The effect of subsampling *individuals* on values of *normalised subcommunity beta* ($^q\bar{\beta}_j^{\mathbf{I}}$) diversity at different values of q and sampling effort: (a) The true subcommunity $^q\bar{\beta}_j^{\mathbf{I}}$ diversity of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity $^q\bar{\beta}_j^{\mathbf{I}}$ diversity is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

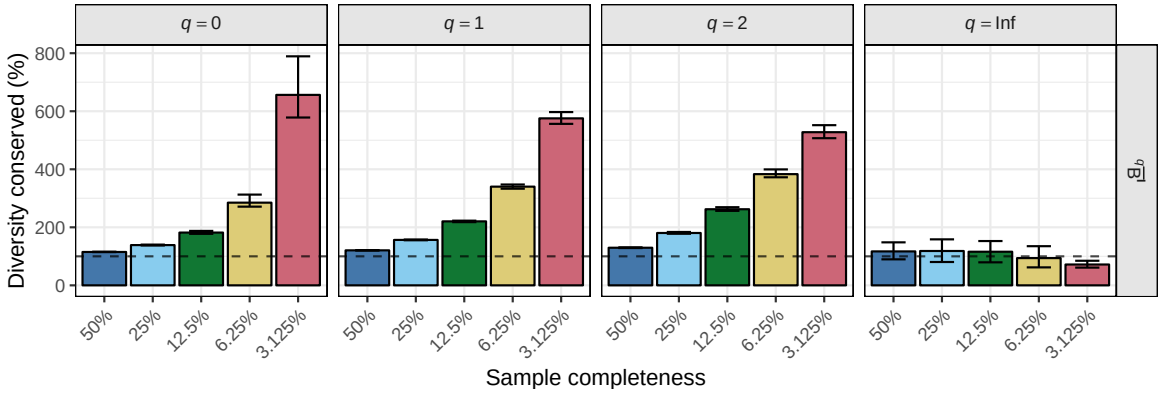


Figure 5.15: The effect of subsampling *individuals* on values of *normalised metacommunity beta* ($^q\bar{B}^{\mathbf{I}}$) diversity at different values of q and sampling effort: The bar plot shows the relative value of $^q\bar{B}^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

5.4.4 Distinctiveness

Subcommunity ${}^{q}\beta_j^{\mathbf{I}}$ describes the average distinctiveness of a subcommunity, relative to the metacommunity as a whole (described in Section 2.2.6.4). Since ${}^{q}\beta_j^{\mathbf{I}} = {}^{1}/_{{}^{q}\rho_j^{\mathbf{I}}}$, results should scale by f (the fraction sampled) when sampling by subcommunity. When subsampling by individual, species that are rare within the subcommunity are likely to be missed and the distinctiveness of sampled subcommunities should therefore be overestimated.

5.4.4.1 Sampling by subcommunity

Like redundancy, at $q \leq 1$, distinctiveness is well conserved when scaled by the fraction sampled for all levels of sampling effort (red dashed lines, Figure 5.16a). However, as with ${}^{q}\bar{\beta}_j^{\mathbf{I}}$, R^2 values drop much faster than for ${}^{q}\rho_j^{\mathbf{I}}$ (Table C.1). At $q \geq 1$, linear regression shows that the scaled estimates are systematically underestimated, with bias increasing with lower sampling effort and higher values of q. As with ${}^{q}\bar{\beta}_j^{\mathbf{I}}$, this is due to subcommunities with very small fully sampled values of representativeness being slightly overestimated, resulting in subcommunities with very large fully sampled values of ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ being greatly underestimated, and subcommunities with very high distinctiveness being massively underestimated (since ${}^{q}\beta_j^{\mathbf{I}} = {}^{{}^{q}\bar{\beta}_j^{\mathbf{I}}}/_{w_j}$).

In Figure 5.16a, ignoring these outliers, distinctiveness is well conserved, following the red dashed line (true diversity scaled by f), for $q \leq 2$ and all levels of sampling effort. However, due to the presence of these outliers, at the lowest level of sampling effort, R^2 values are 0.958, 0.671, 0.524, and 0.09 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.1). Despite this, ROC curves show that the most distinctive subcommunities can be accurately identified when $q \leq 2$ (Figure 5.16b). At 3.125% sampling effort, AUC values are 0.994, 0.994, 0.925, and 0.707 for $q \in \{0, 1, 2, \infty\}$ respectively (Table C.2).

At the metacommunity level, as with at the subcommunity level, results are scaled by sample fraction. For $q \leq 1$, scaled ${}^{q}B^{\mathbf{I}}$ is remarkably well conserved (coloured dashed lines, Figure 5.17). However, at $q = \infty$, values are systematically overestimated. Again, a contradiction is seemingly observed, since ${}^{\infty}R^{\mathbf{I}}$ is perfectly conserved and ${}^{q}\beta_j^{\mathbf{I}} = {}^{1}/_{{}^{q}\rho_j^{\mathbf{I}}}$. However, this is resolved because ${}^{\infty}R^{\mathbf{I}}$ considers the least redundant subcommunity, whereas ${}^{\infty}B^{\mathbf{I}}$ considers the least distinct.

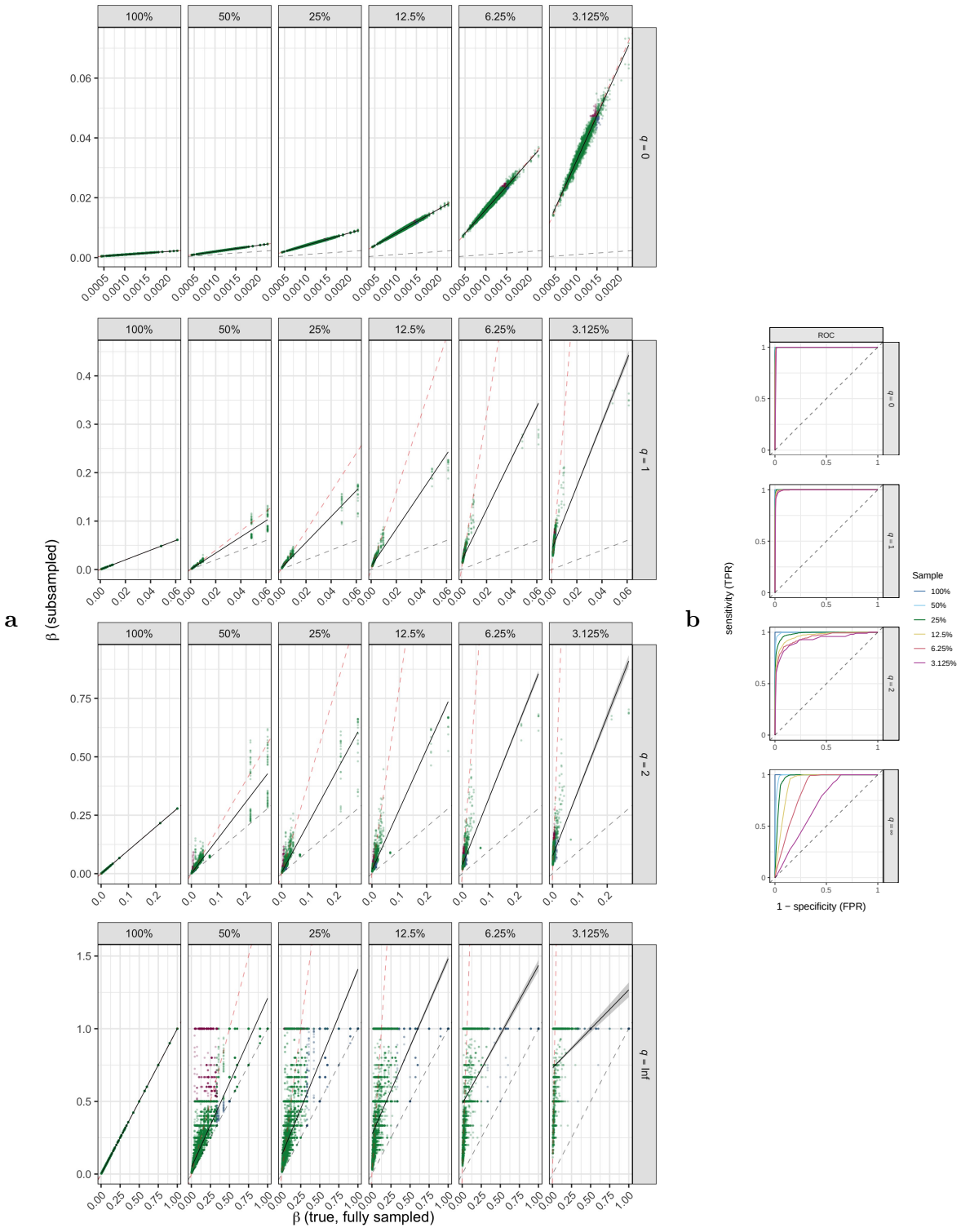


Figure 5.16: The effect of subsampling *subcommunities* on values of *subcommunity distinctiveness* ($^q\beta_j^{\mathbf{I}}$) at different values of q and sampling effort: (a) The true distinctiveness of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, subcommunity distinctiveness is fully conserved when points follow the dashed grey line, whereas the dashed red line scales by the fraction sampled. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

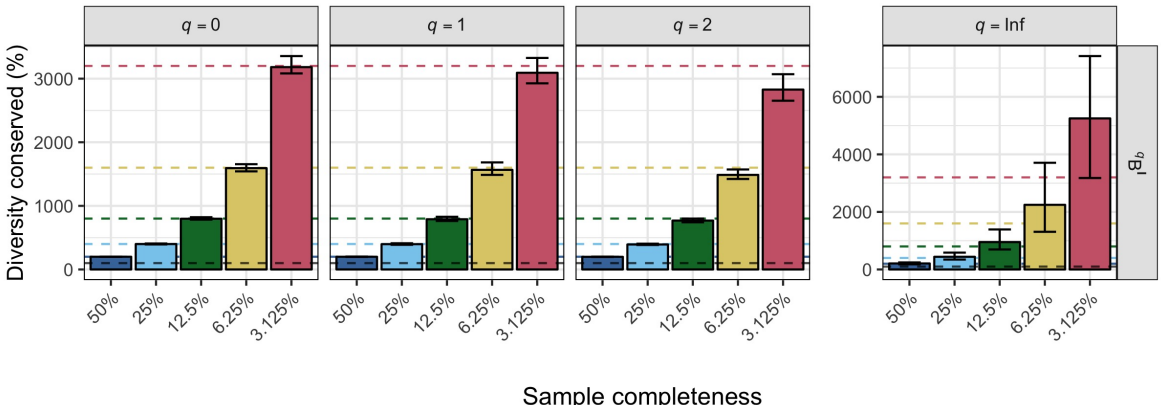


Figure 5.17: The effect of subsampling *subcommunities* on values of *metacommunity distinctiveness* (${}^qB^{\mathbf{I}}$) at different values of q and sampling effort: The bar plot shows the relative value of ${}^qB^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line, whereas coloured dashed lines at 50%, 25%, 12.5%, 6.25%, and 3.125% are scaled by fraction sampled.

5.4.4.2 Sampling by individual

As with redundancy, distinctiveness is poorly conserved for all levels of q and sampling effort (Figure 5.19a), with R^2 values of 0.022 and 0.087 at 12.5% sampling effort for $q \in \{0, \infty\}$, respectively (Table C.1). ROC curves show that $q \leq 2$ are best able to identify the most distinctive subcommunities for sampling effort greater than 25% (Figure 5.19b). Here, AUC values are calculated as 0.975, 0.974, 0.923, and 0.706 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.2). At the metacommunity level, ${}^qB^{\mathbf{I}}$ is again overestimated for $q \in \{0, 1, 2\}$, as sampling effort is decreased. Values are best conserved at $q = \infty$, where on average ${}^qB^{\mathbf{I}}$ is well conserved down to 12.5% sampling effort (Section 5.18), after which values are underestimated.

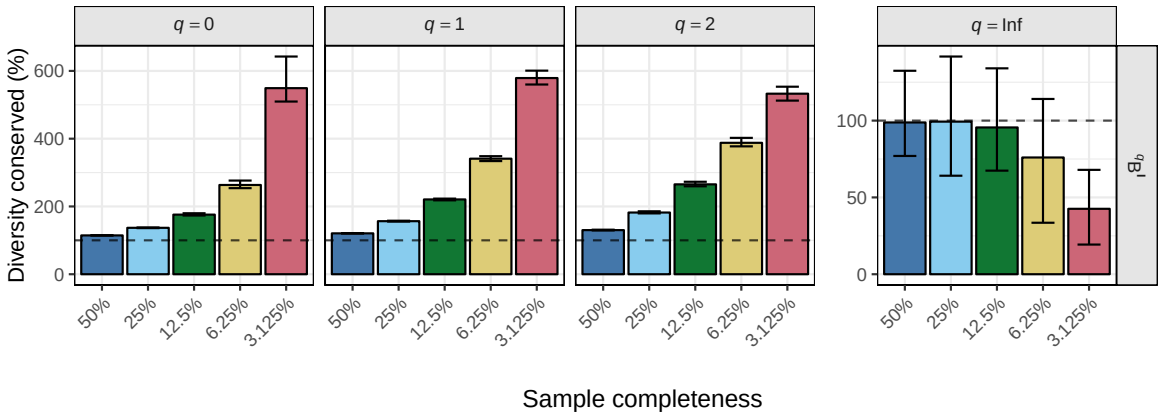


Figure 5.18: The effect of subsampling *individuals* on values of *metacommunity distinctiveness* (${}^qB^{\mathbf{I}}$) at different values of q and sampling effort: The bar plot shows the relative value of ${}^qB^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

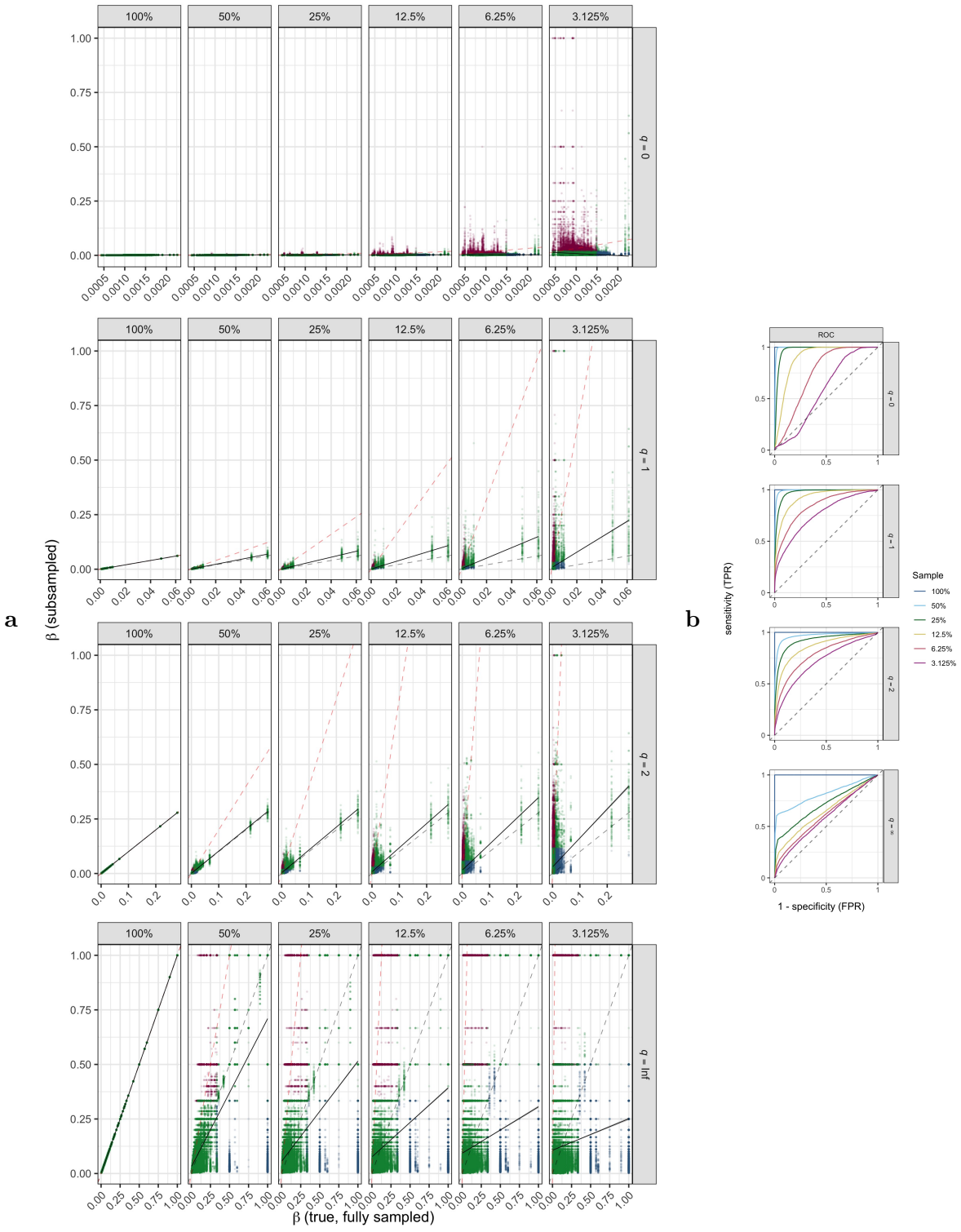


Figure 5.19: The effect of subsampling *individuals* on values of *subcommunity distinctiveness* ($^q\beta_j^{\mathbf{I}}$) at different values of q and sampling effort: (a) The true distinctiveness of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity distinctiveness is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

5.4.5 Gamma diversity

Gamma diversity is traditionally a metacommunity-level concept. At this level, even sampling only 3.125% of the assemblage includes over 6000 trees and as such should give reasonable estimates whether subsampling by subcommunity or individual for $q \geq 1$. For species richness ($q = 0$), rare species will continue to accumulate (and affect the value of measured diversity) in such a diverse study site outwith this 3.125%.

Subcommunity ${}^{q}\gamma_j^{\mathbf{I}}$ describes the average metacommunity contribution of each tree present in a subcommunity (described in Section 2.2.5). As such, values should be well conserved when subsampling by individual (since this measure of diversity is per tree sampled). But as sample size decreases and more species that are rare in the subcommunity are likely to be missed, there will be increasingly large amounts of variability in measured ${}^{q}\gamma_j^{\mathbf{I}}$ – since these species might be rare across the metacommunity and therefore have a high contribution per tree to metacommunity gamma diversity, or they might be common and have a much lower contribution. Therefore, there are two sources of variability, the actual metacommunity contribution and which individuals are present in the subcommunity sample, whereas when subsampling by subcommunity, only the first of these is a problem (because each subcommunity is fully sampled). As a result, there should be lower variability when subsampling by subcommunity than by individual, and like metacommunity ${}^{q}G^{\mathbf{I}}$, there should also be better results for $q \geq 1$ when metacommunity diversity estimates should be accurate (since at higher values of q, only the most abundant species are considered) and so the second source of variability is reduced.

5.4.5.1 Sampling by subcommunity

As expected, ${}^{q}G^{\mathbf{I}}$ is almost perfectly conserved for $q \geq 1$, and at $q = 0$ it is increasingly underestimated as sampling effort is decreased (Figure 5.21). Linear regression shows that ${}^{q}\gamma_j^{\mathbf{I}}$ is well conserved for $q \geq 1$ at all levels of sampling effort (Figure 5.20a). At the lowest level of sampling effort, R^2 values are 0.89 and 0.955 for $q = 1$ and $q = 2$, respectively (Table C.1). At $q = 0$ on the other hand, values are highly variable and tend to underestimate fully sampled ${}^{q}\gamma_j^{\mathbf{I}}$. The latter is unsurprising, since the ${}^{0}G^{\mathbf{I}}$ is itself underestimated.

ROC curves reflect these results, and ${}^{q}\gamma_j^{\mathbf{I}}$ is able to identify which subcommunities contribute most strongly to the diversity of the metacommunity almost perfectly for all levels of sampling effort when $q \geq 1$ (Figure 5.20b). At 3.125% sampling effort, AUC values are 0.839, 0.976, 0.993, and 1 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.2).

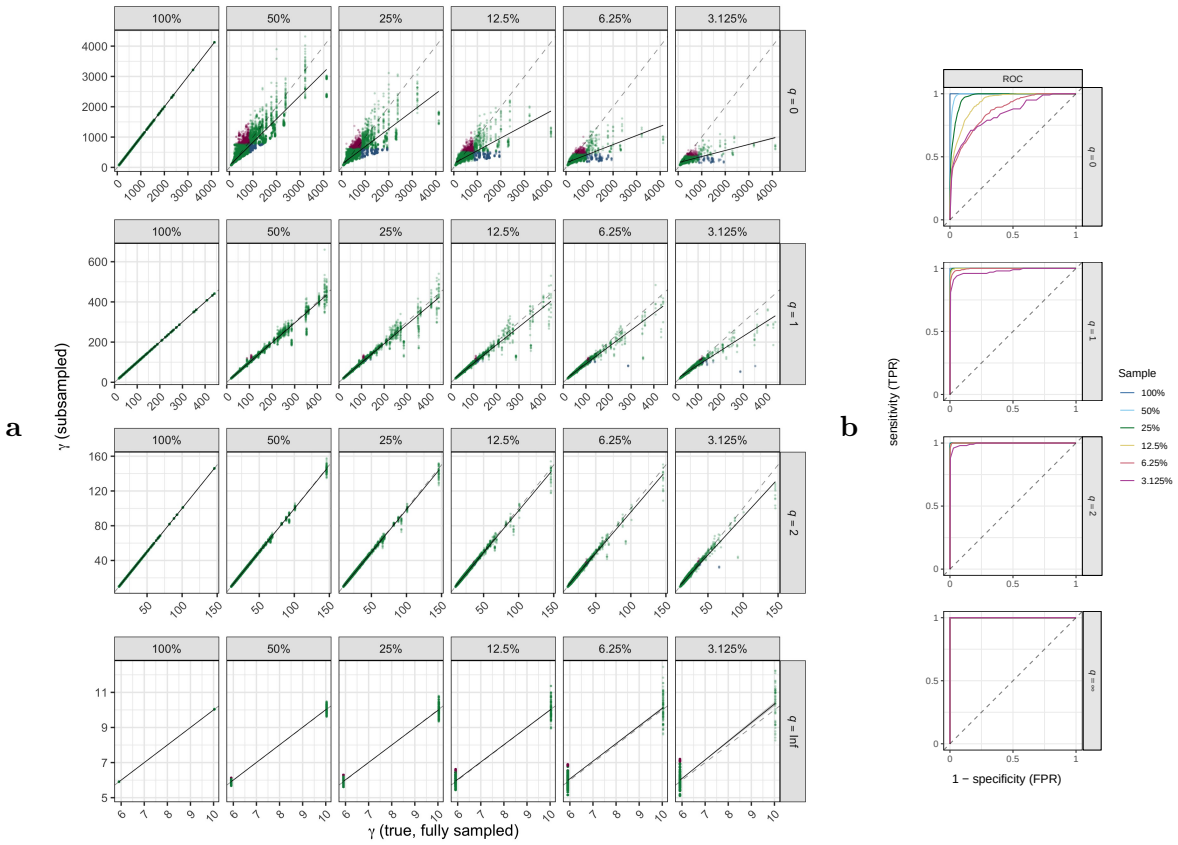


Figure 5.20: The effect of subsampling *subcommunities* on values of *subcommunity gamma* ($^q\gamma_j^{\mathbf{I}}$) diversity at different values of q and sampling effort: (a) The true subcommunity $^q\gamma_j^{\mathbf{I}}$ diversity of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity $^q\gamma_j^{\mathbf{I}}$ diversity is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

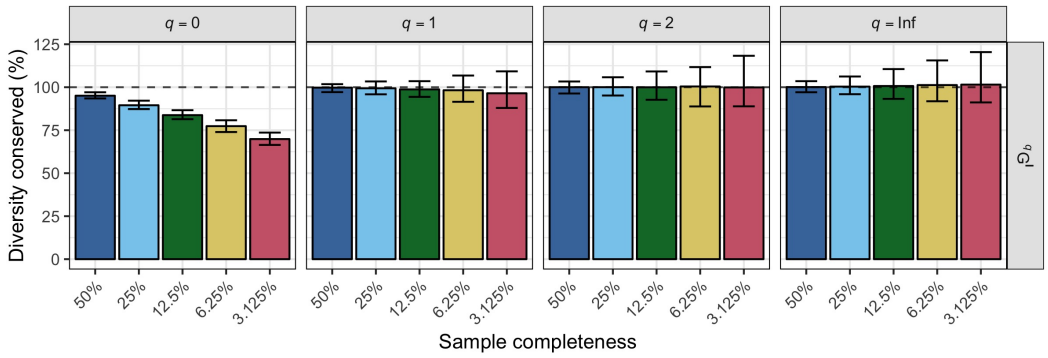


Figure 5.21: The effect of subsampling *subcommunities* on values of *metacommunity gamma* ($^qG^{\mathbf{I}}$) diversity at different values of q and sampling effort: The bar plot shows the relative value of $^qG^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

5.4.5.2 Sampling by individual

As expected, at the metacommunity level, ${}^{q}G^{\mathbf{I}}$ is again almost perfectly conserved for $q \geq 1$ (${}^{q}G^{\mathbf{I}}$, Figure 5.22) and again performs less well at $q = 0$ for low sampling effort due to the diversity of the study site.

Again, as expected, sampling by individual yields higher variability in subcommunity estimates of ${}^{q}\gamma_j^{\mathbf{I}}$ for all values of q and sampling effort than sampling by subcommunity (Figure 5.23a). Even at 25% sampling effort, R^2 values are 0.306, 0.871, 0.803, and 0.200 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.1). These values are less than subsampling by subcommunity for 3.125% sampling effort. ROC curves and AUC values show that ${}^{q}\gamma_j^{\mathbf{I}}$ is able to accurately identify subcommunities that contribute highest to the diversity of the metacommunity when $q \geq 1$ for sampling effort greater than 25% (Figure 5.23b). The AUC values are calculated as 0.685, 0.889, 0.891, and 0.952, for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.2).

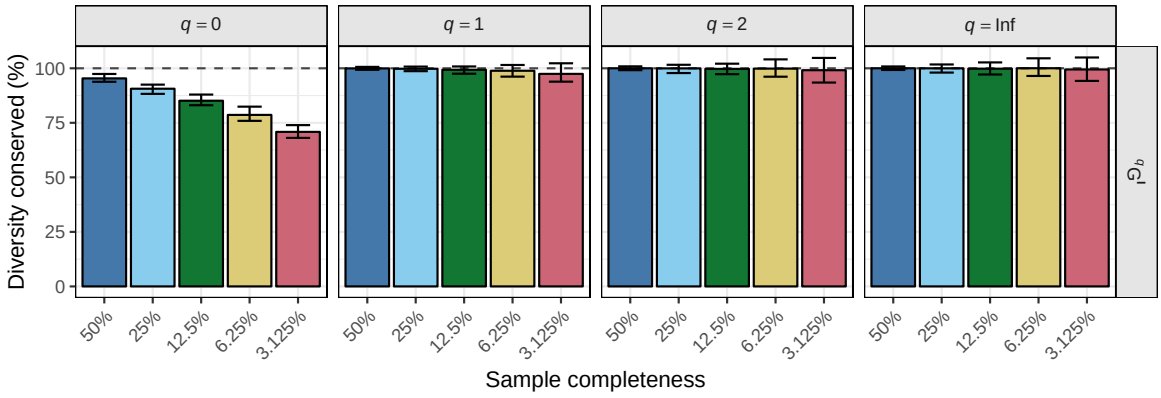


Figure 5.22: The effect of subsampling *individuals* on values of *metacommunity gamma* (${}^{q}G^{\mathbf{I}}$) diversity at different values of q and sampling effort: The bar plot shows the relative value of ${}^{q}G^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

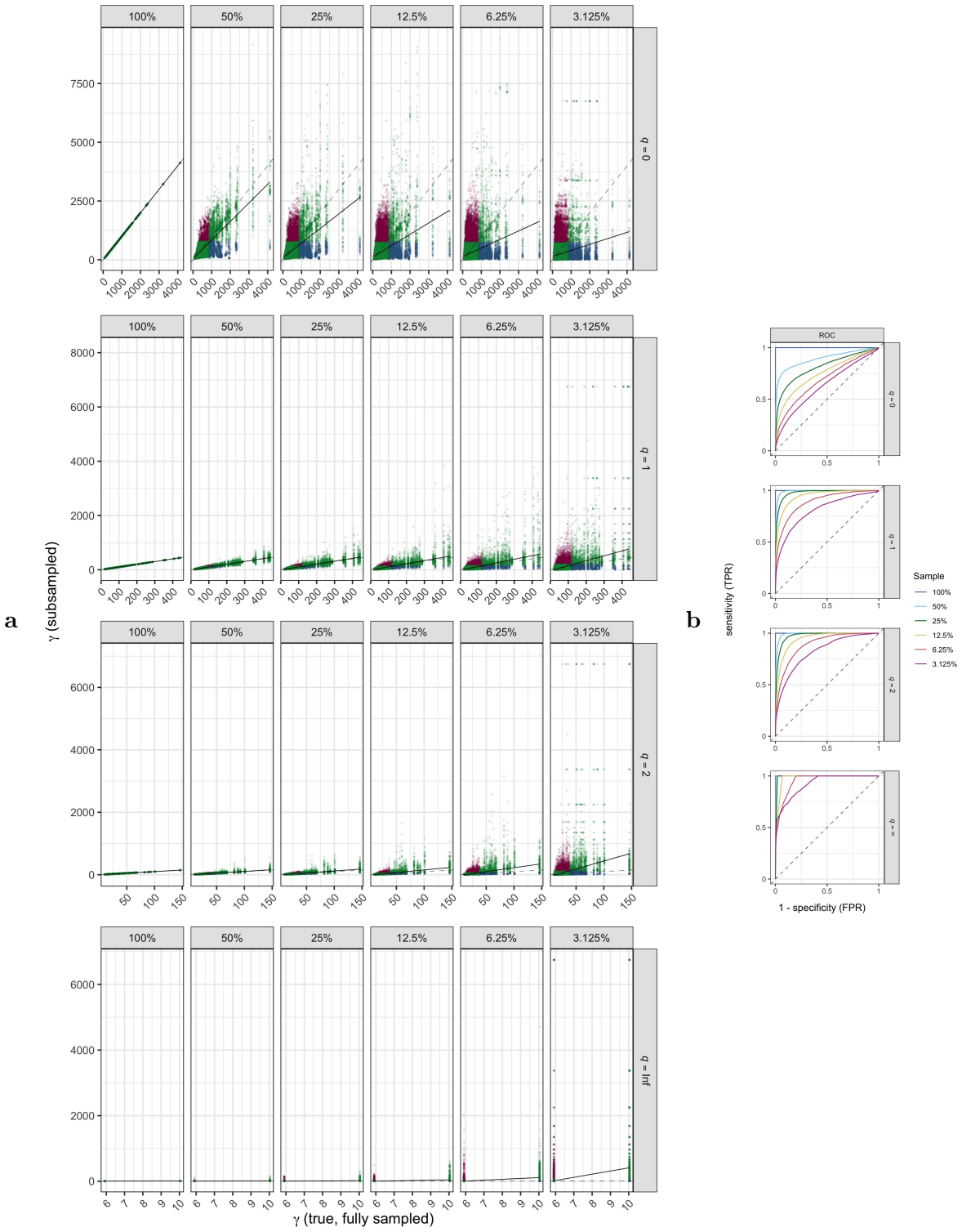


Figure 5.23: The effect of subsampling *individuals* on values of *subcommunity gamma* ($^{q}\gamma_j^{\mathbf{I}}$) diversity at different values of q and sampling effort: (a) The true subcommunity $^{q}\gamma_j^{\mathbf{I}}$ diversity of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity $^{q}\gamma_j^{\mathbf{I}}$ diversity is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m $\times$ 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

5.4.6 Alpha diversities

Subcommunity ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ describes the effective number of species in each subcommunity in isolation (described in Section 2.2.4.1). This measure takes a minimum value of one when only a single species is present and a maximum value of S when species are evenly distributed (within the subcommunity). In the latter case, diversities can only be underestimated as fewer individuals are sampled and species become less evenly distributed. Therefore, when subsampling by individual, subsampled ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ will be severely affected by rarefaction issues. When subsampling by subcommunity, subsampled ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ should be unaffected.

Subcommunity ${}^{q}\alpha_j^{\mathbf{I}}$ is an estimate of naïve-community metacommunity diversity (Raw alpha diversity, Section 2.2.4.2). When subsampling by individual, similar results should be obtained for ${}^{q}\alpha_j^{\mathbf{I}}$ as for ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$, since ${}^{q}\alpha_j^{\mathbf{I}} = {}^{q}\bar{\alpha}_j^{\mathbf{I}}/w_j$, whereas, when subsampling by subcommunity, ${}^{q}\alpha_j^{\mathbf{I}}$ should be scaled by f, the fraction sampled.

5.4.6.1 Sampling by subcommunity

The diversity of a subcommunity in isolation, ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ is perfectly conserved for all values of q regardless of sampling effort (Figure 5.24a). This is unsurprising, since ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ depends only on the proportional abundance of species in subcommunity j, and so, omitting other subcommunities from sampling has no effect on results. This remains true at all values of q, where all R^2 values are 1.000 (Table C.1). Likewise, ROC curves show that ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ is able to accurately identify subcommunities with of ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ for all levels of sampling effort and values of q (Figure 5.24b), where all AUC values are 1.000 (Table C.2).

As sampling effort is decreased, ${}^{q}\alpha_j^{\mathbf{I}}$ is almost perfectly conserved, following the red dashed line (true diversity scaled by $\frac{1}{f}$), for all values of q (Figure 5.26a). The R^2 values are 0.976, 0.993, 0.996, and 0.996 for $q \in \{0, 1, 2, \infty\}$, respectively, for the lowest level of sampling effort. This slight variation is observed because w_j does not change consistently for all subcommunities as fewer subcommunities are sampled. Furthermore, ROC curves show that ${}^{q}\alpha_j^{\mathbf{I}}$ is able to accurately identify the most diverse subcommunities for all levels of sampling effort and values of q (Figure 5.26b), where all AUC values are 1.000 (Table C.2).

At the metacommunity-level, ${}^{q}\bar{A}^{\mathbf{I}}$, is almost perfectly conserved for $q \leq 2$, for all levels of sampling effort (Figure 5.25). At $q = \infty$, ${}^{q}\bar{A}^{\mathbf{I}}$ is increasingly overestimated as sampling effort is decreased. This is because ${}^{\infty}\bar{A}_j^{\mathbf{I}}$ considers only the most dominant species in the metacommunity, which may be missed as subcommunities are sampled. Likewise, when scaled by w_j, ${}^{q}A^{\mathbf{I}}$ behaves in the same way (coloured dashed lines, Figure 5.27).

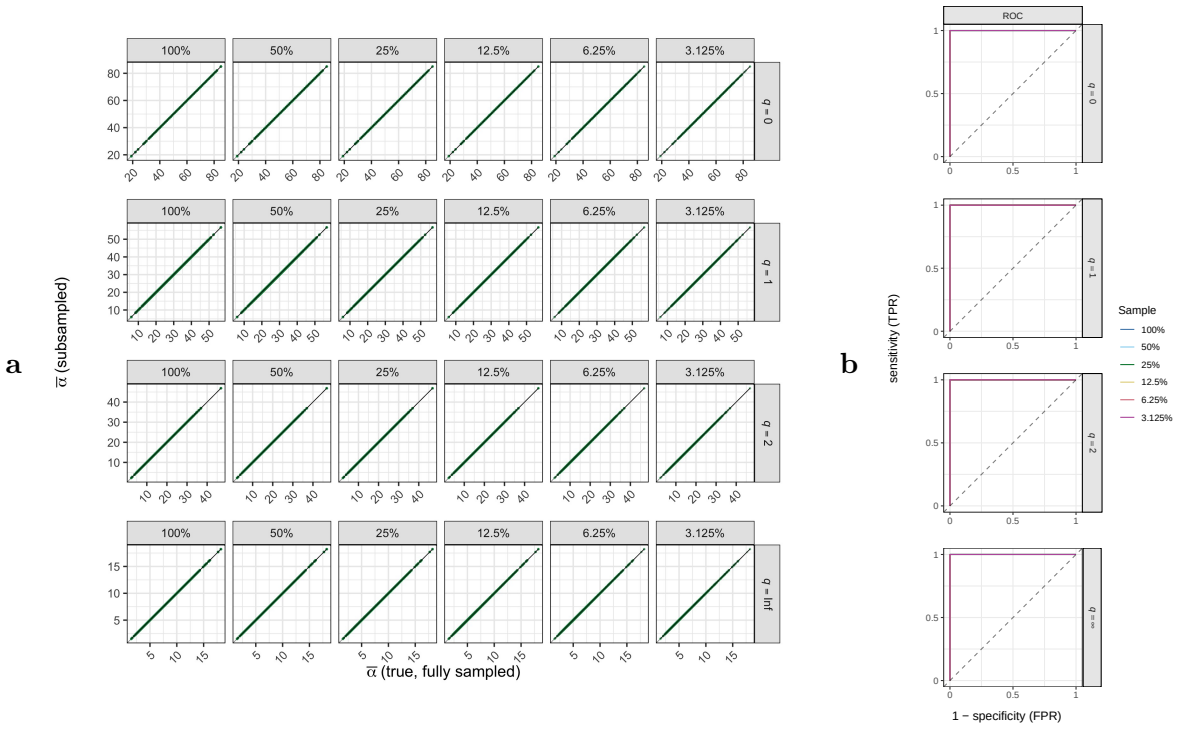


Figure 5.24: The effect of subsampling *subcommunities* on values of *normalised subcommunity alpha* ($^q\bar{\alpha}_j^{\mathbf{I}}$) diversity at different values of q and sampling effort: (a) The true subcommunity $^q\bar{\alpha}_j^{\mathbf{I}}$ diversity of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity $^q\bar{\alpha}_j^{\mathbf{I}}$ diversity is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

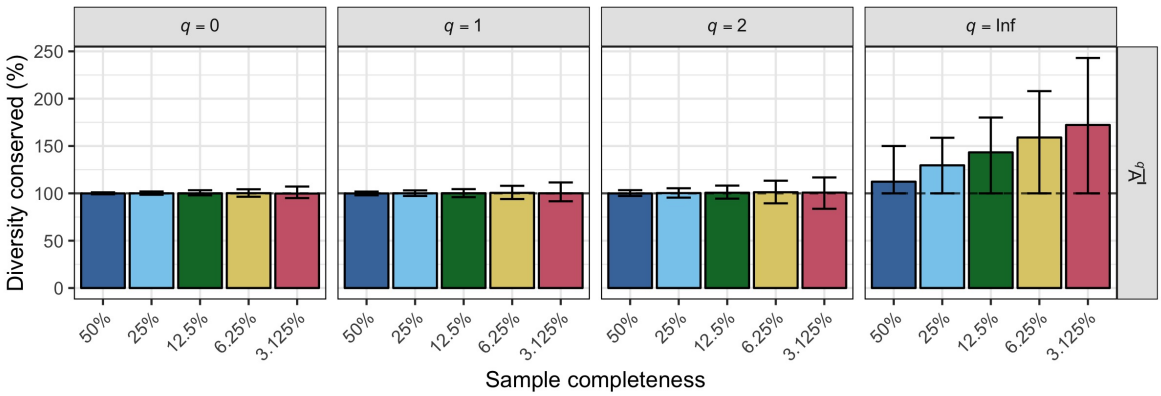


Figure 5.25: The effect of subsampling *subcommunities* on values of *normalised metacommunity alpha* ($^q\bar{A}^{\mathbf{I}}$) diversity at different values of q and sampling effort: The bar plot shows the relative value of $^q\bar{A}^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

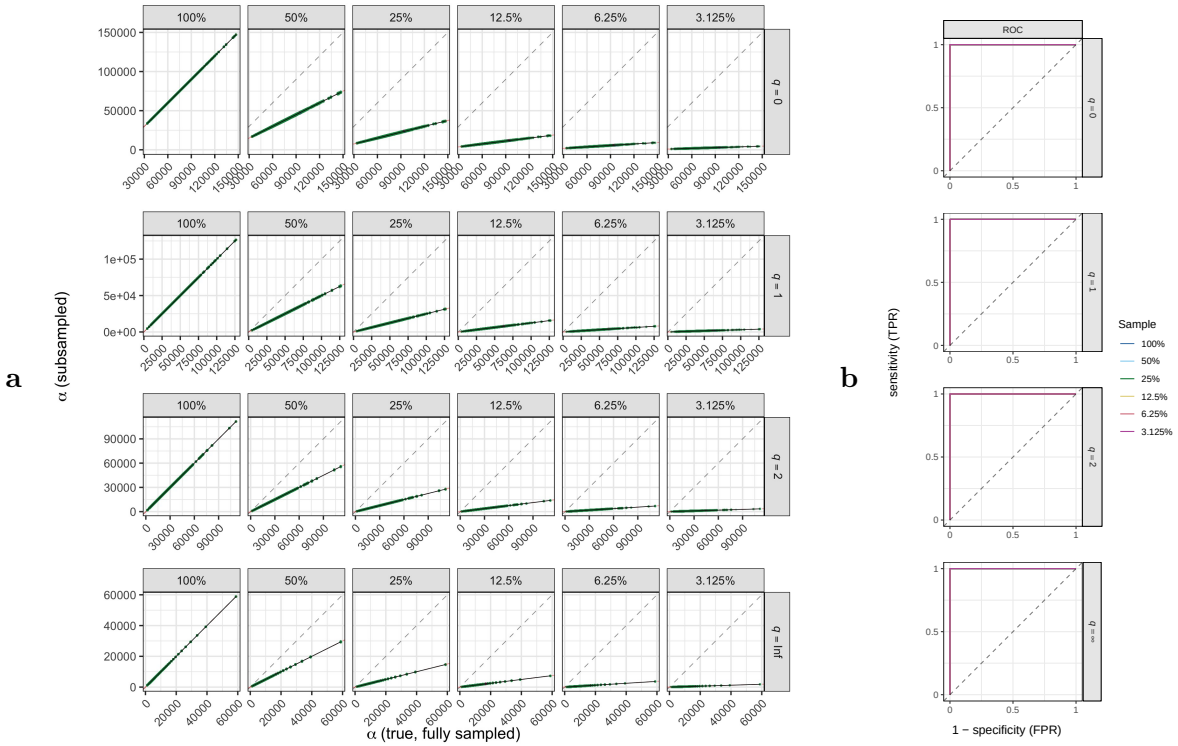


Figure 5.26: The effect of subsampling *subcommunities* on values of *raw subcommunity alpha* (${}^{q}\alpha_j^{\mathbf{I}}$) diversity at different values of q and sampling effort: (a) The true subcommunity ${}^{q}\alpha_j^{\mathbf{I}}$ diversity of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, subcommunity ${}^{q}\alpha_j^{\mathbf{I}}$ diversity is fully conserved when points follow the dashed grey line, whereas the dashed red line scales by the fraction sampled. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

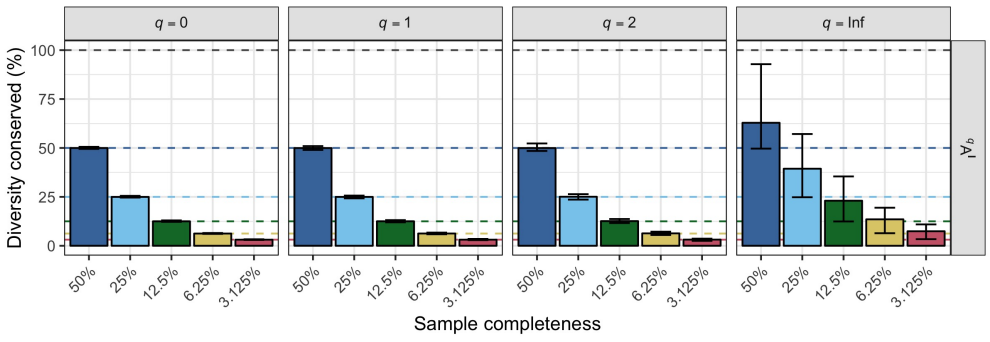


Figure 5.27: The effect of subsampling *subcommunities* on values of *raw metacommunity alpha* (${}^{q}A^{\mathbf{I}}$) diversity at different values of q and sampling effort: The bar plot shows the relative value of ${}^{q}A^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line, whereas coloured dashed lines at 50%, 25%, 12.5%, 6.25%, and 3.125% are scaled by fraction sampled.

5.4.6.2 Sampling by individual

As expected, ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ is increasingly underestimated with decreasing sampling effort (Figure 5.28a). This effect is particularly strong when $q = 0$, since ${}^{0}\bar{\alpha}_j^{\mathbf{I}}$ weights rare and common species equally, whereas higher values of q place more emphasis on the more common species. Therefore, as fewer individuals are sampled from each subcommunity, and rare species are missed, ${}^{0}\bar{\alpha}_j^{\mathbf{I}}$ is underestimated and more variable. At 50% sampling effort, R^2 values are 0.839, 0.871, 0.877, and 0.809 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.1). Nevertheless, ROC curves show that for sampling effort greater than 25%, rank (in the sense of identifying the most interesting subcommunities) is reasonably well conserved for all values of q, though $q = 0$ is best overall (Figure 5.28b). The AUC values at 50% sampling effort are 0.978, 0.982, 0.982, and 0.974 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.2). Again, as expected, ${}^{q}\alpha_j^{\mathbf{I}}$ behaves in a similar manner (Figure 5.29a), with R^2 values of 0.842, 0.909, 0.908, and 0.862 for $q \in \{0, 1, 2, \infty\}$, respectively (Table C.1). In contrast to ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$, the ROC curves and AUC values show that, when sampling effort is less than 25%, $q = 0$ is just as poor as other values of q at identifying subcommunities with high values of ${}^{q}\alpha_j^{\mathbf{I}}$ (Figure 5.29b and Table C.2).

Note that at $q = \infty$, horizontal striations are visible whereby subcommunities with different values of fully sampled ${}^{\infty}\bar{\alpha}_j^{\mathbf{I}}$ appear equivalent under lower sampling effort. This is because ${}^{\infty}\bar{\alpha}_j^{\mathbf{I}}$ considers only the most abundant (or most ordinary) species in each subcommunity. And at low sampling effort, it is more likely that the relative abundance of the most dominant species in multiple subcommunities will be equal. This is also observed for ${}^{\infty}\alpha_j^{\mathbf{I}}$ (Figure 5.29a).

At the metacommunity level, ${}^{q}\bar{A}^{\mathbf{I}}$ describes the effective number of species in each subcommunity in isolation, averaged across all subcommunities. These values are best conserved when $q = \infty$ where only the most dominant species in the metacommunity is considered (Figure 5.30), whereas, for $q \leq 2$, where measures are more sensitive to rare species, ${}^{q}\bar{A}^{\mathbf{I}}$ is increasingly underestimated. Likewise for metacommunity ${}^{q}A^{\mathbf{I}}$ (Figure 5.31).

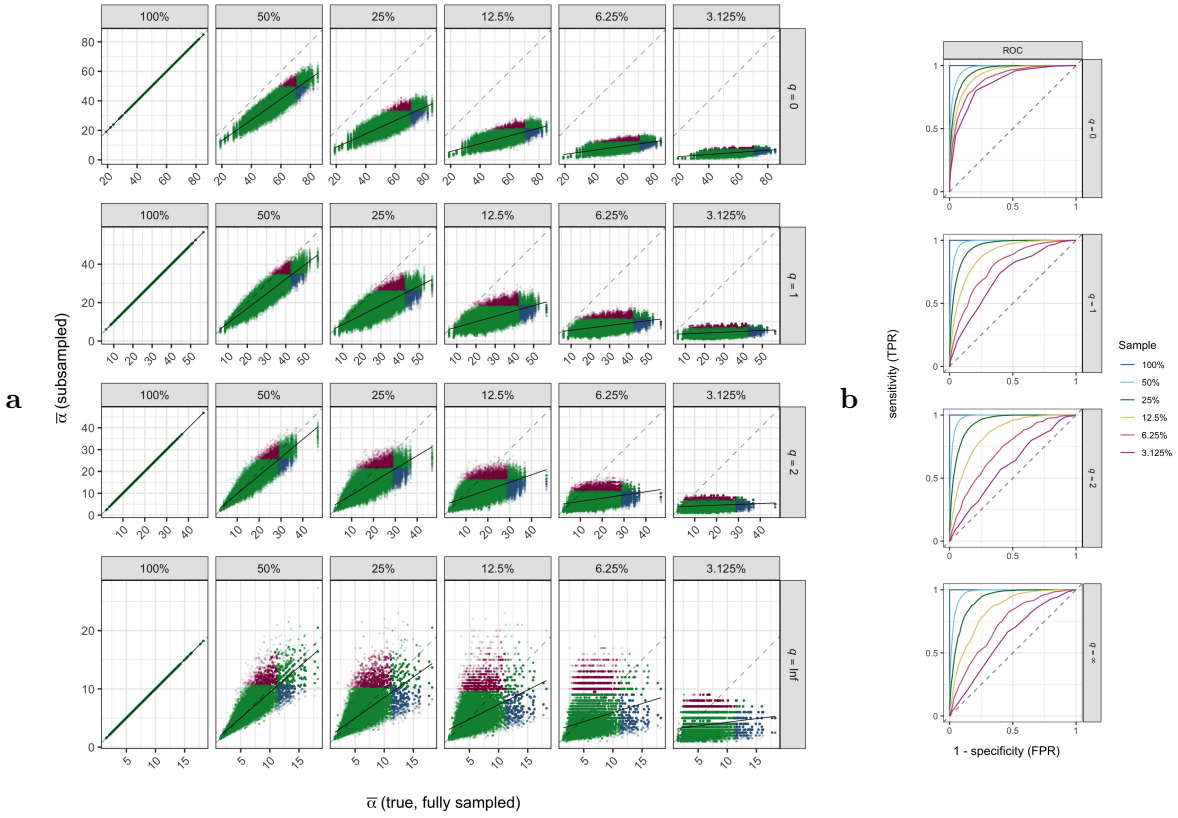


Figure 5.28: The effect of subsampling *individuals* on values of *normalised subcommunity alpha* ($^q\bar{\alpha}_j^{\mathbf{I}}$) diversity at different values of q and sampling effort: (a) The true subcommunity $^q\bar{\alpha}_j^{\mathbf{I}}$ diversity of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity $^q\bar{\alpha}_j^{\mathbf{I}}$ diversity is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

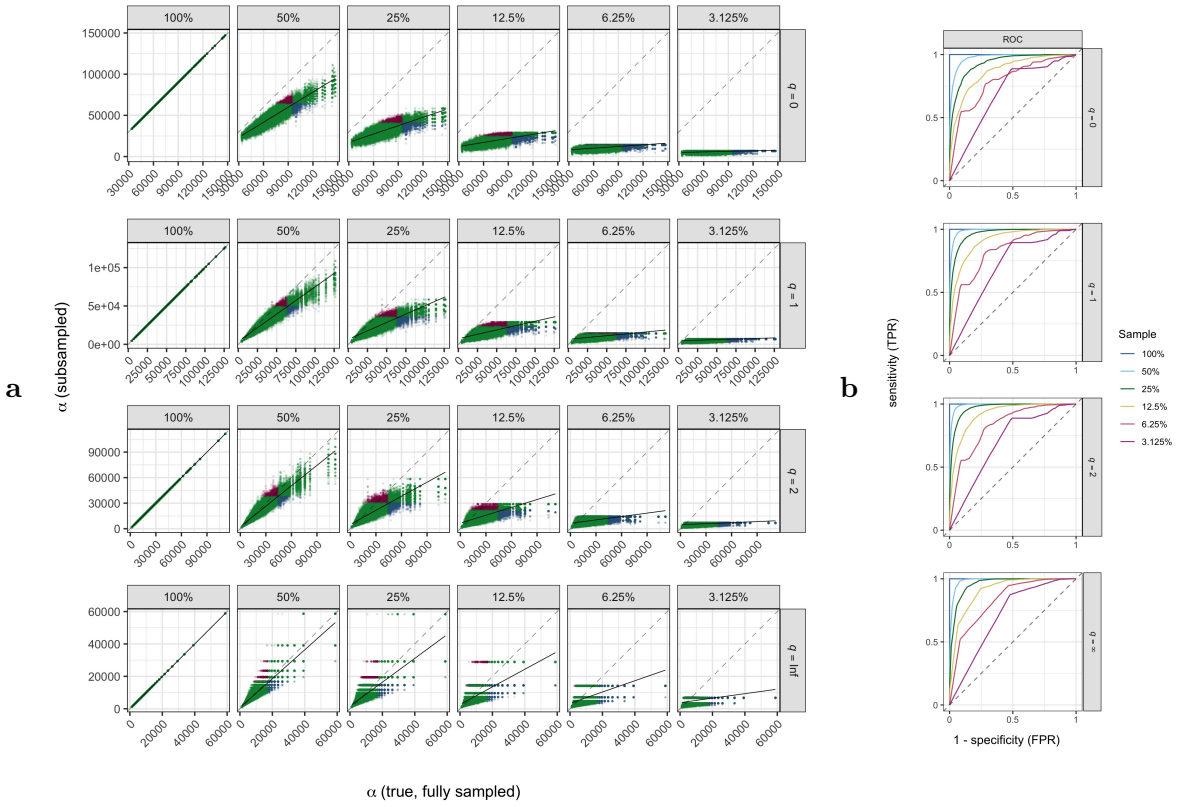


Figure 5.29: The effect of subsampling *individuals* on values of *raw subcommunity alpha* ($^q\alpha_j^{\mathbf{I}}$) diversity at different values of q and sampling effort: (a) The true subcommunity $^q\alpha_j^{\mathbf{I}}$ diversity of each quadrat (x-axis) is compared to partially sampled values (y-axis), where sampling effort is shown above each plot. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error, and subcommunity $^q\alpha_j^{\mathbf{I}}$ diversity is fully conserved when points follow the dashed grey line. Each point corresponds to a single 20 m $\times$ 20 m quadrat in the BCI study site. The colour of each point shows whether the most important communities are still identified as being important after subsampling. True positives and true negatives are coloured green, false positives and false negatives are coloured red and blue, respectively. (b) ROC curves highlight the trade-off between specificity and sensitivity as the threshold x varies between 0 and 100%. A ROC curve following the dashed grey line reflects results no better than chance.

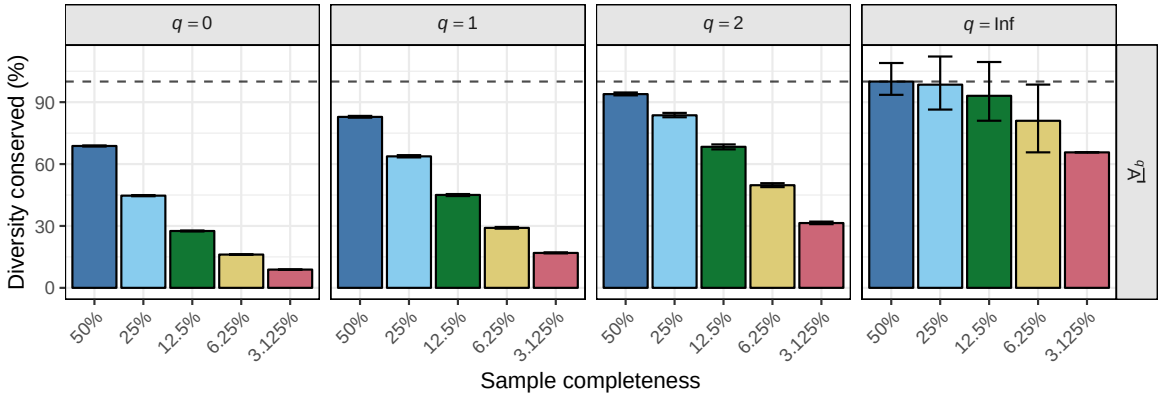


Figure 5.30: The effect of subsampling *individuals* on values of *normalised metacommunity alpha* ($^{q}\bar{A}^{\mathbf{I}}$) diversity at different values of q and sampling effort: The bar plot shows the relative value of $^{q}\bar{A}^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

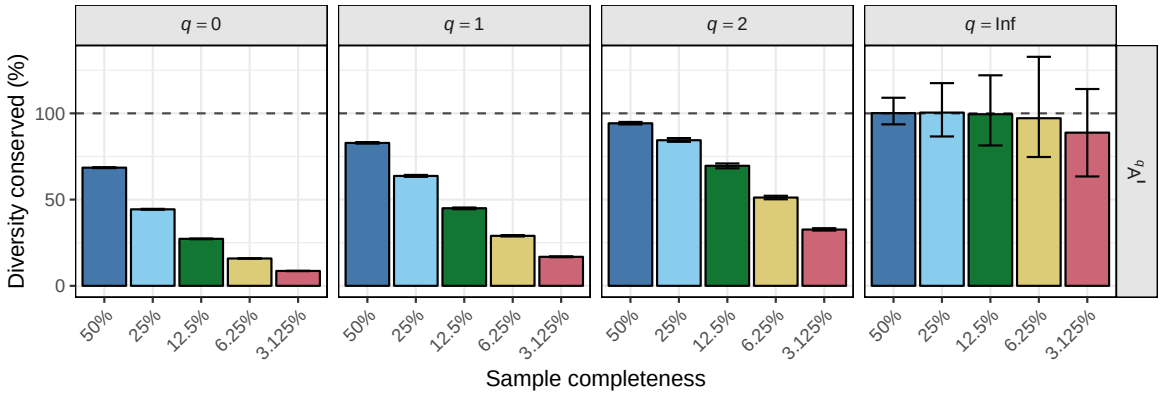


Figure 5.31: The effect of subsampling *individuals* on values of *raw metacommunity alpha* ($^{q}A^{\mathbf{I}}$) diversity at different values of q and sampling effort: The bar plot shows the relative value of $^{q}A^{\mathbf{I}}$ (averaged over 100 repeats) at each level of sampling effort. Error bars indicate a 95% confidence interval. The value of diversity is fully conserved when bars reach the dashed grey line.

5.5 Systematic bias in measures of beta diversity

Figures 5.32 and 5.33 show the relative sampling accuracy of Reeve et al.'s (2016) subcommunity- and metacommunity-level measures of diversity when sampling by individual, Figures 5.34 and 5.35 show relative sampling accuracy when sampling by subcommunity. This summary statistic shows how well each measure is conserved with decreasing sampling effort, allowing general trends to be observed. When sampling by individual, relative sampling accuracy is calculated as the ratio of subsampled to fully sampled diversity values, *e.g.* for normalised alpha, $\overline{SA} = \frac{\bar{\alpha}_{\text{samp}}}{\bar{\alpha}_{\text{full}}}$. When sampling by subcommunity, the relative sampling accuracy is calculated in the same way for all measures except for ${}^{q}\rho_j^{\mathbf{Z}}$, ${}^{q}\beta_j^{\mathbf{Z}}$, and ${}^{q}\alpha_j^{\mathbf{Z}}$. As discussed above, subsampled ${}^{q}\alpha_j^{\mathbf{Z}}$ and ${}^{q}\beta_j^{\mathbf{Z}}$, are scaled by f, whereas subsampled ${}^{q}\rho_j^{\mathbf{Z}}$ is scaled by $\frac{1}{f}$. Therefore, for these measures, the relative sampling accuracy is calculated as $\overline{SA}_\alpha = \frac{\alpha_{samp/fraction}}{\alpha_{full}}$, $\overline{SA}_\rho = \frac{\rho_{samp/fraction}}{\rho_{full}}$ and $\overline{SA}_\beta = \frac{\beta_{samp \times fraction}}{\beta_{full}}$. The pink ribbon denotes a 95% confidence interval and the dashed line highlights a relative sampling accuracy of one, where subsampled diversity is able to accurately predict the diversity of each fully sampled subcommunity.

Chao et al. (2005) note that incomplete data results in a systematic bias to measured diversity. That is, decreasing sampling effort should cause sampled estimates of beta diversity to increase, as rare species are likely to be missed. Indeed, these results show that for $q \leq \infty$, subsampling by individual does cause subcommunity ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ and ${}^{q}\beta_j^{\mathbf{I}}$ diversities to be overestimated as sampling effort is decreased (Figure 5.32). Equivalently, since ${}^{q}\bar{\beta}_j^{\mathbf{I}} = {}^{1}/_{{}^{q}\bar{\rho}_j^{\mathbf{I}}}$ (and ${}^{q}\beta_j^{\mathbf{I}} = {}^{1}/_{{}^{q}\rho_j^{\mathbf{I}}}$), subcommunity ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ and ${}^{q}\rho_j^{\mathbf{I}}$ are increasingly underestimated with reduced sampling effort, and similarly, at the metacommunity-level (Figure 5.33).

However, when subsampling by subcommunity, these effects are no longer present. In fact, as described previously, sampled estimates of beta diversity are very good predictors of fully sampled values at both the subcommunity and metacommunity levels. This is summarised in Figures 5.34 and 5.35. In fact, these observations are reversed – though on a much smaller scale – such that ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ and ${}^{q}\beta_j^{\mathbf{I}}$ are slightly underestimated when sampling effort is low (and ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ and ${}^{q}\rho_j^{\mathbf{I}}$ are slightly overestimated). Furthermore, under extreme subsampling, where the number of subcommunities sampled (N') is less than ${}^{q}\bar{B}^{\mathbf{I}}$ of the fully sampled metacommunity, then the sampled ${}^{q}\bar{B}^{\mathbf{I}}$ must necessarily be an underestimates, since they are constrained by N'.

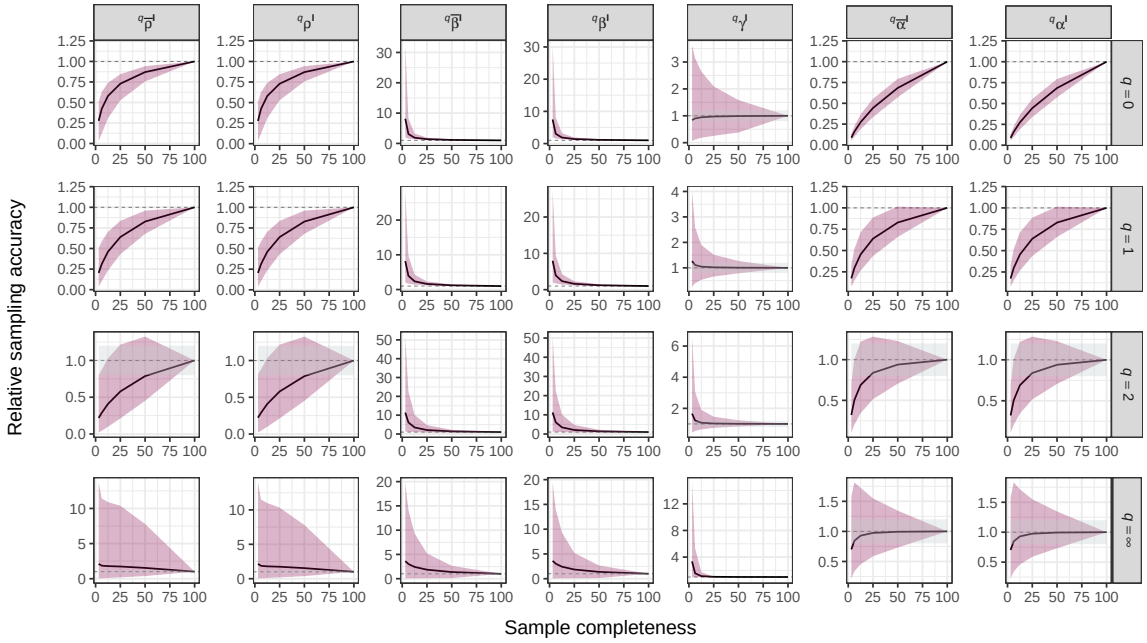


Figure 5.32: The effect of subsampling *individuals* on measured *subcommunity* diversity. Relative sampling accuracy against sampling effort for different measures of diversity, values of q: The shaded ribbon denotes the 95% confidence interval. Diversity values are fully conserved when points follow the dashed grey line.

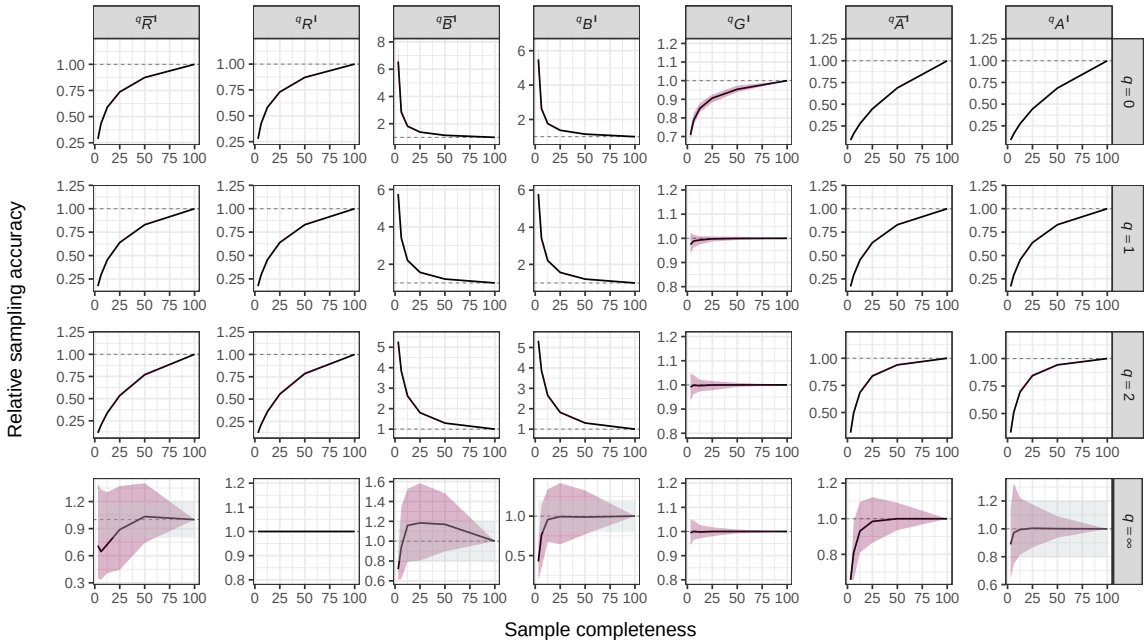


Figure 5.33: The effect of subsampling *individuals* on measured *metacommunity* diversity. Relative sampling accuracy against sampling effort for different measures of diversity, values of q: The shaded ribbon denotes the 95% confidence interval. Diversity values are fully conserved when points follow the dashed grey line.

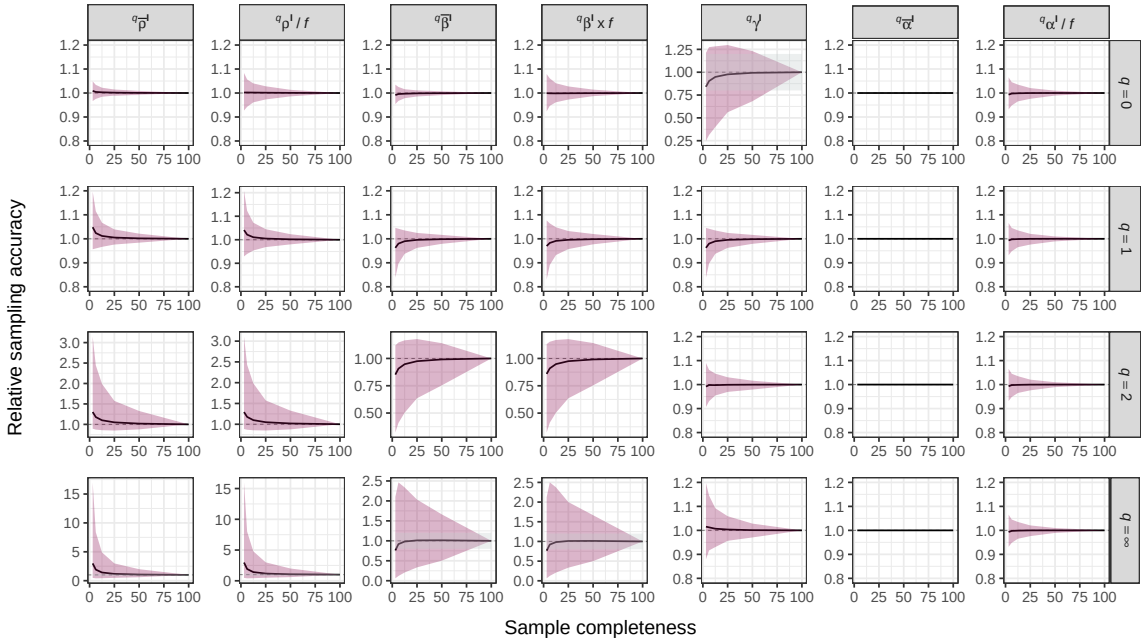


Figure 5.34: The effect of subsampling *subcommunities* on measured *subcommunity* diversity. Relative sampling accuracy against sampling effort for different measures of diversity, values of q, and sampling effort: The shaded ribbon denotes the 95% confidence interval. Diversity values are fully conserved when points follow the dashed grey line.

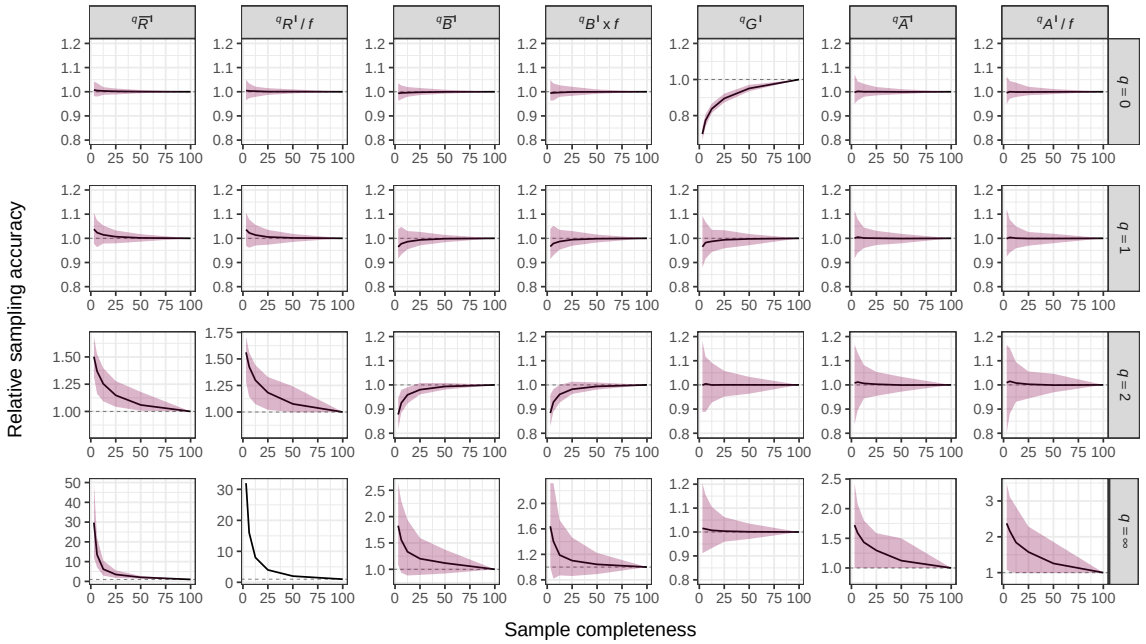


Figure 5.35: The effect of subsampling *subcommunities* on measured *metacommunity* diversity. Relative sampling accuracy against sampling effort for different measures of diversity, values of q, and sampling effort: The shaded ribbon denotes the 95% confidence interval. Diversity values are fully conserved when points follow the dashed grey line.

5.6 Conclusion

In this chapter, the Barro Colorado Island Forest dynamics plot dataset was used to assess the performance of Reeve et al.'s (2016) measures of diversity under sampling pressure. This is a fully sampled dataset, which makes it ideal for such purposes. In these analyses, data were subsampled in two different ways: (1) sampling by individual, where a fraction of individuals are sampled from every subcommunity, and (2) sampling by subcommunity, where a fraction of whole subcommunities are sampled from the metacommunity. In both cases, subsamples were taken at 50%, 25%, 12.5%, 6.25%, and 3.125% and therefore, at each level of sampling effort the same number of individuals were recognised in each sampling strategy. To determine how well measured-diversity was conserved (when sampling by individual or subcommunity), the diversity of each subsampled dataset was compared to that of the fully sampled dataset.

Figures 5.36 and 5.37 collate all ROC curves from earlier in the chapter, corresponding to diversity analyses subsampled by individual and subcommunities, respectively. Table C.2 lists the corresponding AUC (Area Under Curve) values for each plot. These results show how well each measure is able to identify the most interesting – *i.e.* most diverse (${}^{q}\bar{\alpha}_j^{\mathbf{I}}$), most distinctive (${}^{q}\beta_j^{\mathbf{I}}$), most representative (${}^{q}\bar{\rho}_j^{\mathbf{I}}$), and so on – under reduced sampling effort, for each of the sampling strategies. Comparing these results shows that (for the BCI study site) it is overwhelmingly better to fully sample a number of subcommunities in their entirety, than partially sample every subcommunity. Moreover, when the measured values themselves are examined, similar results are observed, as described subsequently.

Assessment of diversity measures

In agreement with Chao et al.'s (2005) observations – that beta diversity is increasingly overestimated with decreasing sampling effort – results showed that when sampling by individual, ${}^{q}\bar{\beta}_j^{\mathbf{I}}$ and ${}^{q}\beta_j^{\mathbf{I}}$ are indeed overestimated, for $q \leq \infty$ (Figure 5.32); Equivalently, ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ and ${}^{q}\rho_j^{\mathbf{I}}$ were underestimated, since ${}^{q}\bar{\rho}_j^{\mathbf{I}} = {}^{1}/_{{}^{q}\bar{\beta}_j^{\mathbf{I}}}$ and ${}^{q}\rho_j^{\mathbf{I}} = {}^{1}/_{{}^{q}\beta_j^{\mathbf{I}}}$. At the metacommunity level, where measures are calculated as a weighted power mean of the subcommunity level values, equivalent results were observed for ${}^{q}B^{\mathbf{I}}$, ${}^{q}\bar{B}^{\mathbf{I}}$, ${}^{q}R^{\mathbf{I}}$, and ${}^{q}\bar{R}^{\mathbf{I}}$ (Figure 5.33). On the other hand, when sampling by subcommunity, normalised beta diversities (${}^{q}\bar{\rho}_j^{\mathbf{I}}$ and ${}^{q}\bar{\beta}_j^{\mathbf{I}}$) accurately predicted fully sampled values for all levels of sampling effort, for $q \leq 1$ (Figure 5.34). Raw beta diversities (${}^{q}\rho_j^{\mathbf{I}}$ and ${}^{q}\beta_j^{\mathbf{I}}$), when scaled by sampling effort, yield accurate estimates of the fully sampled values for all levels of sampling effort, for $q \leq 1$. Likewise, for the metacommunity level measures (Figure 5.35).

It is known from sampling theory that sample-based methods preserve any autocorrelation in species occurrence, and therefore result in fewer species being observed across multiple samples than might be observed using individual-based methods (Gotelli & Colwell, 2001). When the metadata is examined, at the lowest level of sampling effort (3.125%) when only

39 subcommunities are sampled from a total of 1250, the average number of species per subcommunity remained approximately the same (dropping from 54.051 in the fully sampled dataset to 53.899 at 3.125% sampling effort, Table 5.2). Conversely, for when sampling by individual, the average number of species drops from 54.051 to 4.661, per subcommunity (Table 5.1). It is perhaps unsurprising then, that measured values were better conserved when sampling by subcommunity, than when sampling by individual.

Subcommunity ${}^{q}\gamma_j^{\mathbf{I}}$ is also better conserved when sampling by subcommunities than when sampling by individual. However, at the metacommunity level – since ${}^{q}G^{\mathbf{I}}$ is only dependent on the abundance of species in the metacommunity – as expected, ${}^{q}G^{\mathbf{I}}$ performs well under both sampling strategies. Following this trend, ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ is perfectly conserved at all levels of sampling effort, whereas ${}^{q}\alpha_j^{\mathbf{I}}$ performs well when linearly scaled by f, the fraction sampled. On the other hand, as when subsampling by individual, both ${}^{q}\bar{\alpha}_j^{\mathbf{I}}$ and ${}^{q}\alpha_j^{\mathbf{I}}$ are greatly underestimated as rarefaction issues emerge. At the metacommunity level, results are comparable, though some variation is observed due to averaging across subcommunities.

Potential limitations

Though results are encouraging, particularly when subsampling by subcommunity, additional work is required before any generalisations can be made on these findings. As already discussed, observed species richness is affected by many issues, including the number of individuals in the area, how easily they can be observed, the size of the region under study, the heterogeneity of habitat types, and sampling effort. For example, when sampling by subcommunity, as the size of the sampled area increases, species richness will increase rapidly as new species are encountered. Once the area is large enough to contain all of the species associated with a particular habitat type, species richness will increase further as species from differing habitats are more likely to be encountered (Kohn & Walsh, 1994; Steinmann et al., 2011). In other words, species richness increases non-linearly with the size of the area sampled, as different species pools are aggregated (Gotelli & Colwell, 2001). Here, the 50 ha study site was partitioned into 20 m × 20 m quadrats. In addition to this, it might be useful to consider not only how different quadrat sizes might affect these results, but also the distribution of habitat types.

But what about the size of the dataset itself? The BCI Forest dynamics plot is a large, fully sampled dataset, which here was partitioned into 1250 quadrats. However, in a much smaller study in which there might only be a small number of subcommunities, then diversity measures based on metacommunity abundance might be expected to be more affected by reduced sampling effort (due to the small numbers alone). If so, there may be a point below which it might be better to sample by individuals rather than subcommunities. In which case, care must be taken when considering the individuals over which diversity is being calculated. Individual-based sampling strategies assume that each individual is equally likely

to be encountered (or not to be encountered), which is approximately true in this case, where the individuals sampled are trees. However, for communities in which some species are more difficult to sample than others, this assumption might be more problematic. Though beyond the scope of this study, possible sources of bias include (but are not limited to): behaviour-related sampling-bias due to boldness or aggression (Carter et al., 2012; Biro, 2013), and non-detection bias due to rareness or camouflage (Hefley et al., 2013).

Recommendations

In future work it should be possible to make generalised predictions of how each of these measures behave under decreased sampling effort. In addition to this, it would be interesting to determine how well each measure – particularly the beta diversity measures – compare to those commonly used in the literature. For this particular dataset, it is reasonable to hypothesise that Reeve et al.'s (2016) beta diversity measures – based on subcommunity to metacommunity comparisons – should produce more reliable estimates than traditional measures, under reduced sampling effort. Since, in a dataset this large, a reliable estimate of metacommunity diversity can be calculated at even 3% of subcommunities ($N = 1250$). In this work, however, which explored rarefaction issues associated with the measurement of the subcommunity and metacommunity diversity of the BCI study site, by simulating two different subsampling strategies. Results showed that given the same resource, it is better to sample by subcommunity than by individual. In particular and as predicted, Reeve et al.'s (2016) beta diversity measures were well conserved, even at very low sampling effort.

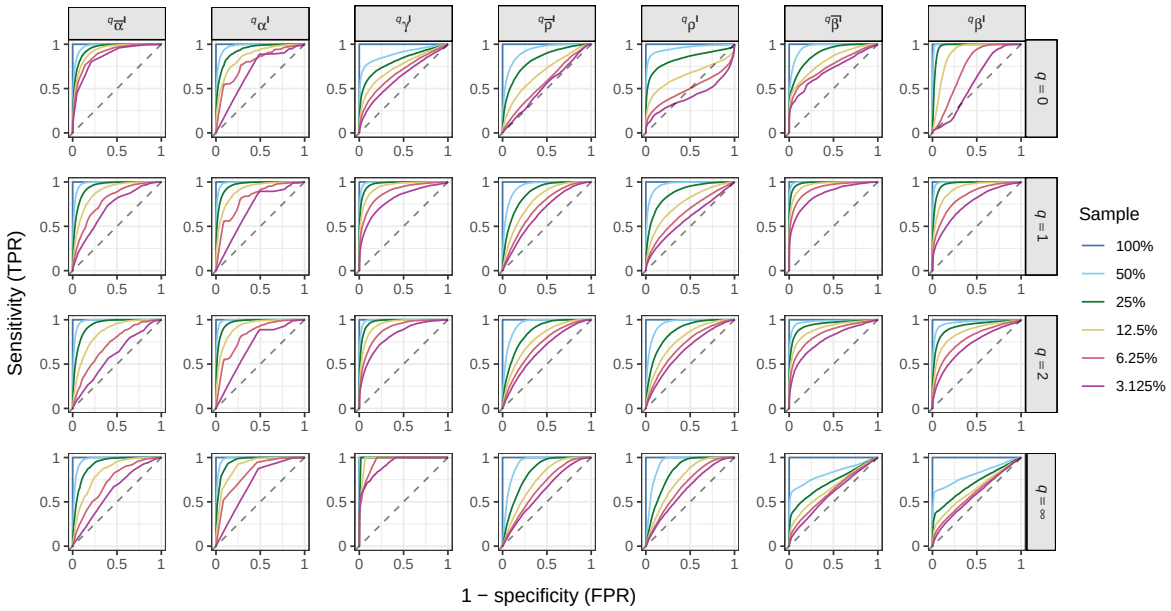


Figure 5.36: ROC curves showing the effect of subsampling *individuals* on the trade-off between specificity and sensitivity for different measures of diversity and values of q.

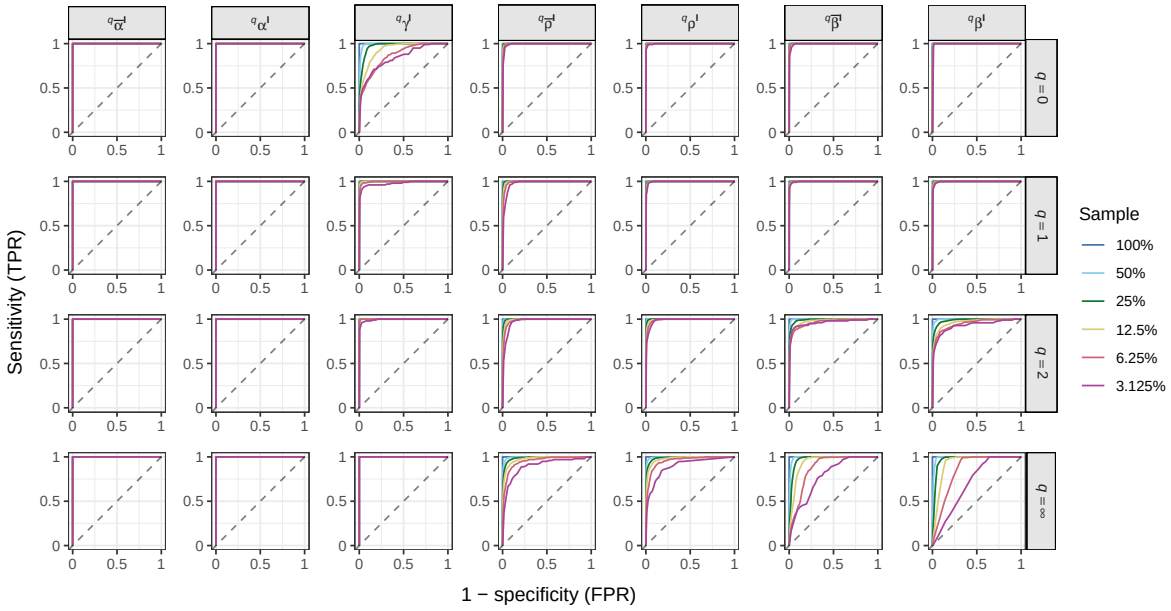


Figure 5.37: ROC curves showing the effect of subsampling *subcommunities* on the trade-off between specificity and sensitivity for different measures of diversity and values of q.

Chapter 6

Discussion

> 'Diversities are mere numbers and should be distinguished from the theories which they support.'
>
> — Hill (1973)

The concept of biodiversity is a complex and multifaceted one, resulting from many interacting processes. Given these difficulties, many groups have attempted to find alternative means with which to quantify diversity. Historically, Shannon (1948) entropy (a measure of information content) and Simpson's (1949) index of concentration (a measure of probability) have been used to quantify this diversity. However, these measures don't satisfy the various mathematical properties required for a measure of diversity to behave as one would expect. It turns out that these measures, when transformed, exist on a continuum of measures known as Hill (1973) numbers. These values are quantified in units of 'effective numbers', which satisfy the aforementioned properties and allow diversity to be expressed intuitively.

Comparing the diversity of one community to another can be considered in terms of Whittaker's (1960) alpha (within subcommunities), beta (between subcommunities), and gamma (across subcommunities, or equivalently within the metacommunity) diversities. Recently, Reeve et al. (2016) proposed new measures of metacommunity alpha, beta, and gamma diversities, that emerge naturally from Rényi's (1961) generalisations of Shannon (1948) entropy and Leinster & Cobbold's (2012) expression of similarity-sensitive diversity. Critically, these measures can be decomposed into their subcommunity contributions, revealing the metacommunity's underlying subcommunity structure and dynamics. Consequently, this framework is able to identify unique signals in population structure that traditional measures might not detect, particularly because the framework (following Leinster & Cobbold, 2012) incorporates a similarity matrix ($\mathbf{Z}$), which allows any kind of similarity to be considered, be it taxonomic, genetic, phenotypic, and so on. The work in this thesis is based on this new framework of diversity measures.

In addition to the average diversity across subcommunities, ${}^{q}\bar{A}^{\mathbf{Z}}$ (equivalent to Whittaker's alpha) and the diversity of the metacommunity as a whole, ${}^{q}G^{\mathbf{Z}}$ (equivalent to Whittaker's gamma), Reeve et al.'s (2016) framework can be used to calculate the average redundancy of subcommunities (${}^{q}R^{\mathbf{Z}}$), the average representativeness of subcommunities (${}^{q}\bar{R}^{\mathbf{Z}}$), the average distinctiveness of subcommunities (${}^{q}B^{\mathbf{Z}}$), and the effective number of distinct subcommunities

(${}^{q}\bar{B}^{\mathbf{Z}}$). At the subcommunity level, these measures can be deconstructed to calculate the diversity of a subcommunity in isolation (${}^{q}\bar{\alpha}_{j}^{\mathbf{Z}}$), the contribution per individual in a subcommunity toward metacommunity diversity (a novel measure of gamma diversity, ${}^{q}\gamma_{j}^{\mathbf{Z}}$), the redundancy of a subcommunity (${}^{q}\rho_{j}^{\mathbf{Z}}$), the representativeness of a subcommunity (${}^{q}\bar{\rho}_{j}^{\mathbf{Z}}$), the distinctiveness of a subcommunity (${}^{q}\beta_{j}^{\mathbf{Z}}$), and an estimate of the effective number of distinct subcommunities (${}^{q}\bar{\beta}_{j}^{\mathbf{Z}}$).

In Chapter 2, each of these measures were described in detail alongside simple examples. I also briefly discussed a software package, `rdiversity`, which I developed in R to calculate these measures. This package is available on CRAN.

6.1 Case studies

In Chapter 3, three distinct case studies were selected to showcase the flexibility of Reeve et al.'s (2016) framework. The challenge was to determine whether or not this framework could be usefully applied to such different datasets, each requiring very different signals to be detected. The focus of the first case study was a classic biodiversity problem, to examine the compositional structure of the Barro Colorado Island (BCI) Forest dynamics plot. The second two case studies developed novel diversity-based solutions to more unusual problems.

In contrast to other traditional measures of diversity, Reeve et al.'s (2016) framework can be used to investigate subcommunity structure. This was illustrated in the first case study (Section 3.2), by examining the spatial and temporal diversity of the BCI Forest dynamics plot – a fully sampled 50 ha study site. First, spatial diversity was measured in the naïve-type case (where species are considered distinct), so that the properties of each measure could be clearly observed. Two areas were identified as being particularly interesting, and unrepresentative of the study site: invasive species in the top-left of the site, as well as a swamp in the centre-left with unique plant types. Following this, taxonomic diversity was assessed by incorporating Shimatani's (2001) measure of taxonomic distance. From these results, only the swamp continued to contribute highly to the biodiversity (${}^{q}\gamma_{j}^{\mathbf{Z}_{tax}}$) of the study site, highlighting the evolutionary uniqueness of swamp-based species. The area in the top-left, though distinct (${}^{q}\beta_{j}^{\mathbf{Z}_{tax}}$), no longer contributed as highly, comprising invasive species that were found to be taxonomically similar to the rest of the plot. Temporal diversity was then calculated as ${}^{q}B^{\mathbf{I}}$, a measure of turnover, where the swamp was identified as having the highest turnover in species composition over time. This was confirmed by separate analyses showing the spatial representativeness (${}^{q}\bar{\rho}_{j}^{\mathbf{I}}$) of the swamp increasing at each time point, and examining the species composition of the swamp.

The second case study (Section 3.3) examined the population demographics of England and Wales from 2001 census data. This dataset was selected for the ease with which results could be

verified after the analysis. Two datasets were compared, each comprising the total population of England and Wales in 2001, but differing in the way individuals were partitioned into age groups and geographic regions: the parish dataset had a high spatial resolution but fewer age classes, whilst the CAS ward dataset had a lower spatial resolution but a higher number of age classes. To account for these differences, similarity was scaled to behave equivalently between the two different datasets. This was done by calibrating a scaling parameter to a pre-defined effective number of age classes, ${}^{q}G^{\mathbf{Z}} = 8$. Low subcommunity representativeness (${}^{q}\bar{\rho}_j^{\mathbf{Z}}$) was used to identify areas with unusual age distributions such as those inhabited by the young adults in proximity to universities, and retirement communities dominated by the elderly.

Finally, the third case study (Section 3.4) investigated the transmission of antimicrobial resistance (AMR) phenotypes in sympatric human and animal host populations. Subcommunity redundancy (${}^{q}\rho_j^{\mathbf{Z}}$) and distinctiveness (${}^{q}\beta_j^{\mathbf{Z}}$) were used to detect emerging resistance in host populations, with similarity defined from phenotypic resistance profiles. Agreeing with published results, it was found that host populations were not well mixed and the animal population was unlikely to have been the source of antimicrobial resistance in the human population. These results show that diversity-based methods can be used to tackle problems that are not typically diversity related. More generally, these studies highlight the flexibility and robustness of Reeve et al.'s (2016) framework of diversities, by demonstrating how it can easily be extended to handle new diversity problems by defining similarity to suit each problem.

6.2 Phylogenetic beta diversity

Measures of phylogenetic beta diversity can be categorised into two groups: tree-based measures such as Unifrac (the unique fraction distance, Lozupone & Knight, 2005) and Phylosor (the phylogenetic Sørensen index, Bryant et al., 2008), and distance-based measures such as MPD (mean pairwise distance, Webb, 2000) and MNTD (mean nearest taxon distance, Webb et al., 2002). In Chapter 4, I developed new methods of phylogenetic diversity analysis that extend Reeve et al.'s (2016) framework and quantify phylogenetic beta diversity from both of these perspectives: tree-based measures ($tree$) that build on Leinster & Cobbold's (2012) measure of phylogenetic similarity; and distance-based measures (PPD_l and PPD_e). I compared these new measures to those commonly used in the literature and assessed their robustness using different phylogenetic simulations. Results showed that all of these new measures were better able to detect phylogenetic signal in community structure than traditional measures in almost all circumstances (varying the number of tips in the phylogeny, the number of subcommunities, evolutionary rate, whether a phylogeny is ultrametric or non-ultrametric, whether data is incidence-based or abundance-based, nestedness vs. turnover, and so on).

To conclude this chapter, I applied these new measures of phylogenetic beta diversity to a familiar problem – to investigate the transmission of antimicrobial resistance between human and animal populations, to determine whether epidemic strains of DT104 were maintained separately or transmitted freely between hosts (Section 4.6). By excluding the evolutionary history prior to the epidemic, it was possible to examine only the information directly related to possible host-to-host transmission during the epidemic. These results were validated against published results, again showing that epidemic strains of DT104 were maintained separately in human and animal populations.

There was no time to do a similar investigation, comparing commonly used genetic diversity measures to those based on Reeve et al.'s (2016) framework, but this would be an obvious next step. Nonetheless, it seems intuitive to apply these measures directly to the genetic data that was used to generate the phylogenetic trees. This would avoid possible sources of error associated with generating a phylogeny and show whether there is any difference in using phylogenetic vs. genetic diversity measures. A small case study in Appendix D illustrates how Reeve et al.'s (2016) framework can be used to quantify genetic diversity, using measures analogous to the phylogenetic distance-based beta diversities developed in Chapter 4. A straightforward relationship between phylogenetic and genetic diversity should exist, and so future work might include an extensive analyses of how these genetic diversity measures compare to the phylogenetic distance-based beta diversities, and more generally to measures of genetic beta diversity commonly used in the literature.

6.3 Sampling properties

In this thesis, most studies utilised large, often fully sampled datasets. However, this is not usually the case. Incomplete data and low sample completeness is a common problem and as a result, estimators are often required. In Chapter 5, I used the fully sampled Barro Colorado Island Forest dynamics plot to determine the effect of undersampling on Reeve et al.'s (2016) diversity framework. The 50 ha site was partitioned into 20 m$\times$ 20 m quadrats to assess the robustness of Reeve et al.'s (2016) framework of measures to subsampling. Specifically, I investigated how well measured diversity is preserved, how well order is preserved between subcommunities, and how well correlated fully sampled and subsampled diversity measures are, under subsampling. In addition to this, I examined whether it was better to (1) partially sample every subcommunity, or (2) fully sample a proportion of subcommunities. Subsamples were taken at 50%, 25%, 12.5%, 6.25%, and 3.125% sample completeness in order to determine how each measure performed under each sampling strategy. Results showed that given the same amount of sampling effort, it was always better to sample by subcommunity than to sample by individual, and doing so results in accurate diversity estimates, even for low sample completeness.

In Chapter 4, I investigated how phylogenetic diversity measures were affected by subsampling species from the tree, and found that the framework was better than other measures, although still not very powerful. It would be interesting to consider how phylogenetic beta diversities are affected by subsampling. However, a practical problem was presented to me, where field epidemiologists were interested in knowing how large an outbreak of a disease was from incomplete data. This is really a gamma diversity problem, but nonetheless, it seemed worthwhile investigating as a preliminary case study. The results of this study are presented in Appendix E, where sampling pressure is applied to viral genetic and phylogenetic data in the context of the 2001 UK FMDV outbreak, to determine the best way of inferring the total size of the epidemic from incomplete data.

6.4 Conclusion

In conclusion, this family of diversity measures is powerful when compared to measures commonly found in the literature, is easily adaptable to new research domains, and can be used with poorly sampled data whilst still achieving a good level of accuracy. Obvious next steps include:

- A similar comparison to commonly used measures of genetic and genomic diversity as was done here with phylogenetic beta diversities. This should be relatively straight forward, though relatively time consuming given the number of genetic diversity measures in the literature and the computational time required to perform these analyses;
- A deeper investigation of the sampling properties of the measures on more varied datasets – both other Center for Tropical Forest Science (CTFS) study plots as well as datasets in other research domains and potentially, simulated datasets where known signals are investigated (as was done in Chapter 4); and
- In the longer term, the development of suitable estimators of the whole family of diversity measures from sampled data.

APPENDIX A

Some derivations

The Rényi-divergence of order α (Equation 3.3 in Rényi, 1961) is written:

$$I_\alpha(\mathcal{Q}|\mathcal{P}) = \frac{1}{\alpha - 1} \log_2 \left(\sum_{k=1}^{n} \frac{q_k^\alpha}{p_k^{\alpha-1}} \right) \tag{A.1a}$$

which he defines as "the information of order α obtained if the distribution $\mathcal{P}$ is replaced by distribution $\mathcal{Q}$. We can rearrange this formula such that:

$$= \left(\sum_{k=1}^{n} \frac{q_k^\alpha}{p_k^{\alpha-1}} \right)^{\frac{1}{\alpha-1}} \tag{A.1b}$$

$$= \left(\sum q_k \left(\frac{q_k^\alpha}{p_k} \right)^{\alpha-1} \right)^{\frac{1}{\alpha-1}} \tag{A.1c}$$

where $\alpha \neq 1$. Then, substituting our standard notation:

$$= \left(\sum_i \left(\bar{P}_{.j}\right)_i \left(\frac{\left(\bar{P}_{.j}\right)_i}{p_i} \right)^{q-1} \right)^{\frac{1}{q-1}} \tag{A.1d}$$

$$= \left(\sum_i \left(\bar{P}_{.j}\right)_i \left(\frac{p_i}{\left(\bar{P}_{.j}\right)_i} \right)^{1-q} \right)^{\frac{1}{q-1}} \tag{A.1e}$$

$$= \left(\left(\sum_i \left(\bar{P}_{.j}\right)_i \left(\frac{p_i}{\left(\bar{P}_{.j}\right)_i} \right)^{1-q} \right)^{\frac{1}{1-q}} \right)^{-1} \tag{A.1f}$$

This is equivalent to the weighted powermean:

$$= \left(\mathrm{M}_{1-q} \left(\bar{P}_{.j}, \frac{p}{\bar{P}_{.j}} \right) \right)^{-1} \tag{A.1g}$$

$$= \mathrm{M}_{1-q} \left(\bar{P}_{.j}, \frac{\bar{P}_{.j}}{p} \right)^{-1} \tag{A.1h}$$

which is Reeve et al.'s (2016) subcommunity $\bar{\beta}$ diversity, or *the effective number of distinct subcommunities in a metacommunity.*

APPENDIX B

Phylogenetic beta diversity

B.1 The effect of varying evolutionary rates on subcommunity structure

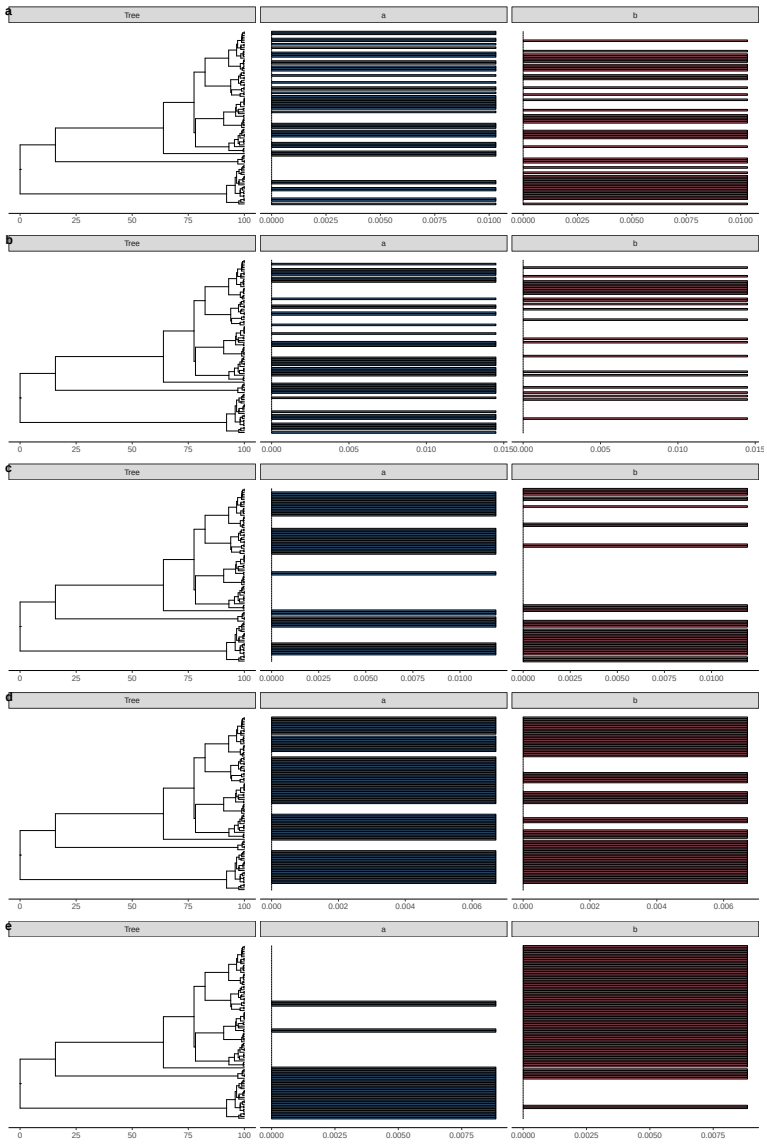


Figure B.1: Example population structure produced by the qualitative-independent (qi) partition strategy: where transition rates are (a) 1, (b) 0.2, (c) 0.1, (d) 0.02, (e) 0.01.

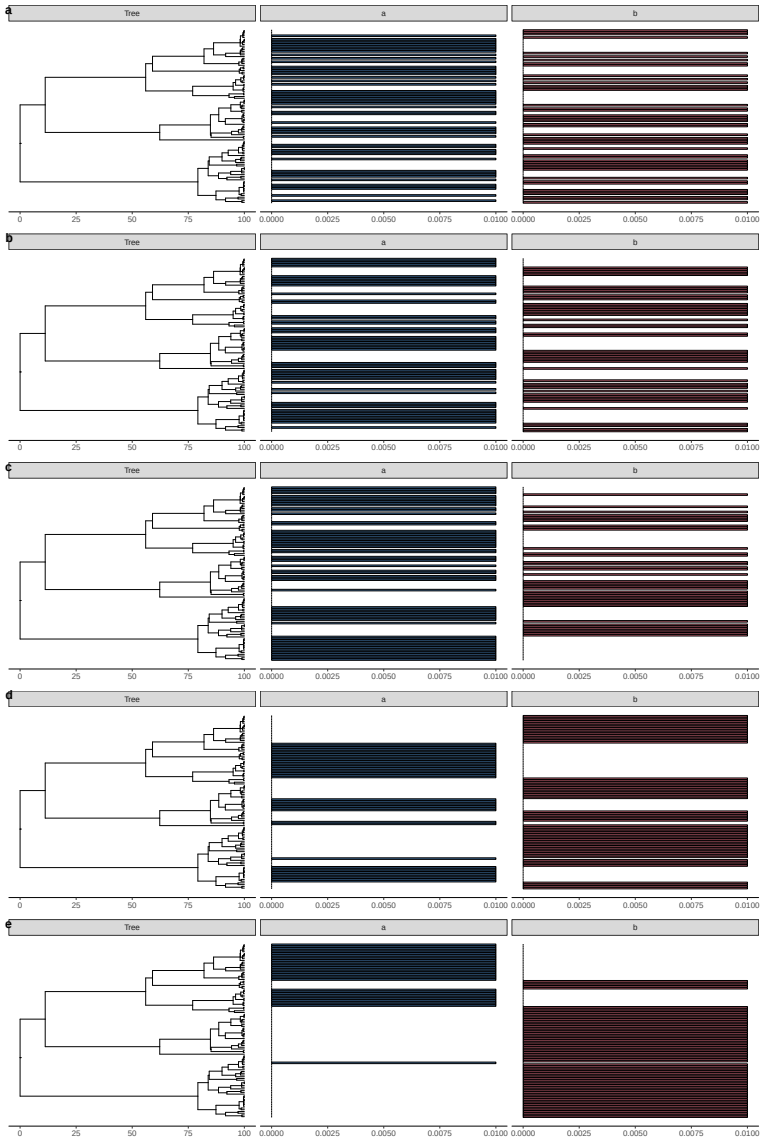


Figure B.2: Example population structure produced by the qualitative-dependent (qd) partition strategy: where transition rates are (a) 1, (b) 0.2, (c) 0.1, (d) 0.02, (e) 0.01.

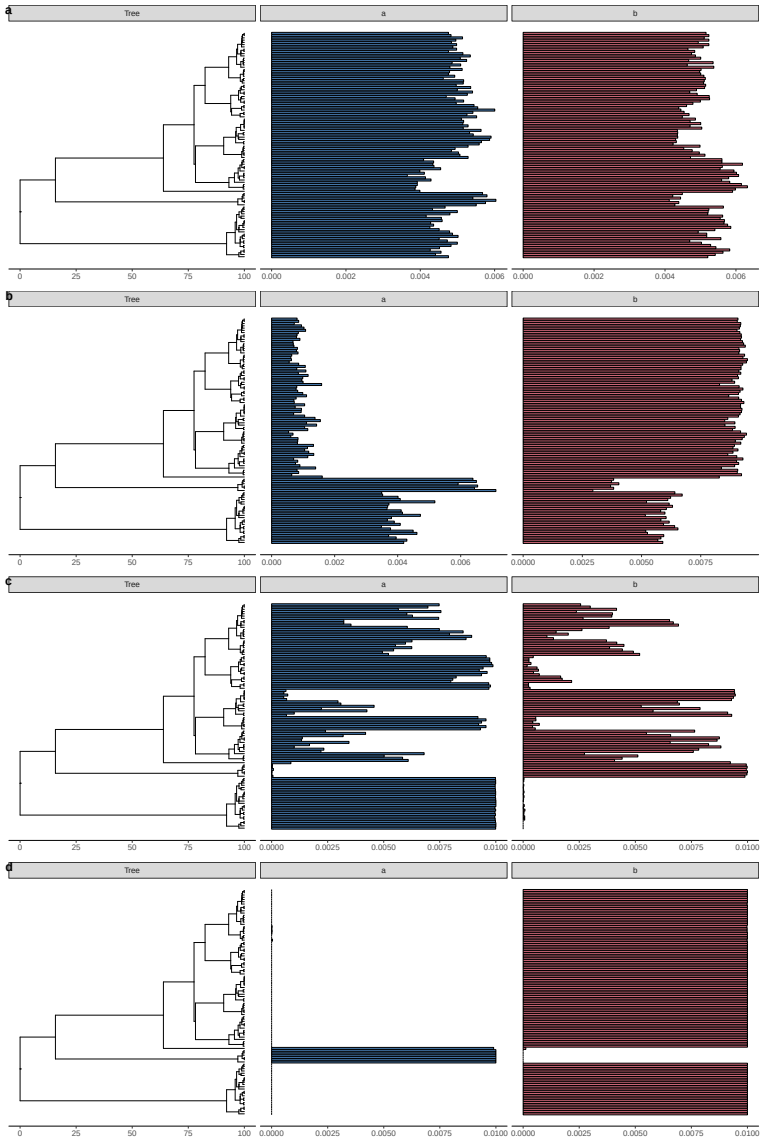


Figure B.3: Example population structure produced by the quantitative-dependent proportional (qdp) partition strategy: where transition rates are (a) 0.002, (b) 0.02, (c) 0.2, and (d) 2.

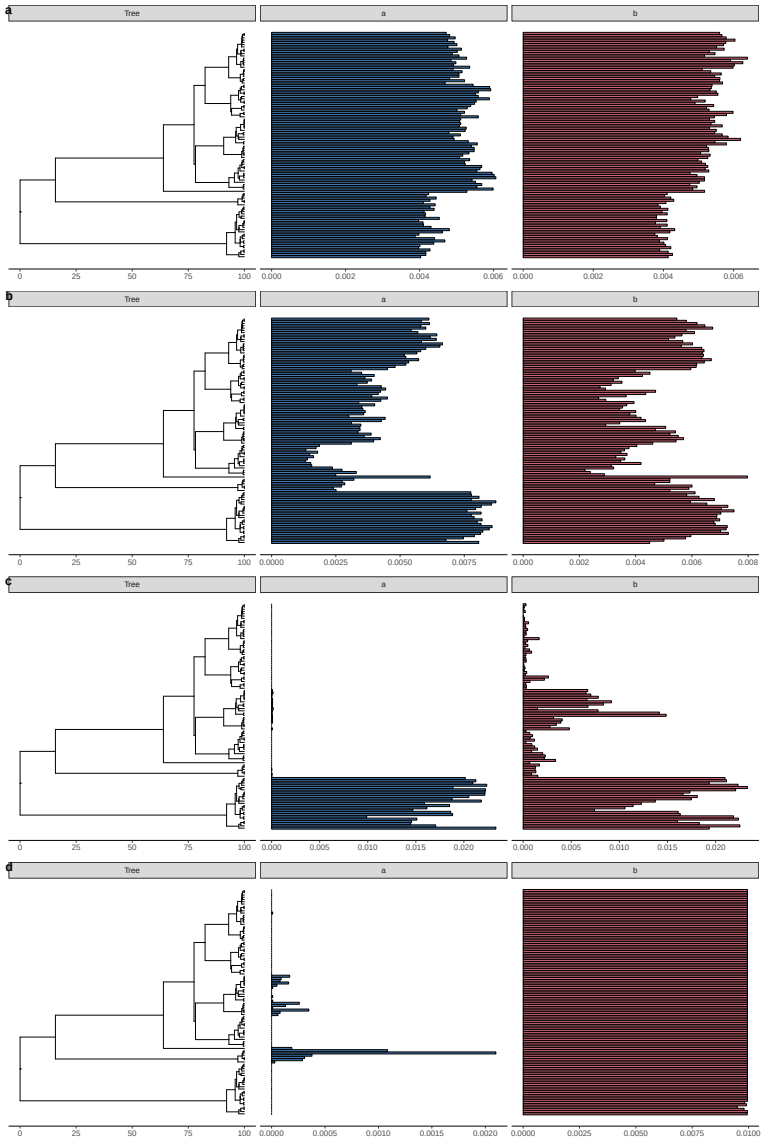


Figure B.4: Example population structure produced by the quantitative-independent proportional (qip) partition strategy: where transition rates are (a) 0.002, (b) 0.02, (c) 0.2, and (d) 2.

B.2 Power values

B.2.1 Experiment 1: Two subcommunities

Table B.1: Experiment 1 - Two subcommunities: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qd_nu	0.75	0.73	0.91	0.91	0.89	0.89
qdb_nu	1.00	0.98	1.00	1.00	1.00	1.00
qdp_nu	1.00	0.98	1.00	1.00	1.00	1.00
qi_nu	0.70	0.66	0.84	0.84	0.86	0.86
qib_nu	0.93	0.91	0.88	0.88	0.91	0.91
qip_nu	0.99	1.00	1.00	1.00	1.00	1.00
qd_u	0.98	0.98	1.00	1.00	1.00	1.00
qdb_u	0.98	0.98	0.96	0.96	0.98	0.98
qdp_u	1.00	1.00	1.00	1.00	1.00	1.00
qi_u	0.98	0.98	1.00	1.00	1.00	1.00
qib_u	0.98	0.98	0.94	0.94	0.96	0.96
qip_u	1.00	1.00	1.00	1.00	1.00	1.00

B.2.2 Experiment 2: Nested subcommunities

Table B.2: Experiment 2 - Nestedness: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qi_nu	0.33	0.23	0.34	0.34	0.39	0.39
qib_nu	0.52	0.41	0.45	0.45	0.57	0.57
qip_nu	0.93	0.38	0.89	0.89	0.91	0.91
qi_u	0.55	0.55	0.52	0.52	0.60	0.60
qib_u	0.69	0.69	0.48	0.48	0.59	0.59
qip_u	0.98	0.98	0.93	0.93	0.96	0.96

B.2.3 Experiment 3: Varying tree size

Table B.3: Experiment 3 - 10 tips: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qd_nu	0.11	0.11	0.05	0.04	0.12	0.12
qdb_nu	0.33	0.38	0.07	0.07	0.34	0.34
qdp_nu	0.40	0.39	0.31	0.31	0.45	0.45
qi_nu	0.08	0.07	0.02	0.02	0.08	0.08
qib_nu	0.21	0.17	0.02	0.02	0.24	0.24
qip_nu	0.38	0.34	0.30	0.30	0.45	0.45
qd_u	0.16	0.16	0.09	0.09	0.20	0.20
qdb_u	0.38	0.35	0.11	0.11	0.32	0.32
qdp_u	0.47	0.47	0.47	0.47	0.59	0.59
qi_u	0.16	0.17	0.09	0.09	0.21	0.21
qib_u	0.30	0.29	0.09	0.08	0.30	0.30
qip_u	0.52	0.52	0.52	0.52	0.57	0.57

Table B.4: Experiment 3 - 25 tips: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qd_nu	0.17	0.20	0.27	0.27	0.30	0.30
qdb_nu	0.59	0.66	0.54	0.54	0.62	0.62
qdp_nu	0.87	0.81	0.89	0.89	0.92	0.92
qi_nu	0.17	0.13	0.20	0.20	0.24	0.24
qib_nu	0.45	0.40	0.38	0.38	0.47	0.47
qip_nu	0.82	0.77	0.88	0.88	0.91	0.91
qd_u	0.45	0.45	0.64	0.64	0.63	0.63
qdb_u	0.69	0.69	0.61	0.61	0.68	0.68
qdp_u	0.88	0.88	0.92	0.92	0.93	0.93
qi_u	0.38	0.38	0.52	0.52	0.50	0.50
qib_u	0.61	0.61	0.39	0.39	0.52	0.52
qip_u	0.86	0.86	0.95	0.95	0.98	0.98

Table B.5: Experiment 3 - 50 tips: PThese results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ (${}^1\bar{R}^{\mathbf{Z}_{tree}}$)	${}^1B^{\mathbf{Z}_{tree}}$ (${}^1R^{\mathbf{Z}_{tree}}$)	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ (${}^1\bar{R}^{\mathbf{Z}_{PPD_l}}$)	${}^1B^{\mathbf{Z}_{PPD_l}}$ (${}^1R^{\mathbf{Z}_{PPD_l}}$)	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ (${}^1\bar{R}^{\mathbf{Z}_{PPD_e}}$)	${}^1B^{\mathbf{Z}_{PPD_e}}$ (${}^1R^{\mathbf{Z}_{PPD_e}}$)
qd_nu	0.49	0.46	0.58	0.58	0.64	0.64
qdb_nu	0.91	0.84	0.89	0.89	0.92	0.92
qdp_nu	0.96	0.95	0.98	0.98	0.98	0.98
qi_nu	0.38	0.34	0.55	0.55	0.55	0.55
qib_nu	0.73	0.70	0.65	0.65	0.73	0.73
qip_nu	0.97	0.89	0.99	0.99	0.98	0.98
qd_u	0.74	0.74	0.91	0.91	0.91	0.91
qdb_u	0.91	0.91	0.87	0.87	0.91	0.91
qdp_u	0.99	0.99	0.99	0.99	0.99	0.99
qi_u	0.70	0.70	0.84	0.84	0.87	0.87
qib_u	0.83	0.83	0.77	0.77	0.82	0.82
qip_u	0.97	0.97	1.00	1.00	1.00	1.00

Table B.6: Experiment 3 - 75 tips: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ (${}^1\bar{R}^{\mathbf{Z}_{tree}}$)	${}^1B^{\mathbf{Z}_{tree}}$ (${}^1R^{\mathbf{Z}_{tree}}$)	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ (${}^1\bar{R}^{\mathbf{Z}_{PPD_l}}$)	${}^1B^{\mathbf{Z}_{PPD_l}}$ (${}^1R^{\mathbf{Z}_{PPD_l}}$)	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ (${}^1\bar{R}^{\mathbf{Z}_{PPD_e}}$)	${}^1B^{\mathbf{Z}_{PPD_e}}$ (${}^1R^{\mathbf{Z}_{PPD_e}}$)
qd_nu	0.65	0.60	0.84	0.84	0.81	0.81
qdb_nu	0.97	0.92	0.94	0.94	0.95	0.95
qdp_nu	1.00	0.98	1.00	1.00	1.00	1.00
qi_nu	0.61	0.59	0.80	0.80	0.79	0.79
qib_nu	0.89	0.82	0.88	0.88	0.91	0.91
qip_nu	0.99	0.98	1.00	1.00	1.00	1.00
qd_u	0.91	0.91	0.98	0.98	0.98	0.98
qdb_u	0.98	0.98	0.96	0.96	0.98	0.98
qdp_u	1.00	1.00	1.00	1.00	1.00	1.00
qi_u	0.91	0.91	0.98	0.98	0.99	0.99
qib_u	0.96	0.96	0.92	0.92	0.94	0.94
qip_u	1.00	1.00	1.00	1.00	1.00	1.00

Table B.7: Experiment 3 - 100 tips: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qd_nu	0.85	0.80	0.91	0.91	0.91	0.91
qdb_nu	1.00	0.96	0.99	0.99	0.99	0.99
qdp_nu	1.00	1.00	1.00	1.00	1.00	1.00
qi_nu	0.71	0.70	0.87	0.87	0.84	0.84
qib_nu	0.98	0.93	0.98	0.98	0.96	0.96
qip_nu	0.99	0.97	0.99	0.99	0.99	0.99
qd_u	0.98	0.98	1.00	1.00	1.00	1.00
qdb_u	0.99	0.99	0.98	0.98	0.98	0.98
qdp_u	1.00	1.00	1.00	1.00	1.00	1.00
qi_u	0.95	0.95	0.98	0.98	0.98	0.98
qib_u	0.97	0.97	0.93	0.93	0.95	0.95
qip_u	1.00	1.00	1.00	1.00	1.00	1.00

B.2.4 Experiment 4: Varying evolutionary rates

Table B.8: Experiment 4 - qualitative rates: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ (${}^1\bar{R}^{\mathbf{Z}_{tree}}$)	${}^1B^{\mathbf{Z}_{tree}}$ (${}^1R^{\mathbf{Z}_{tree}}$)	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ (${}^1\bar{R}^{\mathbf{Z}_{PPD_l}}$)	${}^1B^{\mathbf{Z}_{PPD_l}}$ (${}^1R^{\mathbf{Z}_{PPD_l}}$)	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ (${}^1\bar{R}^{\mathbf{Z}_{PPD_e}}$)	${}^1B^{\mathbf{Z}_{PPD_e}}$ (${}^1R^{\mathbf{Z}_{PPD_e}}$)
qd_nu_1	0.04	0.03	0.05	0.05	0.05	0.05
qi_nu_1	0.07	0.06	0.06	0.06	0.08	0.08
qd_nu_10	0.07	0.05	0.05	0.05	0.07	0.07
qi_nu_10	0.05	0.05	0.06	0.06	0.05	0.05
qd_nu_100	0.98	0.94	0.99	0.99	0.99	0.99
qi_nu_100	0.95	0.89	1.00	1.00	1.00	1.00
qd_nu_5	0.06	0.05	0.07	0.07	0.06	0.06
qi_nu_5	0.08	0.08	0.06	0.06	0.06	0.06
qd_nu_50	0.78	0.80	0.94	0.94	0.92	0.92
qi_nu_50	0.77	0.71	0.89	0.89	0.89	0.89
qd_u_1	0.06	0.06	0.06	0.06	0.05	0.05
qi_u_1	0.04	0.04	0.09	0.09	0.09	0.09
qd_u_10	0.27	0.27	0.44	0.44	0.47	0.47
qi_u_10	0.27	0.27	0.31	0.31	0.34	0.34
qd_u_100	1.00	1.00	1.00	1.00	1.00	1.00
qi_u_100	1.00	1.00	0.99	0.99	1.00	1.00
qd_u_5	0.21	0.21	0.20	0.20	0.20	0.20
qi_u_5	0.12	0.12	0.13	0.13	0.16	0.16
qd_u_50	0.98	0.98	1.00	1.00	1.00	1.00
qi_u_50	0.97	0.97	0.99	0.99	0.99	0.99

Table B.9: Experiment 4 - quantitative rates: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qdb_nu_0.002	0.23	0.24	0.21	0.21	0.24	0.24
qdp_nu_0.002	1.00	1.00	1.00	1.00	1.00	1.00
qib_nu_0.002	0.13	0.16	0.16	0.16	0.16	0.16
qip_nu_0.002	1.00	1.00	1.00	1.00	1.00	1.00
qdb_nu_0.02	0.87	0.85	0.83	0.83	0.89	0.89
qdp_nu_0.02	1.00	1.00	1.00	1.00	1.00	1.00
qib_nu_0.02	0.60	0.60	0.55	0.55	0.65	0.65
qip_nu_0.02	1.00	0.99	1.00	1.00	1.00	1.00
qdb_nu_0.2	1.00	0.98	1.00	1.00	1.00	1.00
qdp_nu_0.2	1.00	0.99	1.00	1.00	1.00	1.00
qib_nu_0.2	0.95	0.91	0.93	0.93	0.95	0.95
qip_nu_0.2	1.00	0.98	1.00	1.00	1.00	1.00
qdb_nu_2	0.99	0.97	0.99	0.99	0.99	0.99
qdp_nu_2	0.98	1.00	1.00	1.00	1.00	1.00
qib_nu_2	0.98	0.95	0.98	0.98	0.98	0.98
qip_nu_2	0.99	0.97	1.00	1.00	1.00	1.00
qdb_u_0.002	0.18	0.18	0.11	0.11	0.16	0.16
qdp_u_0.002	1.00	1.00	1.00	1.00	1.00	1.00
qib_u_0.002	0.09	0.09	0.06	0.06	0.09	0.09
qip_u_0.002	1.00	1.00	1.00	1.00	1.00	1.00
qdb_u_0.02	0.85	0.85	0.78	0.78	0.85	0.85
qdp_u_0.02	1.00	1.00	1.00	1.00	1.00	1.00
qib_u_0.02	0.66	0.66	0.49	0.49	0.62	0.62
qip_u_0.02	1.00	1.00	1.00	1.00	1.00	1.00
qdb_u_0.2	1.00	1.00	0.99	0.99	1.00	1.00
qdp_u_0.2	1.00	1.00	1.00	1.00	1.00	1.00
qib_u_0.2	0.98	0.98	0.91	0.91	0.95	0.95
qip_u_0.2	1.00	1.00	1.00	1.00	1.00	1.00
qdb_u_2	1.00	1.00	1.00	1.00	1.00	1.00
qdp_u_2	1.00	1.00	0.99	0.99	1.00	1.00
qib_u_2	0.98	0.98	0.98	0.98	0.99	0.99
qip_u_2	1.00	1.00	1.00	1.00	1.00	1.00

B.2.5 Experiment 5: Multiple subcommunities

Table B.10: Experiment 5 - two subcommunities: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qd_nu	0.76	0.74	0.94	0.94	0.89	0.89
qdb_nu	0.99	0.96	0.98	0.98	0.99	0.99
qdp_nu	1.00	0.98	1.00	1.00	1.00	1.00
qi_nu	0.73	0.71	0.91	0.91	0.88	0.88
qib_nu	0.96	0.90	0.95	0.95	0.95	0.95
qip_nu	1.00	0.98	1.00	1.00	1.00	1.00
qd_u	0.98	0.98	1.00	1.00	1.00	1.00
qdb_u	1.00	1.00	0.99	0.99	1.00	1.00
qdp_u	1.00	1.00	1.00	1.00	1.00	1.00
qi_u	0.96	0.96	0.98	0.98	0.99	0.99
qib_u	0.96	0.96	0.94	0.94	0.98	0.98
qip_u	1.00	1.00	1.00	1.00	1.00	1.00

Table B.11: Experiment 5 - four subcommunities: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qd_nu	0.45	0.42	0.61	0.61	0.66	0.66
qdb_nu	1.00	1.00	1.00	1.00	1.00	1.00
qdp_nu	1.00	1.00	1.00	1.00	1.00	1.00
qi_nu	0.33	0.36	0.54	0.54	0.60	0.60
qib_nu	1.00	1.00	1.00	1.00	1.00	1.00
qip_nu	1.00	1.00	1.00	1.00	1.00	1.00
qd_u	0.98	0.98	1.00	1.00	1.00	1.00
qdb_u	1.00	1.00	1.00	1.00	1.00	1.00
qdp_u	1.00	1.00	1.00	1.00	1.00	1.00
qi_u	0.88	0.88	0.91	0.91	0.98	0.98
qib_u	1.00	1.00	1.00	1.00	1.00	1.00
qip_u	1.00	1.00	1.00	1.00	1.00	1.00

Table B.12: Experiment 5 - eight subcommunities: These results show how well tree-based and distance-based phylogenetic beta diversity measures are able to detect phylogenetic signal in community structure. Values are denote statistical power (the probability that phylogenetic signal has been correctly detected).

	Tree-based		**Linear, distance-based**		**Exponential, distance-based**	
	${}^1\bar{B}^{\mathbf{Z}_{tree}}$ $({}^1\bar{R}^{\mathbf{Z}_{tree}})$	${}^1B^{\mathbf{Z}_{tree}}$ $({}^1R^{\mathbf{Z}_{tree}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_l}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_l}})$	${}^1B^{\mathbf{Z}_{PPD_l}}$ $({}^1R^{\mathbf{Z}_{PPD_l}})$	${}^1\bar{B}^{\mathbf{Z}_{PPD_e}}$ $({}^1\bar{R}^{\mathbf{Z}_{PPD_e}})$	${}^1B^{\mathbf{Z}_{PPD_e}}$ $({}^1R^{\mathbf{Z}_{PPD_e}})$
qd_nu	0.20	0.16	0.26	0.26	0.28	0.28
qdb_nu	1.00	1.00	1.00	1.00	1.00	1.00
qdp_nu	1.00	1.00	1.00	1.00	1.00	1.00
qi_nu	0.08	0.09	0.12	0.12	0.14	0.14
qib_nu	1.00	1.00	1.00	1.00	1.00	1.00
qip_nu	1.00	1.00	1.00	1.00	1.00	1.00
qd_u	0.83	0.83	0.95	0.95	0.98	0.98
qdb_u	1.00	1.00	1.00	1.00	1.00	1.00
qdp_u	1.00	1.00	1.00	1.00	1.00	1.00
qi_u	0.62	0.62	0.69	0.69	0.76	0.76
qib_u	1.00	1.00	1.00	1.00	1.00	1.00
qip_u	1.00	1.00	1.00	1.00	1.00	1.00

B.3 Power plots

Note that the bar charts presented in this Appendix include measures of raw *tree*-based measures, which correspond to both ${}^1B^{\mathbf{Z}_{tree}}$ and ${}^1R^{\mathbf{Z}_{tree}}$, since calculated power is identical for each of these measures. Similarly, normalised *tree* corresponds to ${}^1\bar{B}^{\mathbf{Z}_{tree}}$ and ${}^1\bar{R}^{\mathbf{Z}_{tree}}$. Likewise, with raw and normalised PPD_l and PPD_e. This is because at $q = 1$, there exists a special case where ${}^1\bar{B}^{\mathbf{Z}} = ({}^1\bar{R}^{\mathbf{Z}})^{-1}$, and ${}^1B^{\mathbf{Z}} = ({}^1R^{\mathbf{Z}})^{-1}$.

B.3.1 Experiment 1: Two subcommunities (detailed)

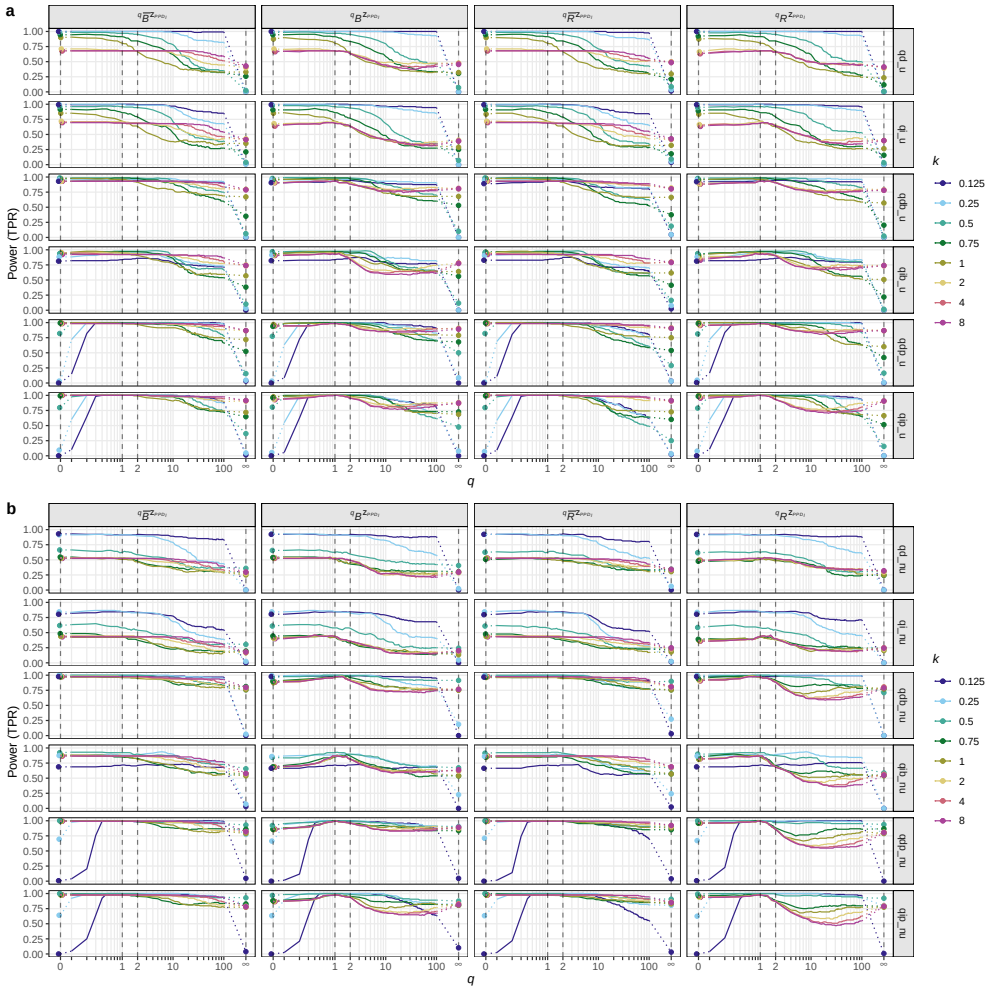


Figure B.5: Plots of power against q, calculated for ${}^q\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^qR^{\mathbf{Z}_{PPD_l}}$, ${}^q\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^qB^{\mathbf{Z}_{PPD_l}}$, in Experiment 1 (for two subcommunities): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

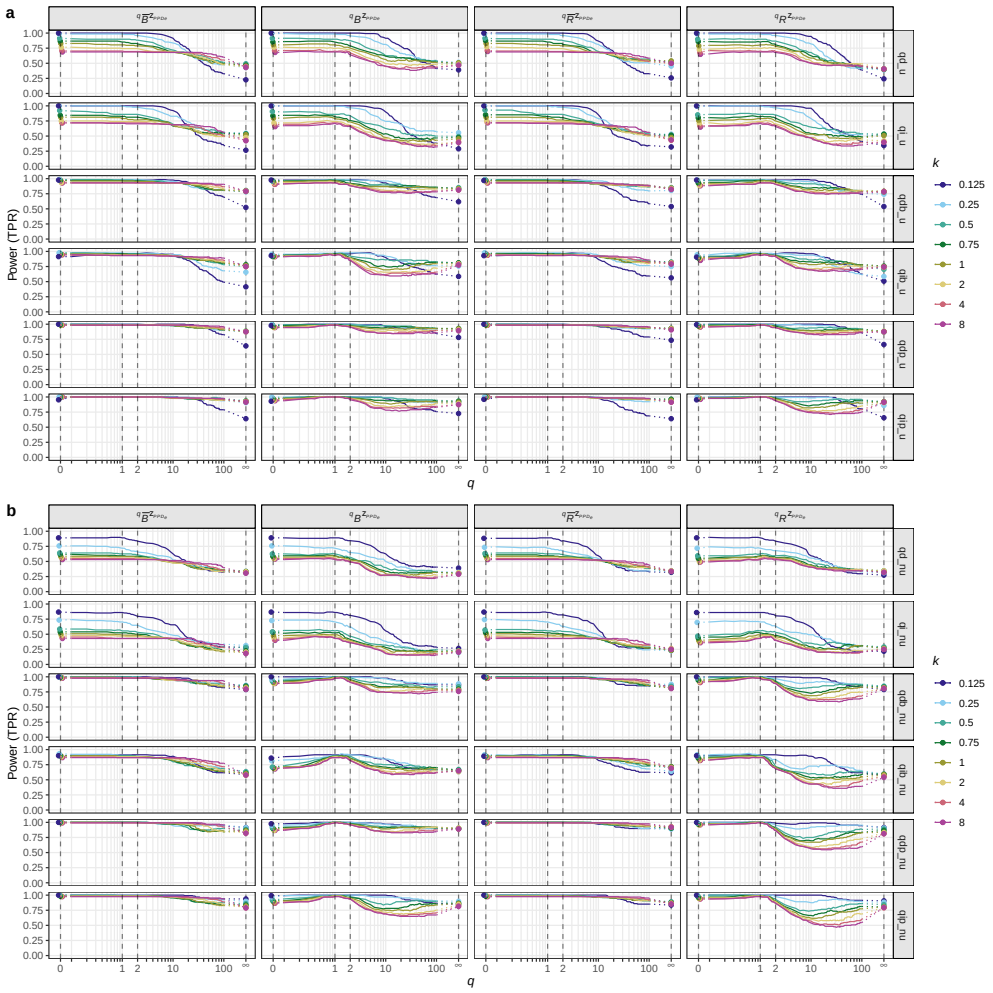


Figure B.6: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 1 (for two subcommunities): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as a exponential transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

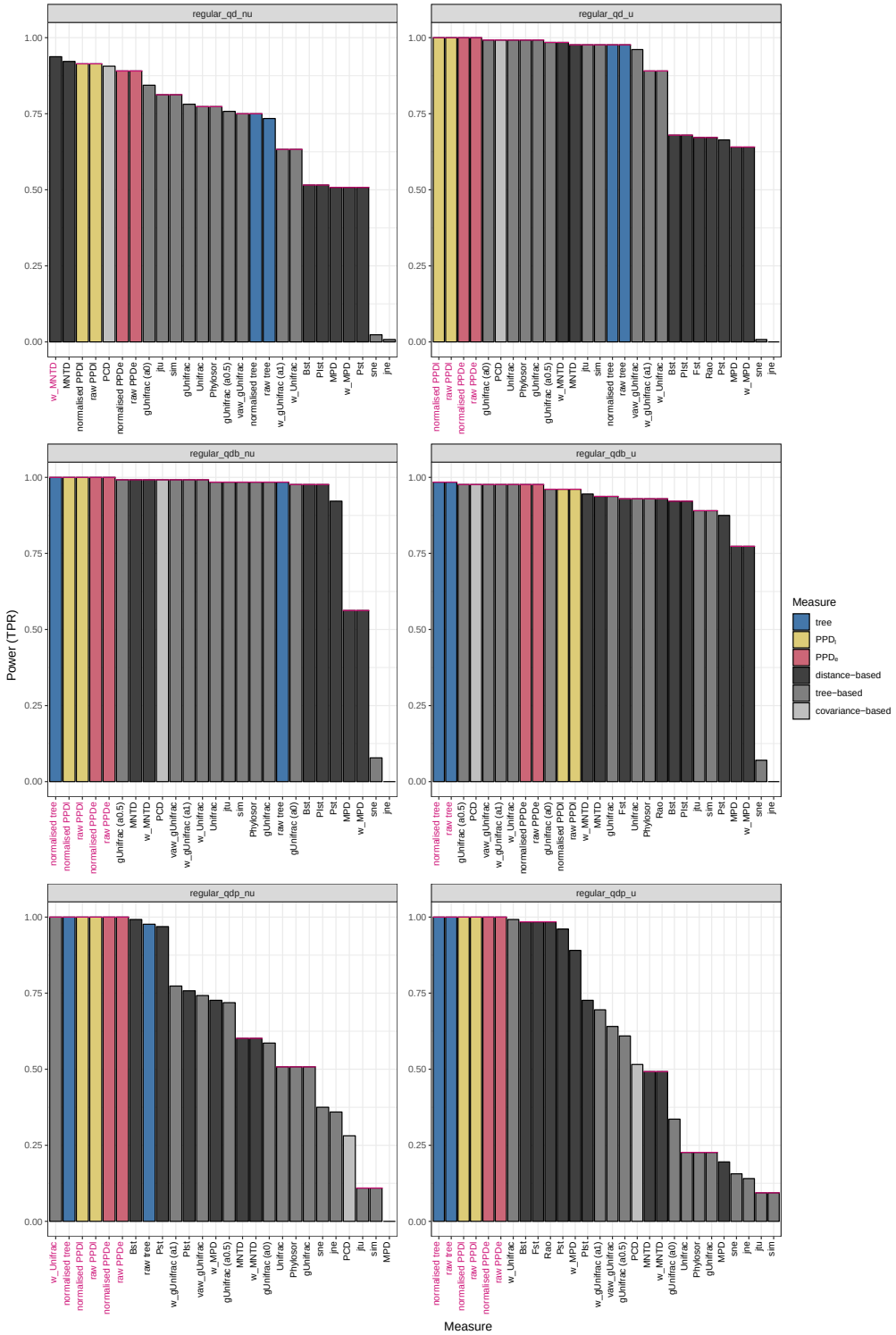


Figure B.7: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 1 (for two subcommunities): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

B.3.2 Experiment 1: Two subcommunities

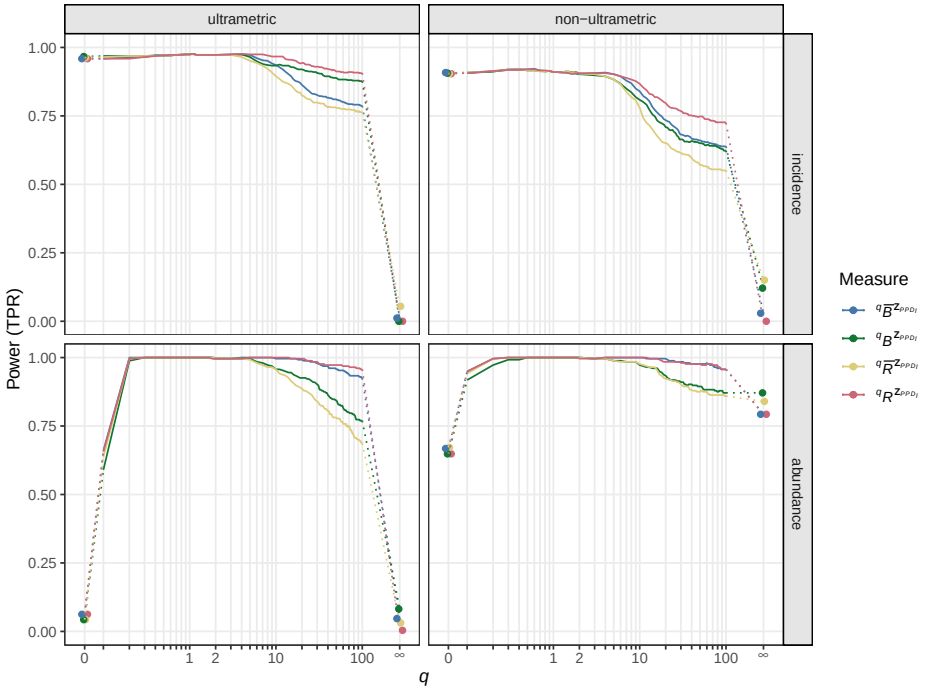


Figure B.8: Summary of power against q, calculated from *linearly* transformed phylogenetic distance-based beta diversity measures: Power is calculated for $^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, $^{q}R^{\mathbf{Z}_{PPD_l}}$, $^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and $^{q}B^{\mathbf{Z}_{PPD_l}}$, and averaged across all subcommunity partitioning strategies within each categorical group.

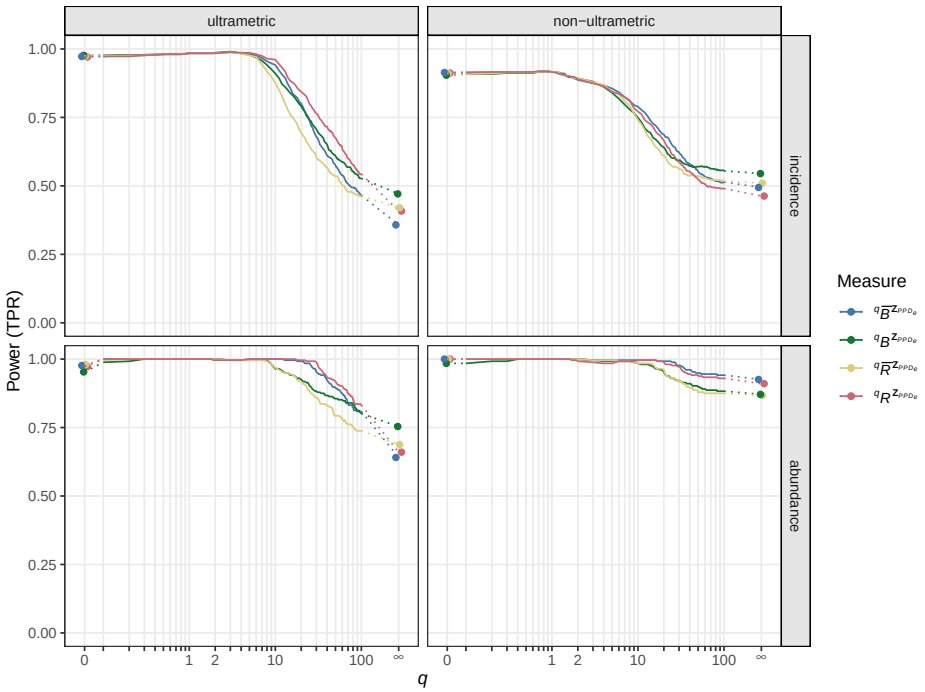


Figure B.9: Summary of power against q, calculated from *exponentially* transformed phylogenetic distance-based beta diversity measures: Power is calculated for $^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, $^{q}R^{\mathbf{Z}_{PPD_e}}$, $^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and $^{q}B^{\mathbf{Z}_{PPD_e}}$, and averaged across all subcommunity partitioning strategies within each categorical group.

B.3.3 Experiment 2: Nested subcommunities

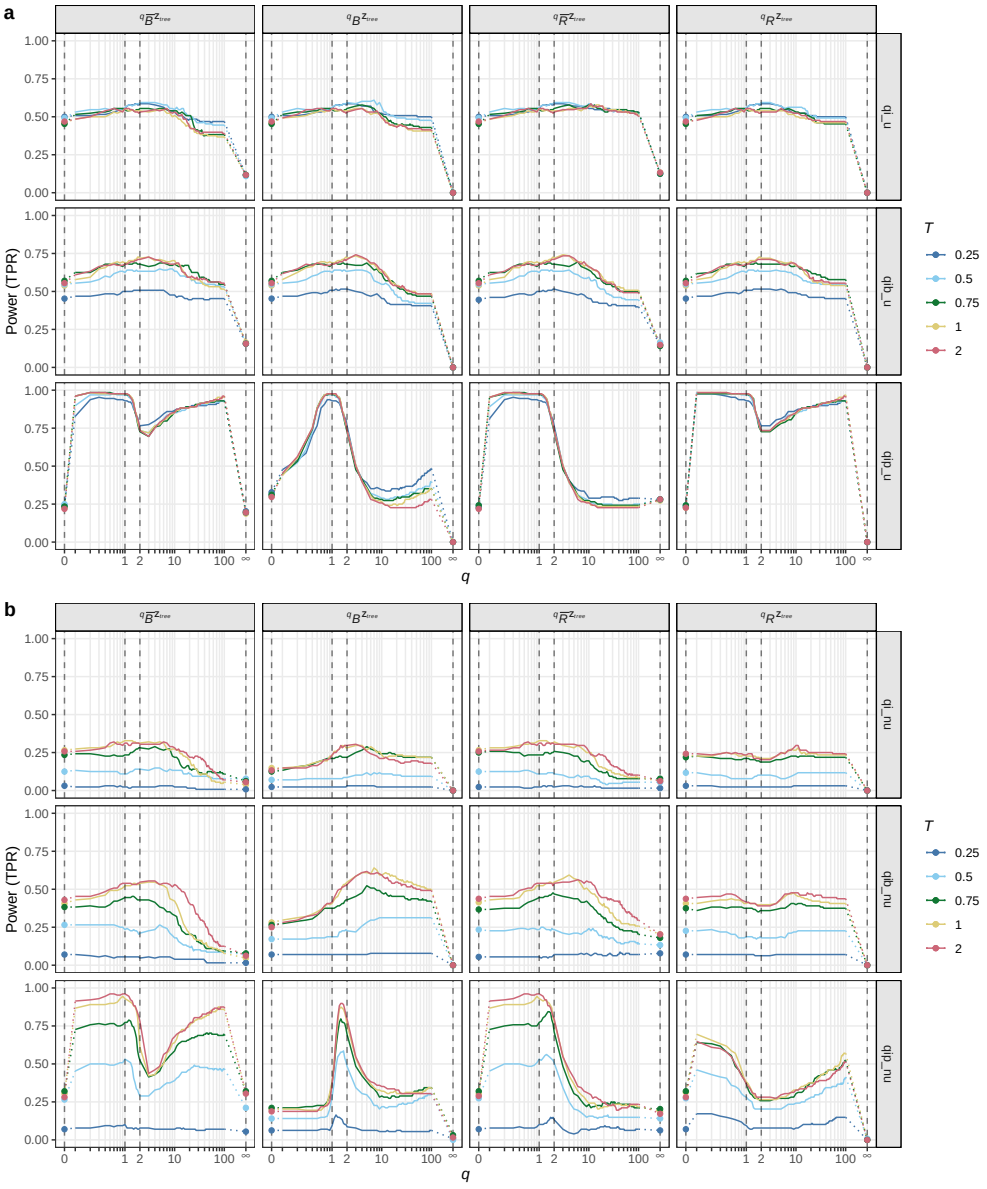


Figure B.10: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 2 (for nested subcommunities): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with independently evolved traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

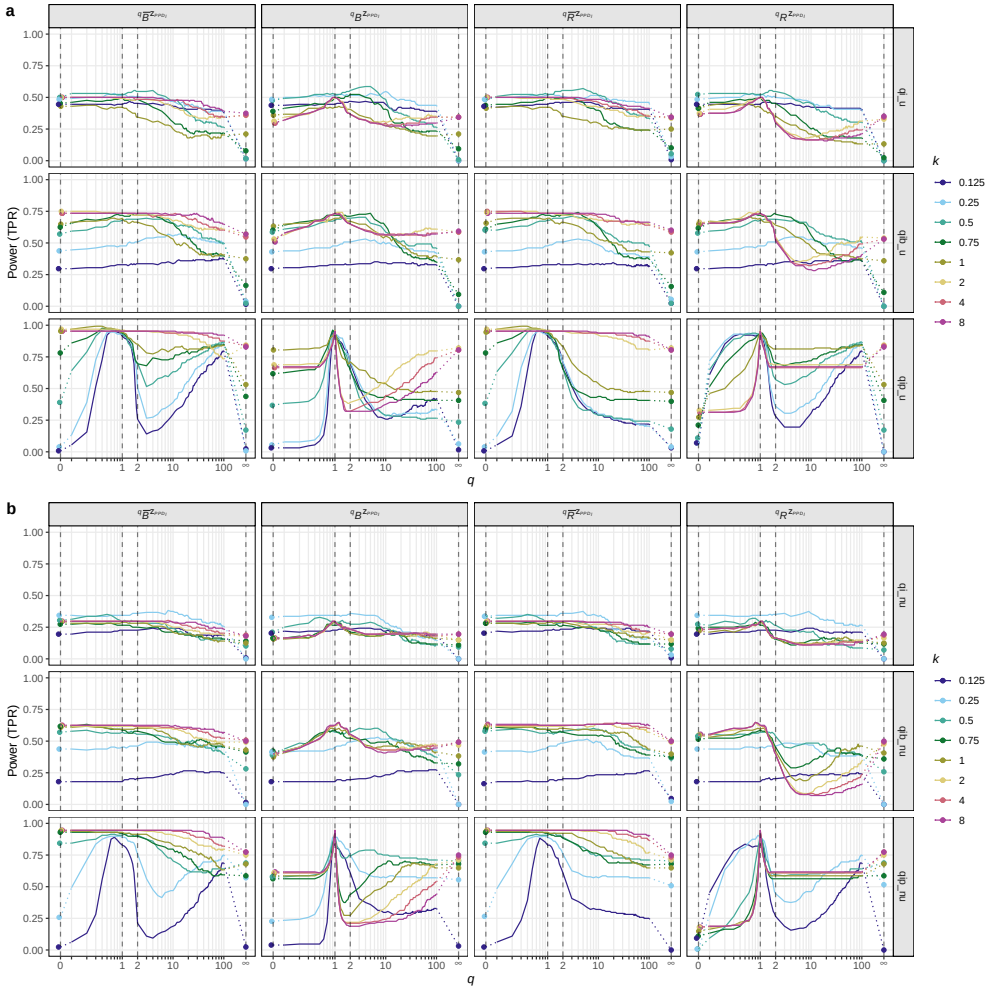


Figure B.11: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 2 (for nested subcommunities): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with independently evolved traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

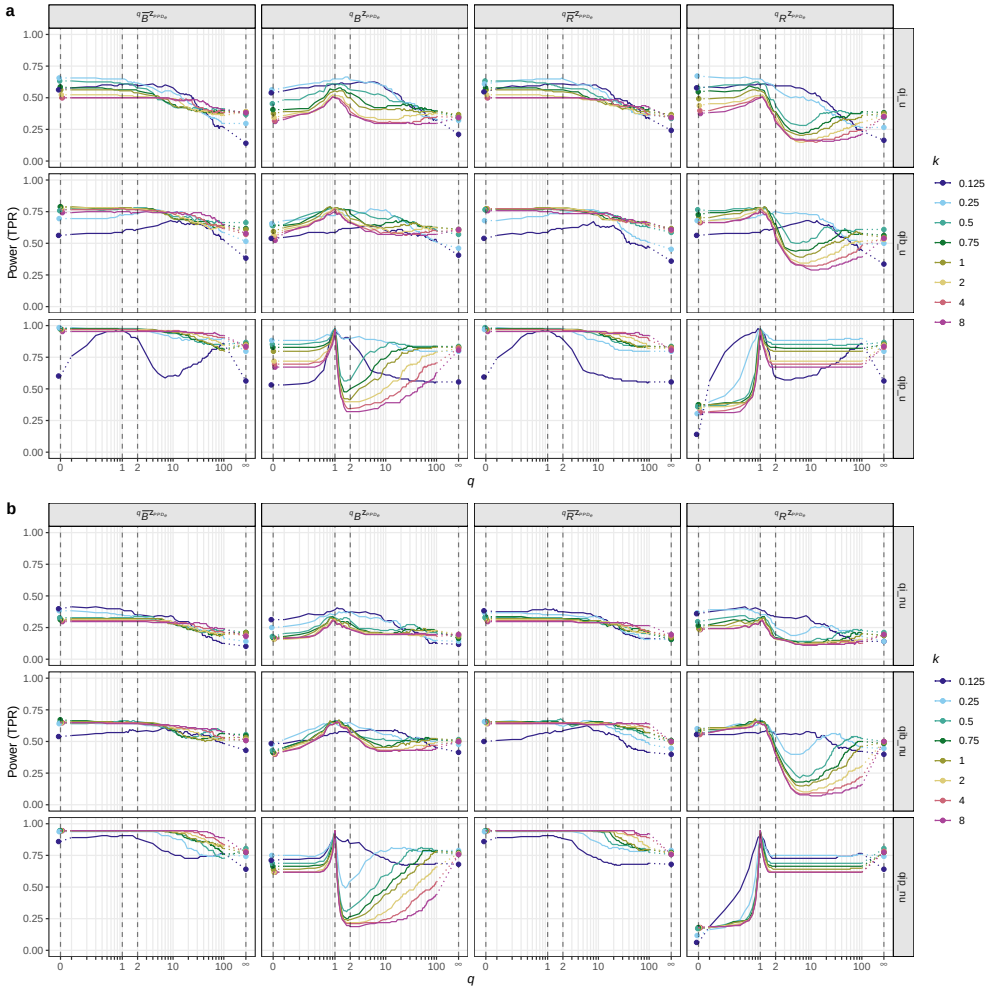


Figure B.12: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 2 (for nested subcommunities): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with independently evolved traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

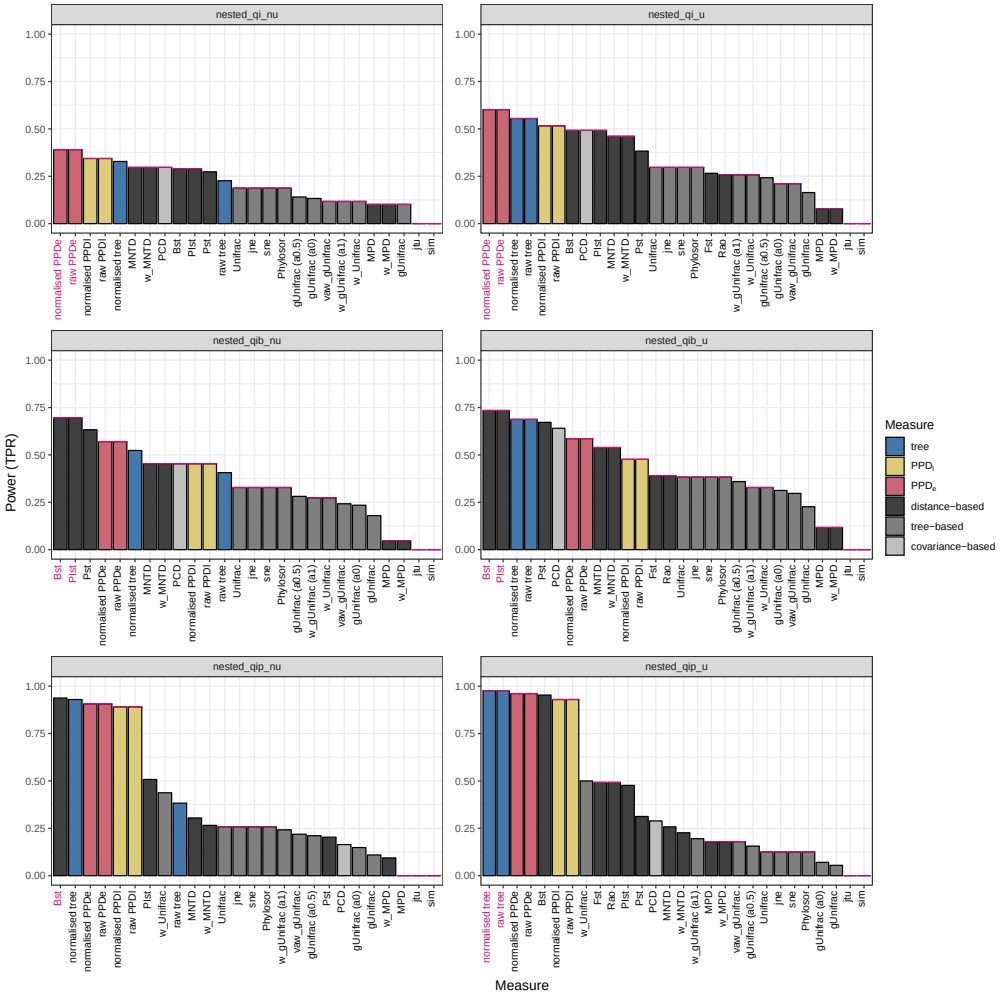


Figure B.13: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 2 (for nested subcommunities): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

B.3.4 Experiment 3: Varying tree size

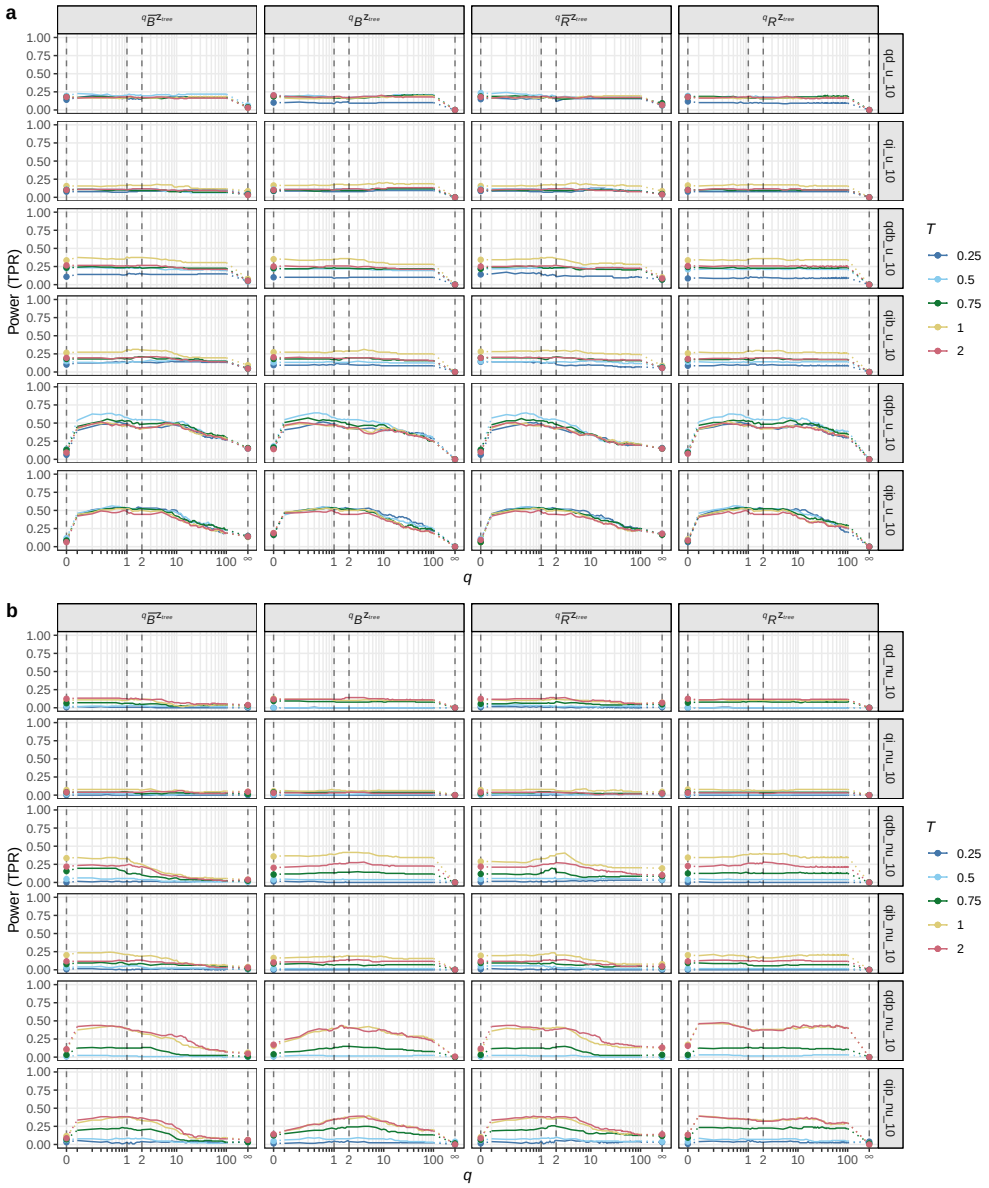


Figure B.14: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 3 (for 10-tip phylogenies): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

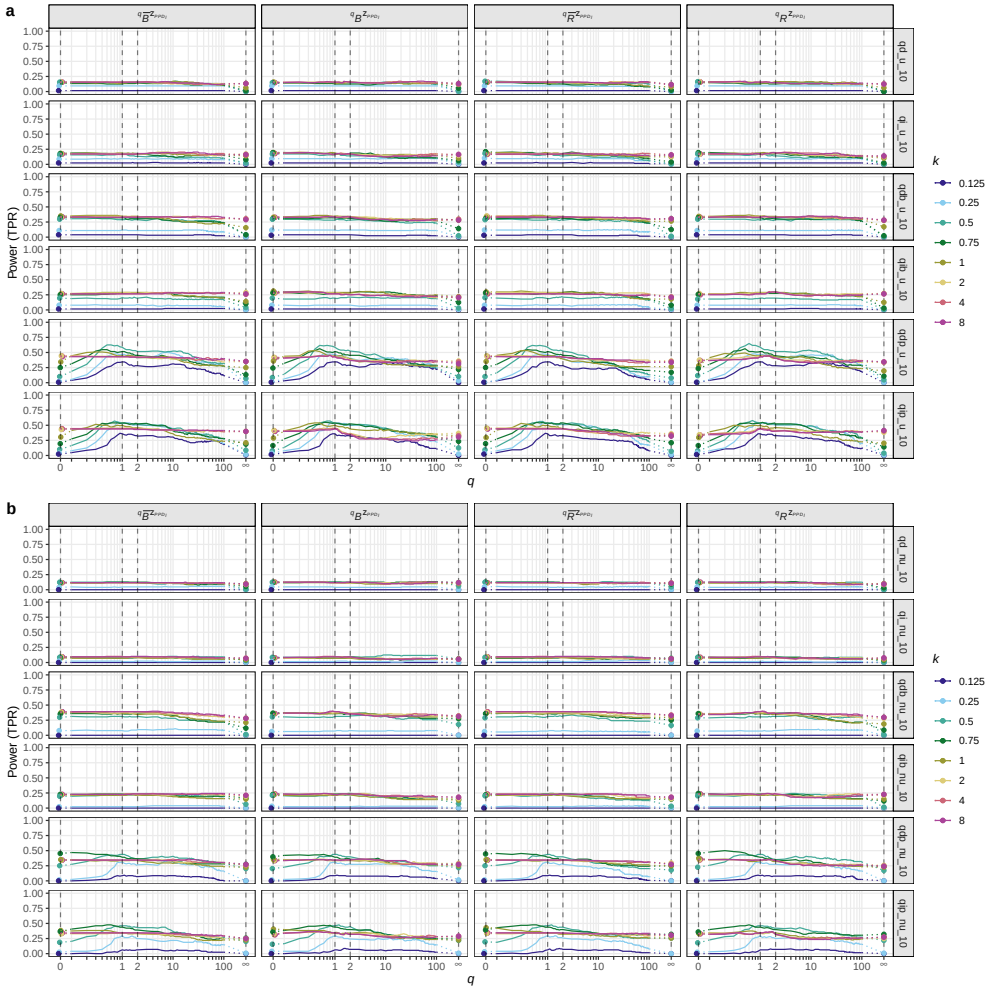


Figure B.15: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 3 (for 10-tip phylogenies): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

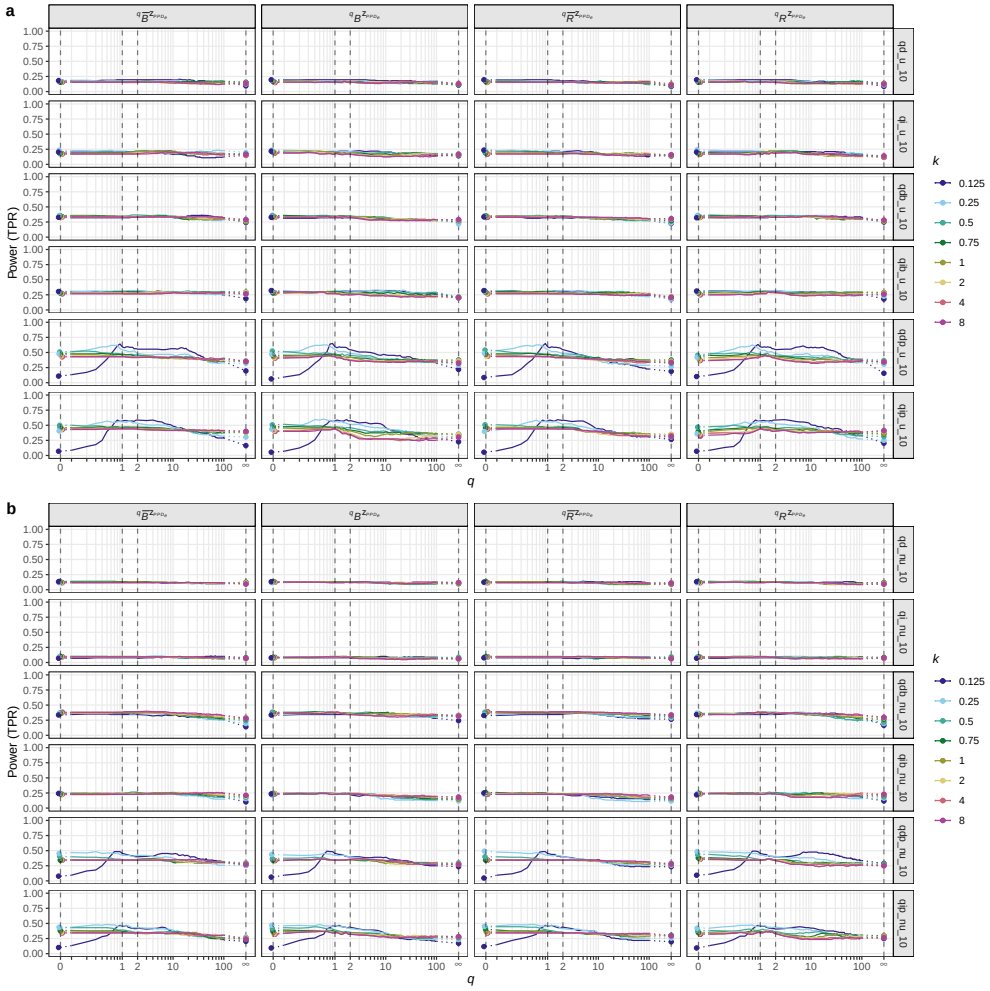


Figure B.16: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 3 (for 10-tip phylogenies): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

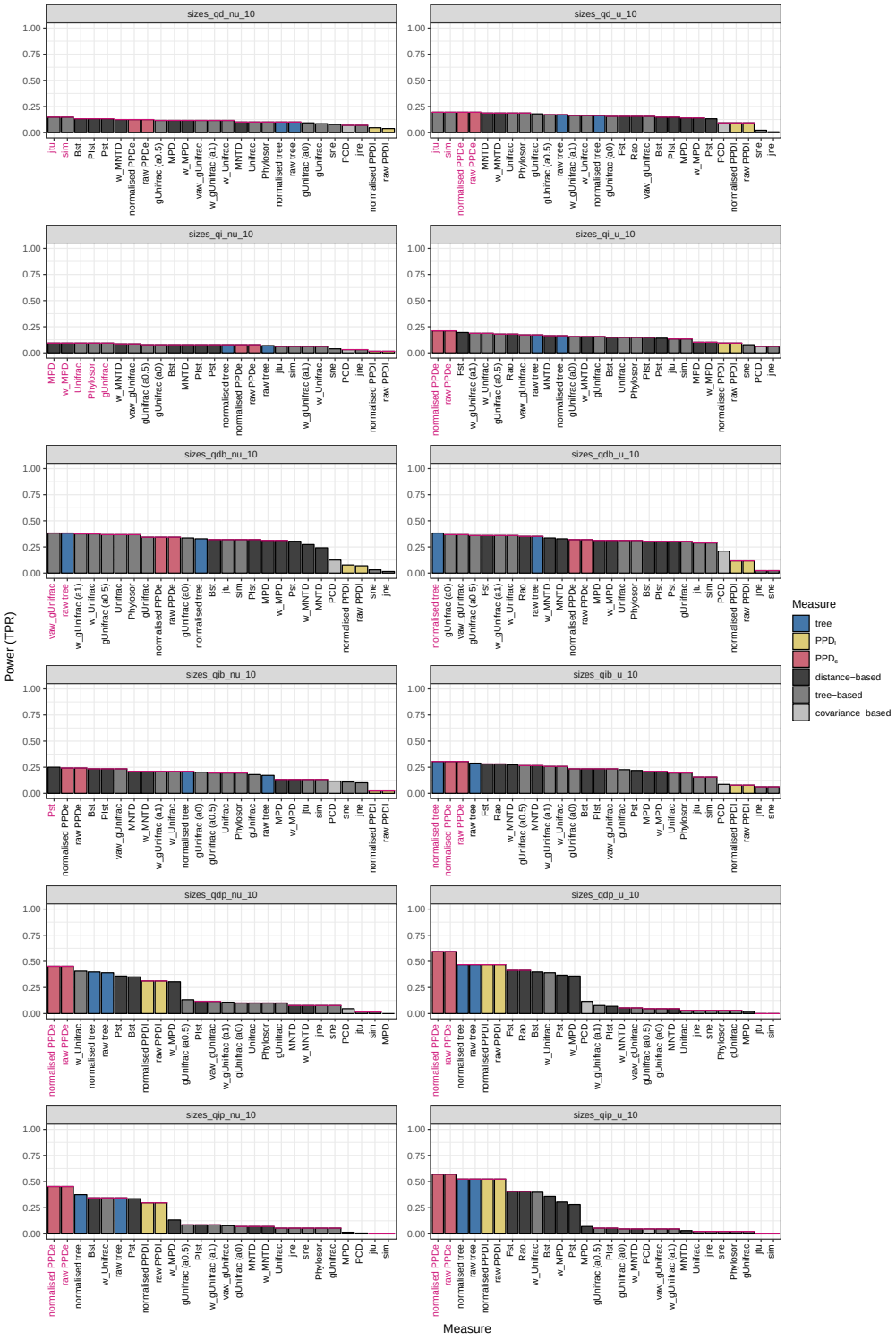


Figure B.17: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 3 (for 10-tip phylogenies): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

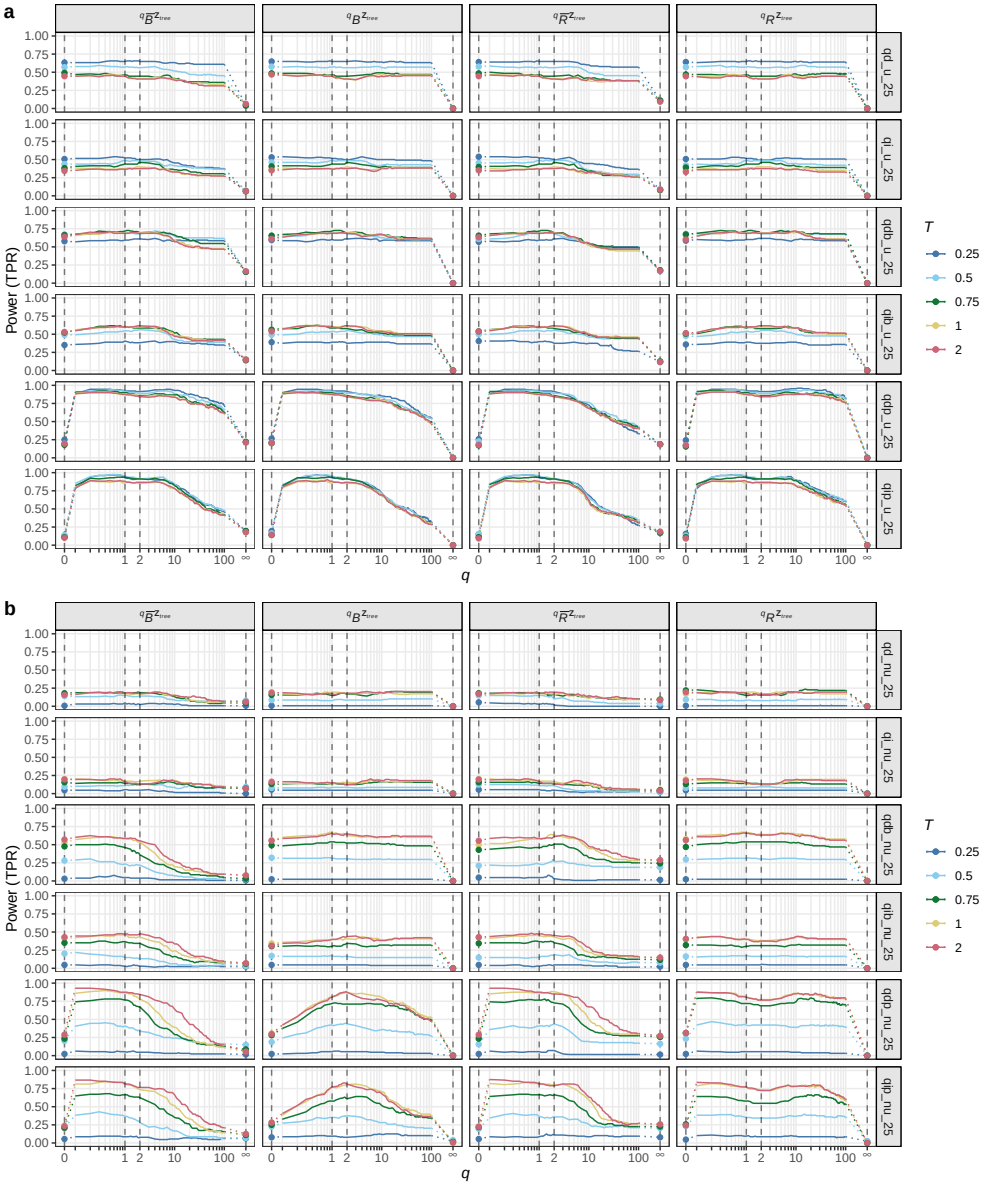


Figure B.18: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 3 (for 25-tip phylogenies): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

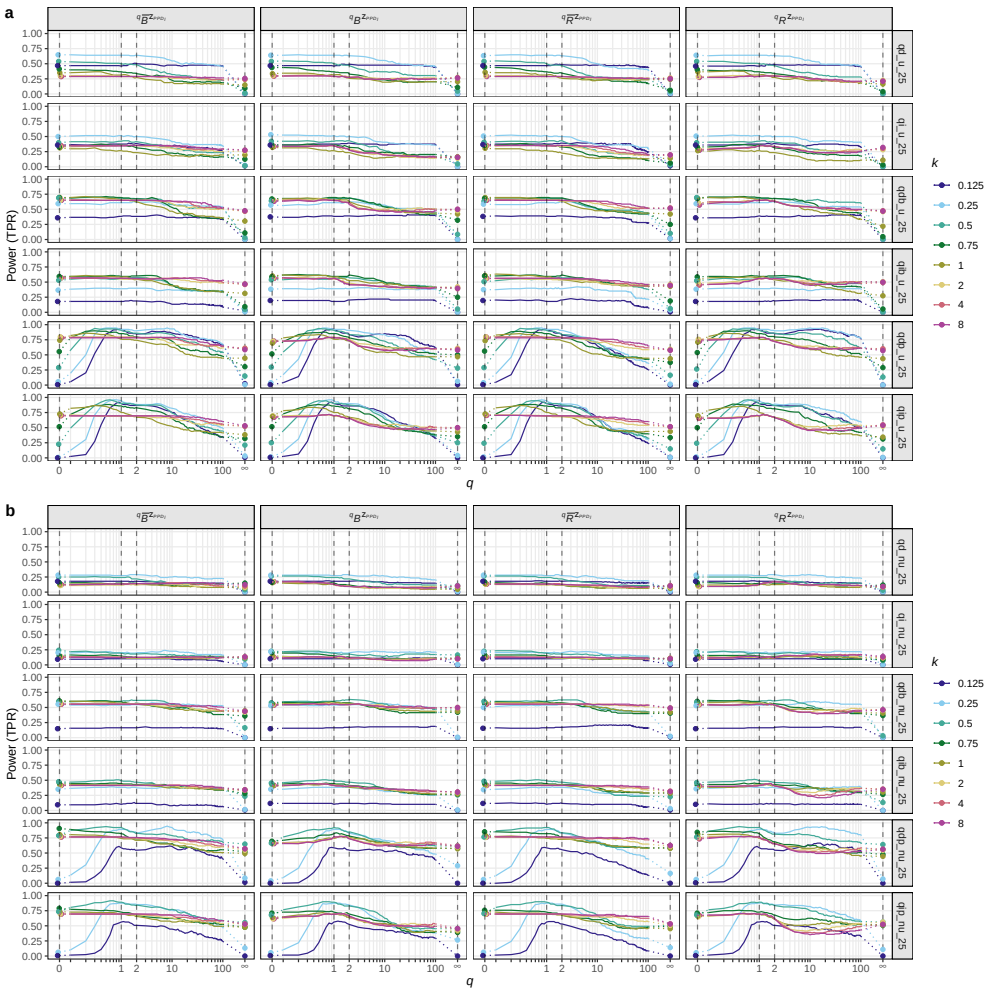


Figure B.19: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 3 (for 25-tip phylogenies): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

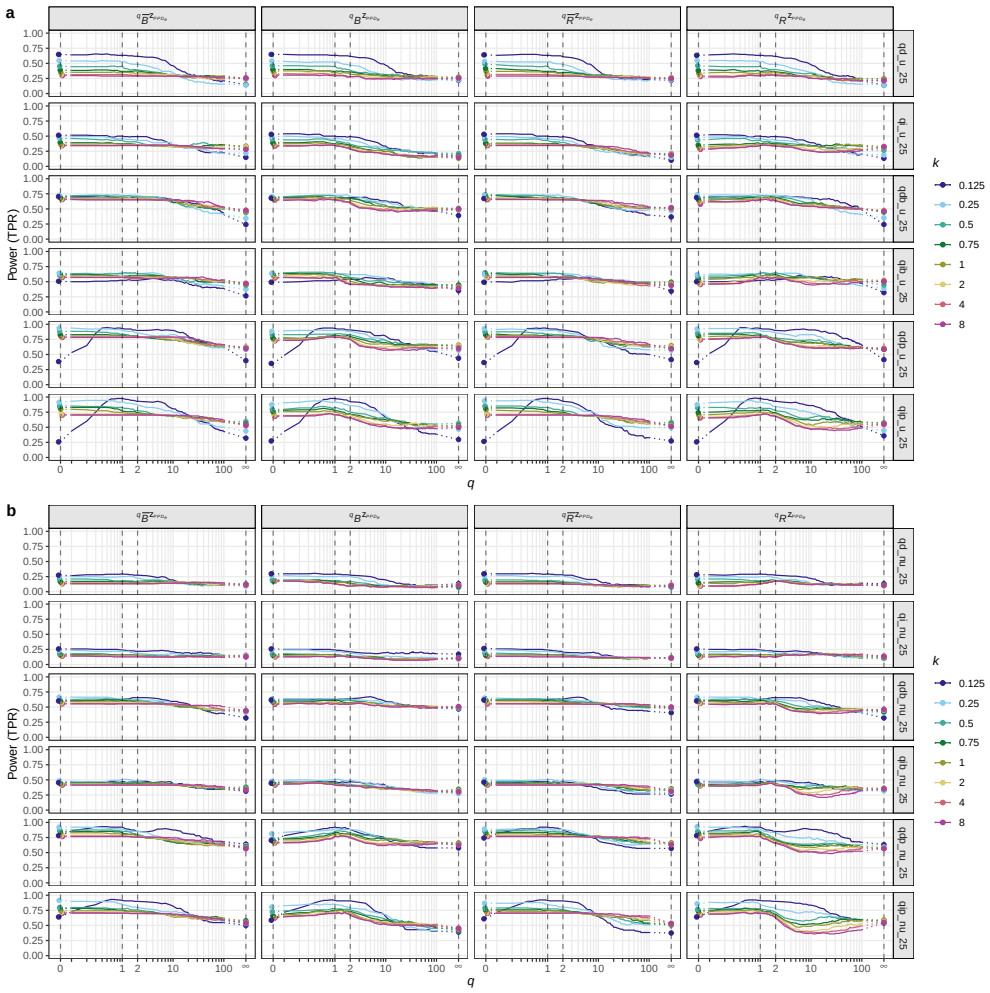


Figure B.20: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 3 (for 25-tip phylogenies): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

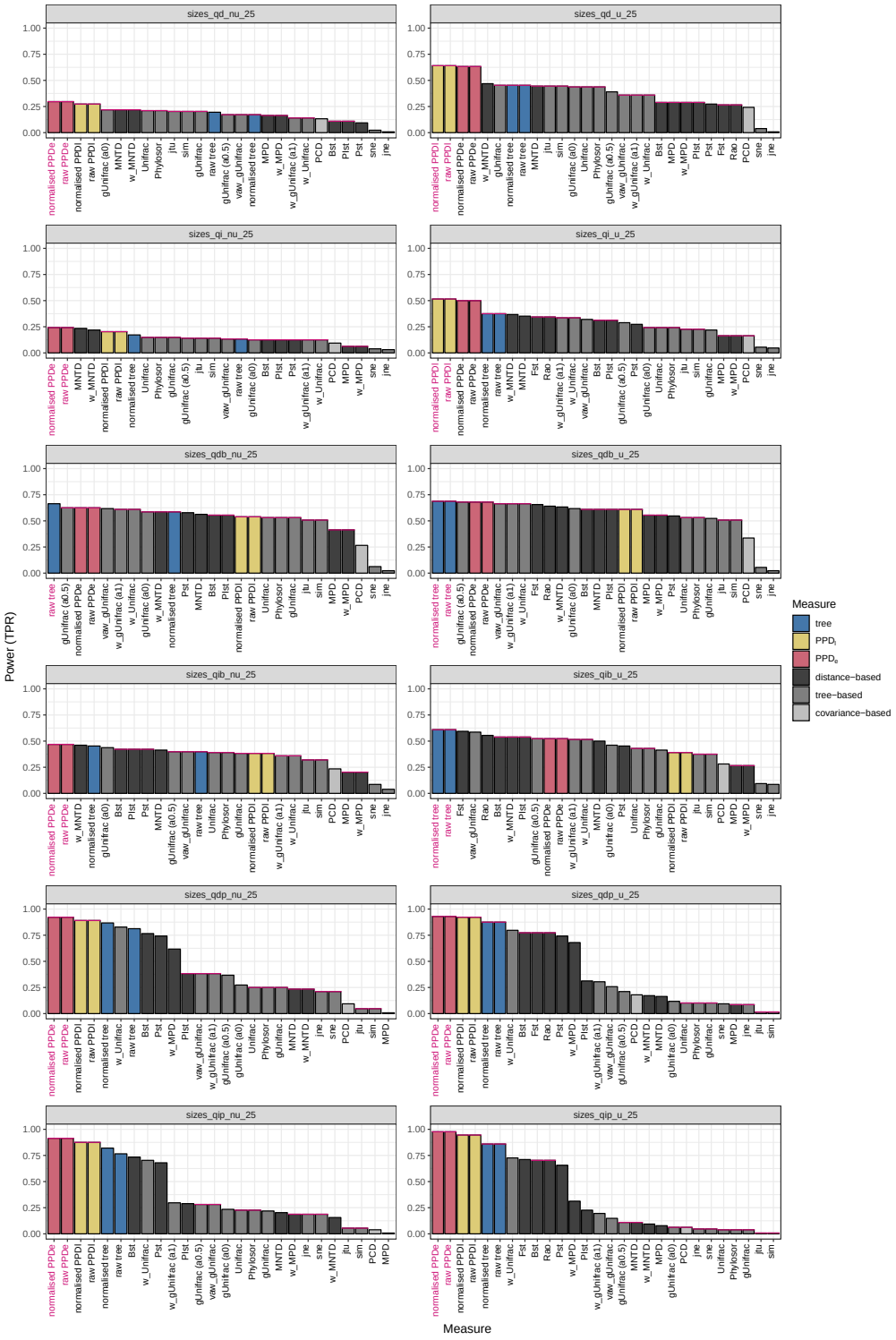


Figure B.21: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 3 (for 25-tip phylogenies): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

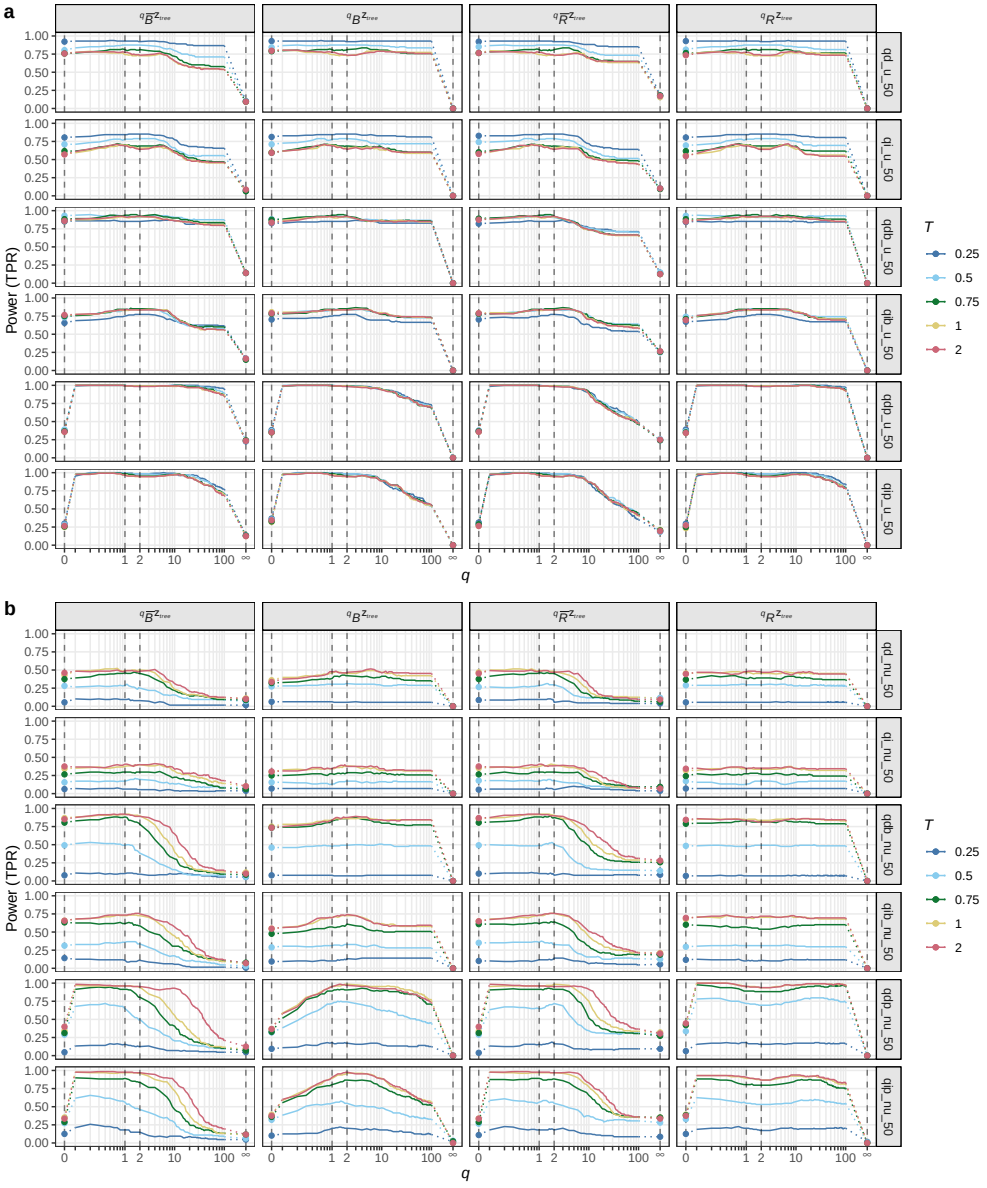


Figure B.22: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 3 (for 50-tip phylogenies): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

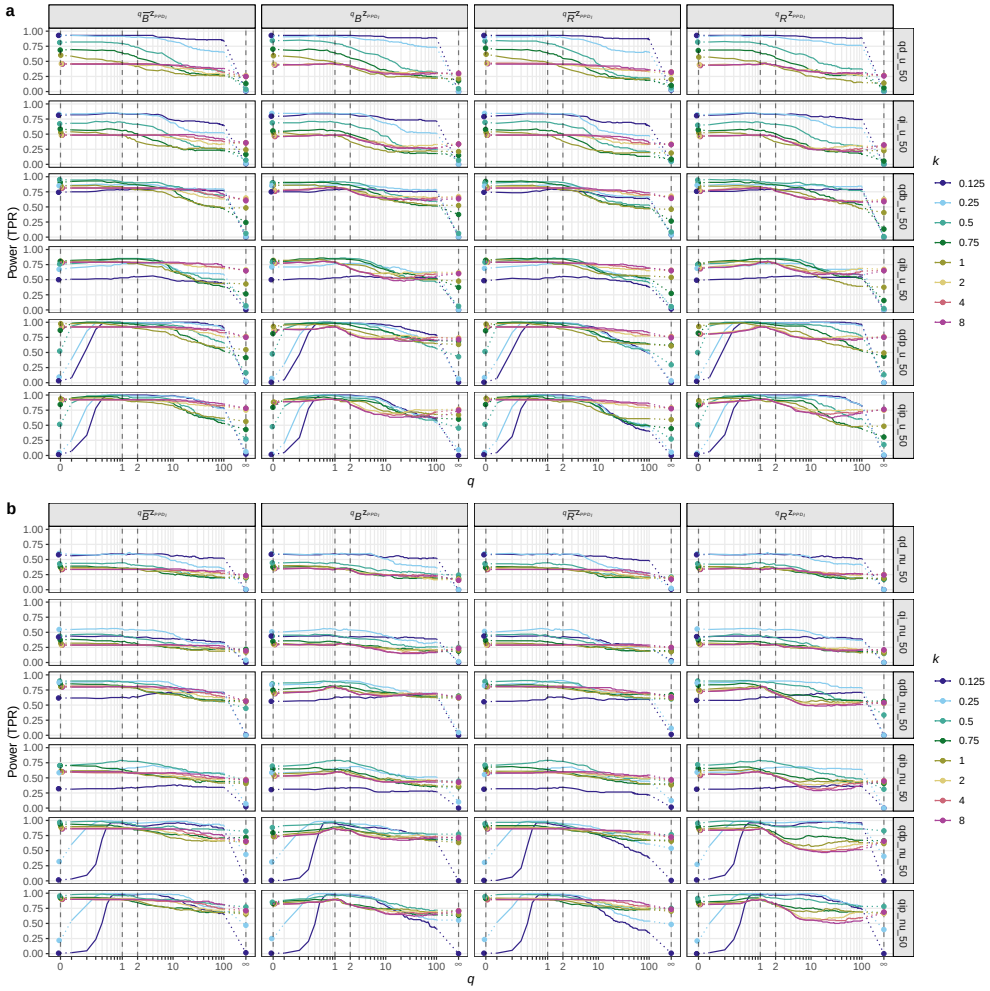


Figure B.23: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 3 (for 50-tip phylogenies): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

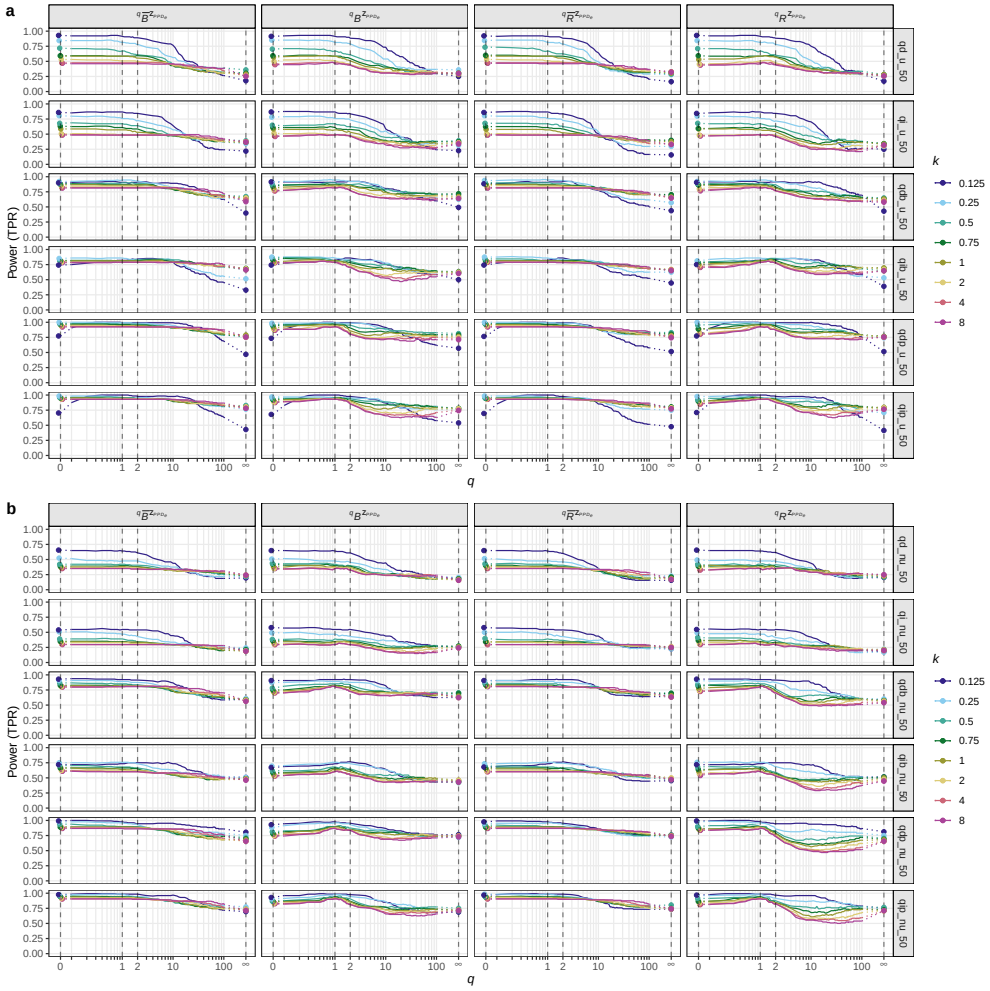


Figure B.24: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 3 (for 50-tip phylogenies): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

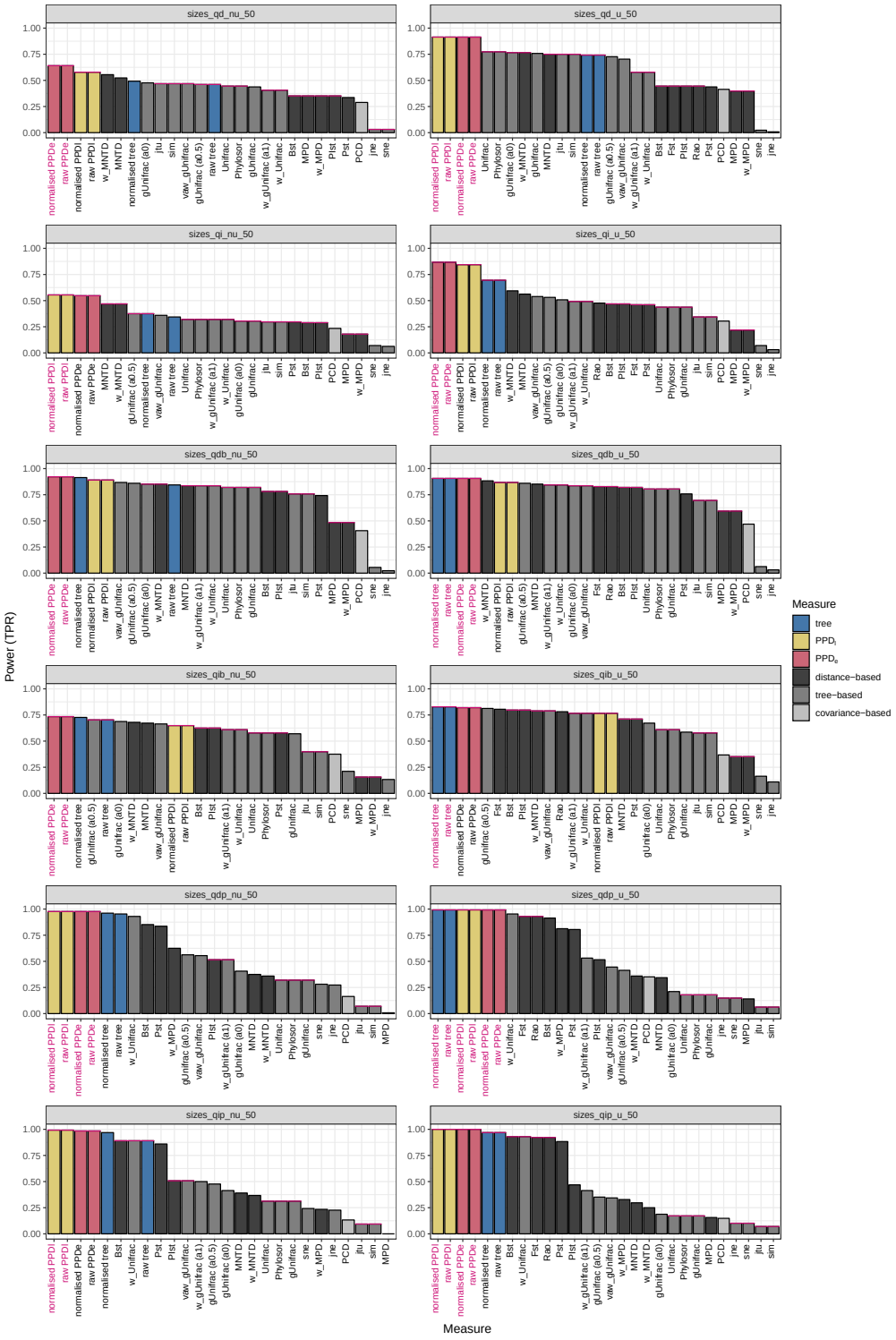


Figure B.25: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 3 (for 50-tip phylogenies): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

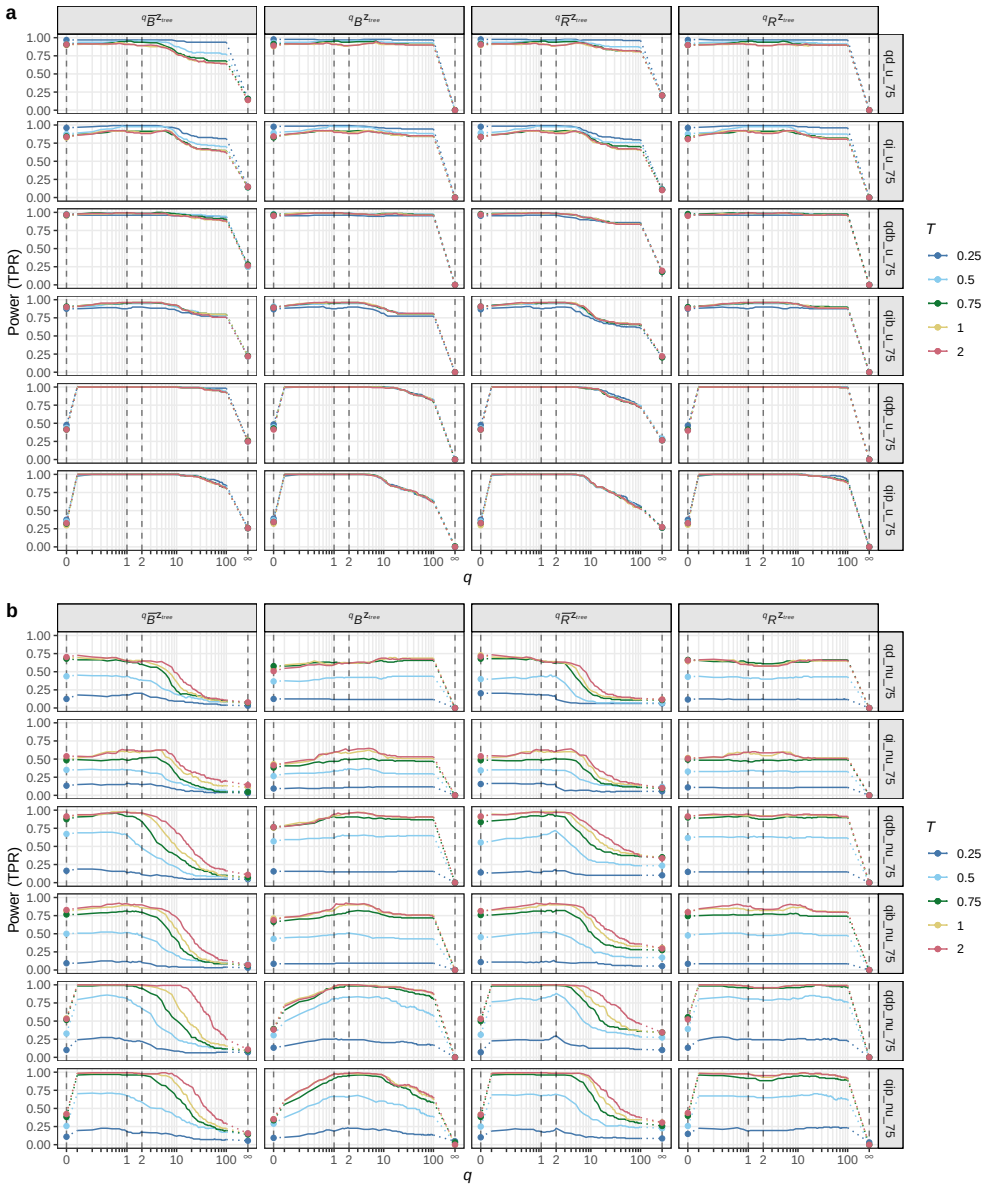


Figure B.26: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 3 (for 75-tip phylogenies): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

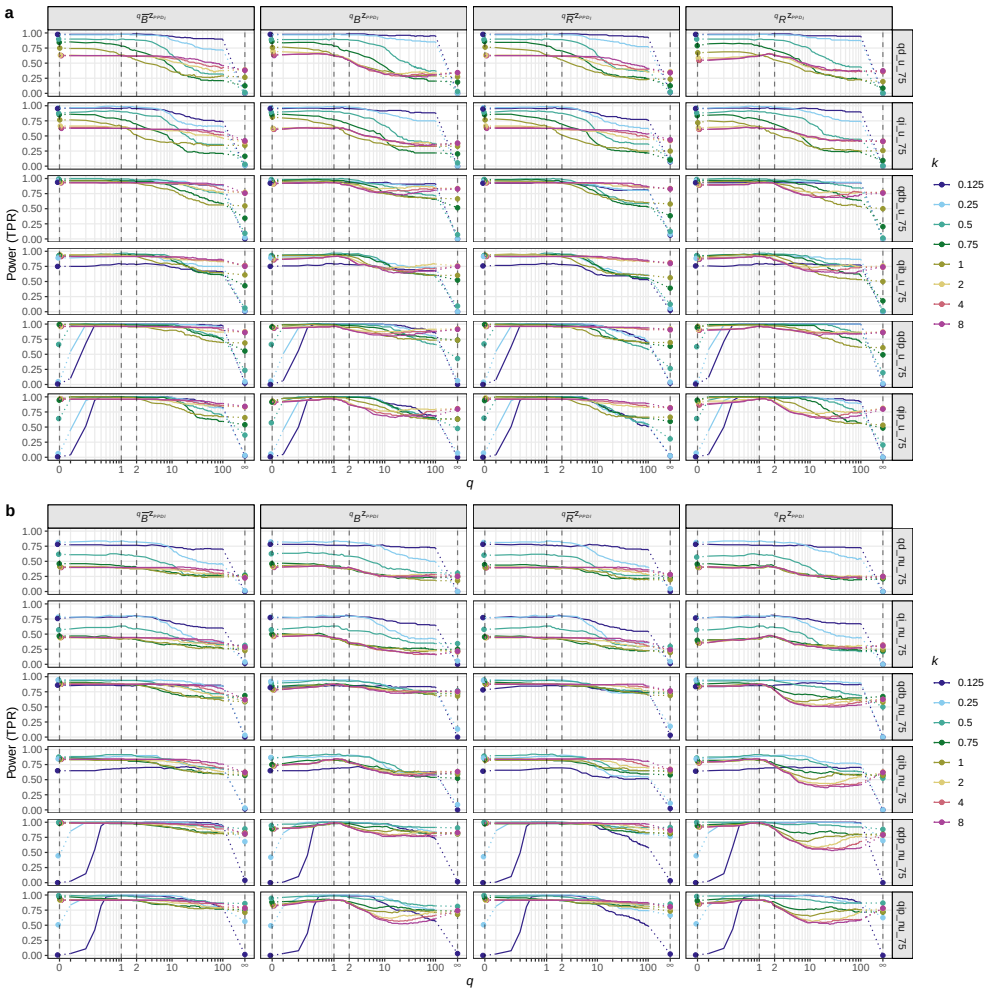


Figure B.27: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 3 (for 75-tip phylogenies): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

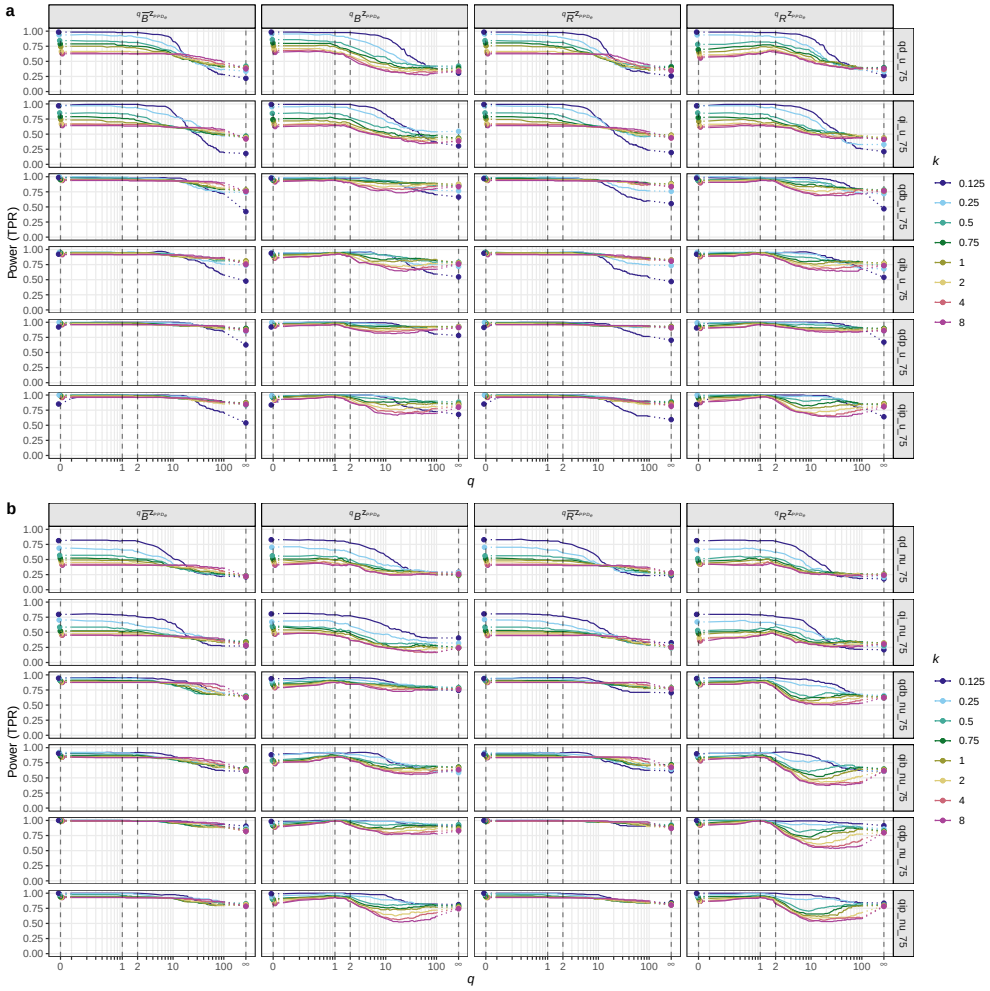


Figure B.28: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 3 (for 75-tip phylogenies): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

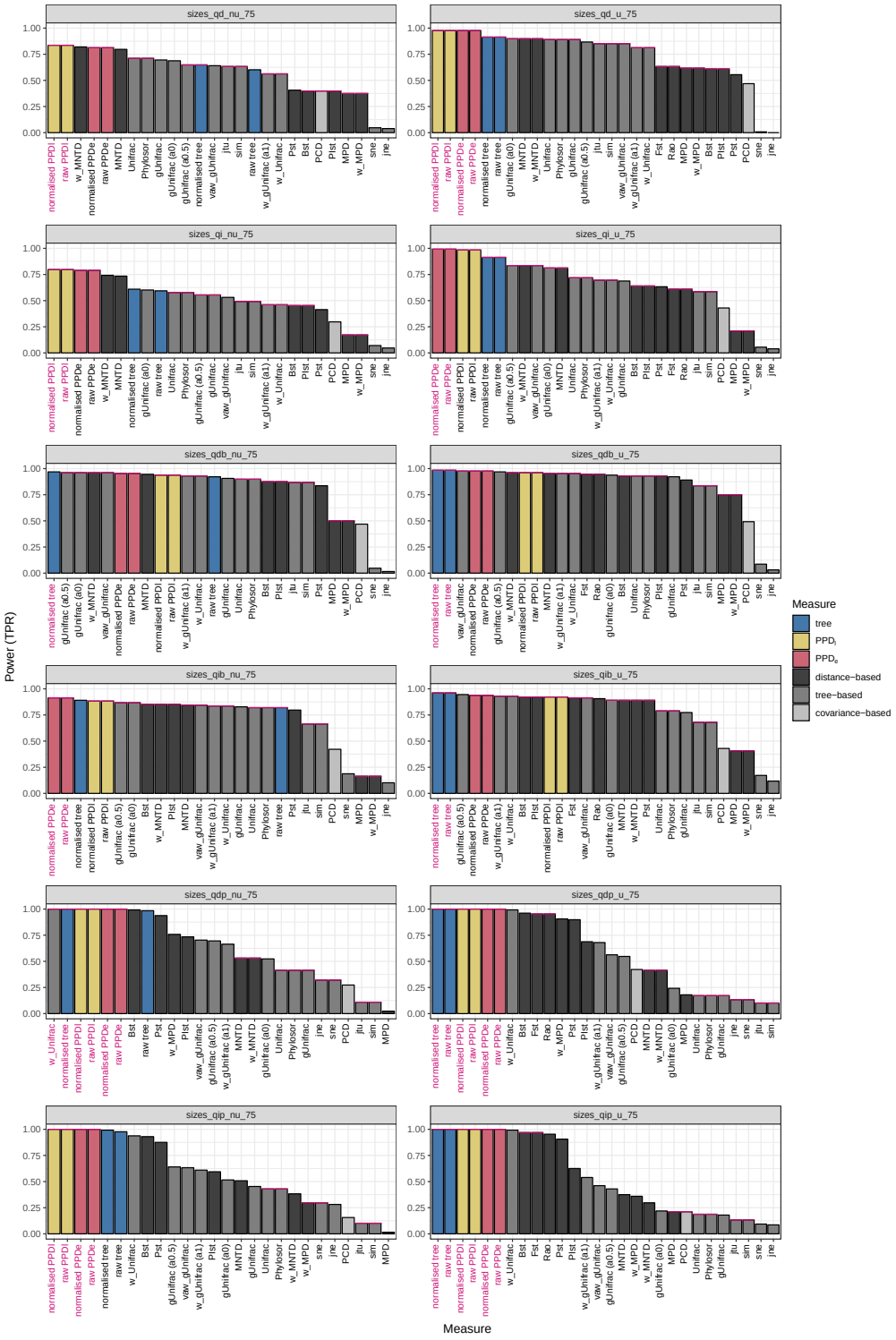


Figure B.29: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 3 (for 75-tip phylogenies): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

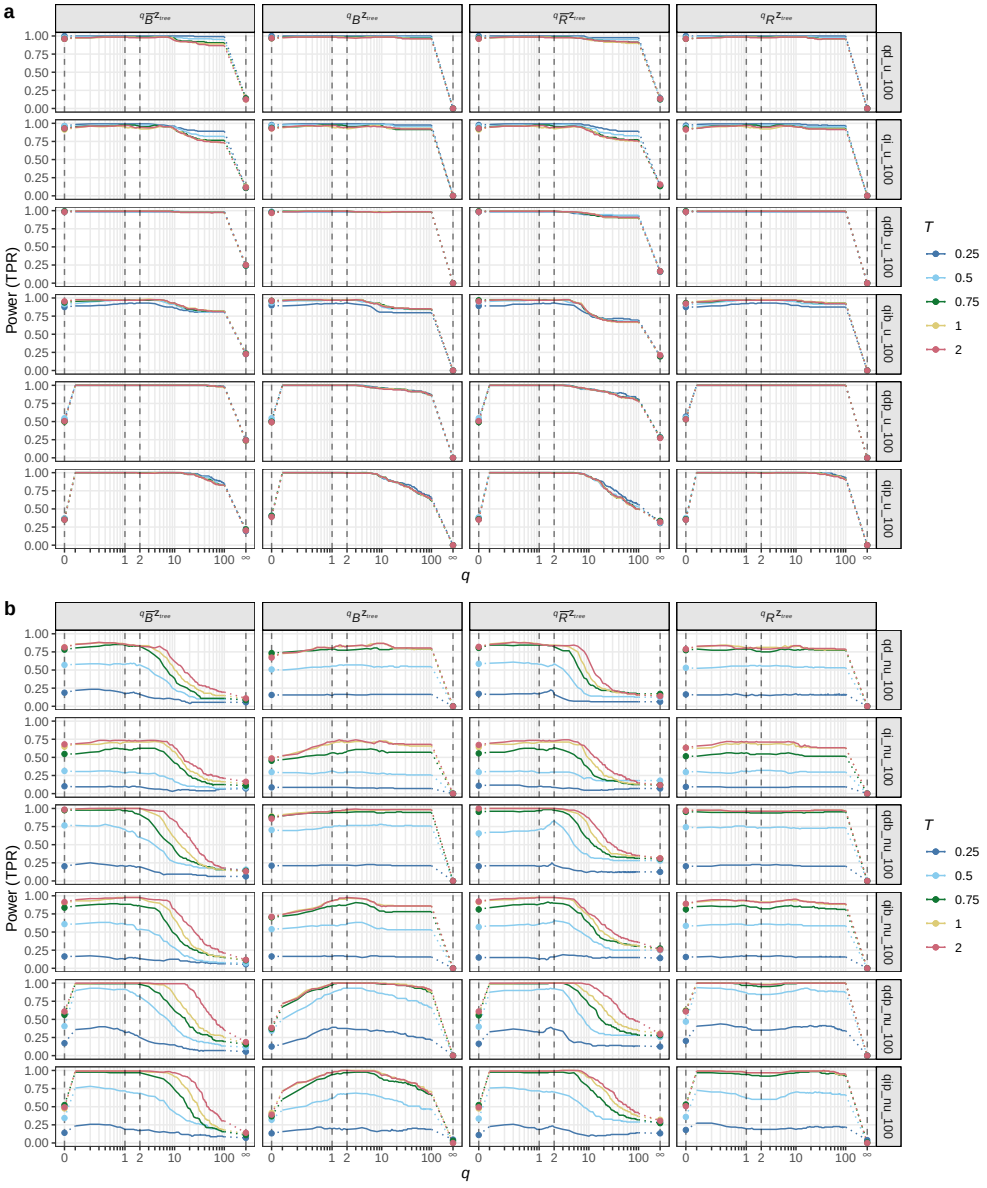


Figure B.30: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 3 (for 100-tip phylogenies): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

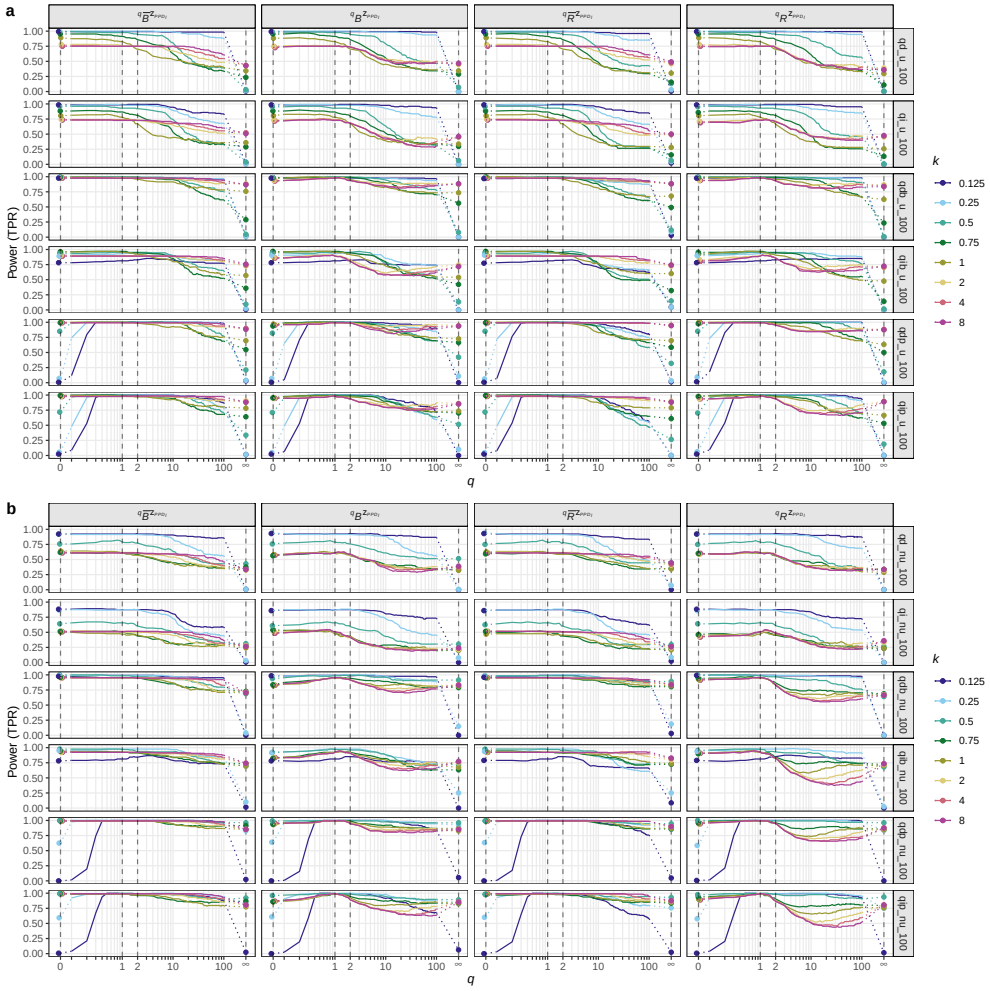


Figure B.31: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 3 (for 100-tip phylogenies): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

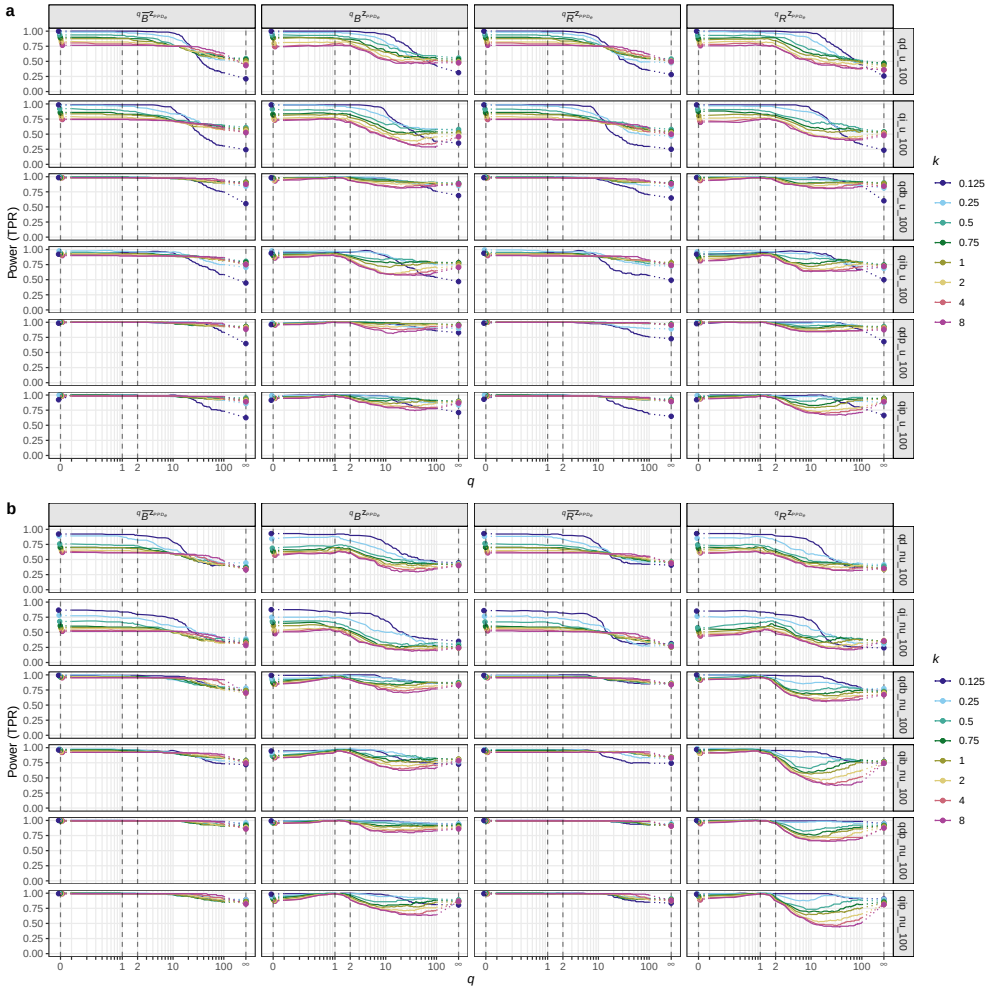


Figure B.32: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 3 (for 100-tip phylogenies): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

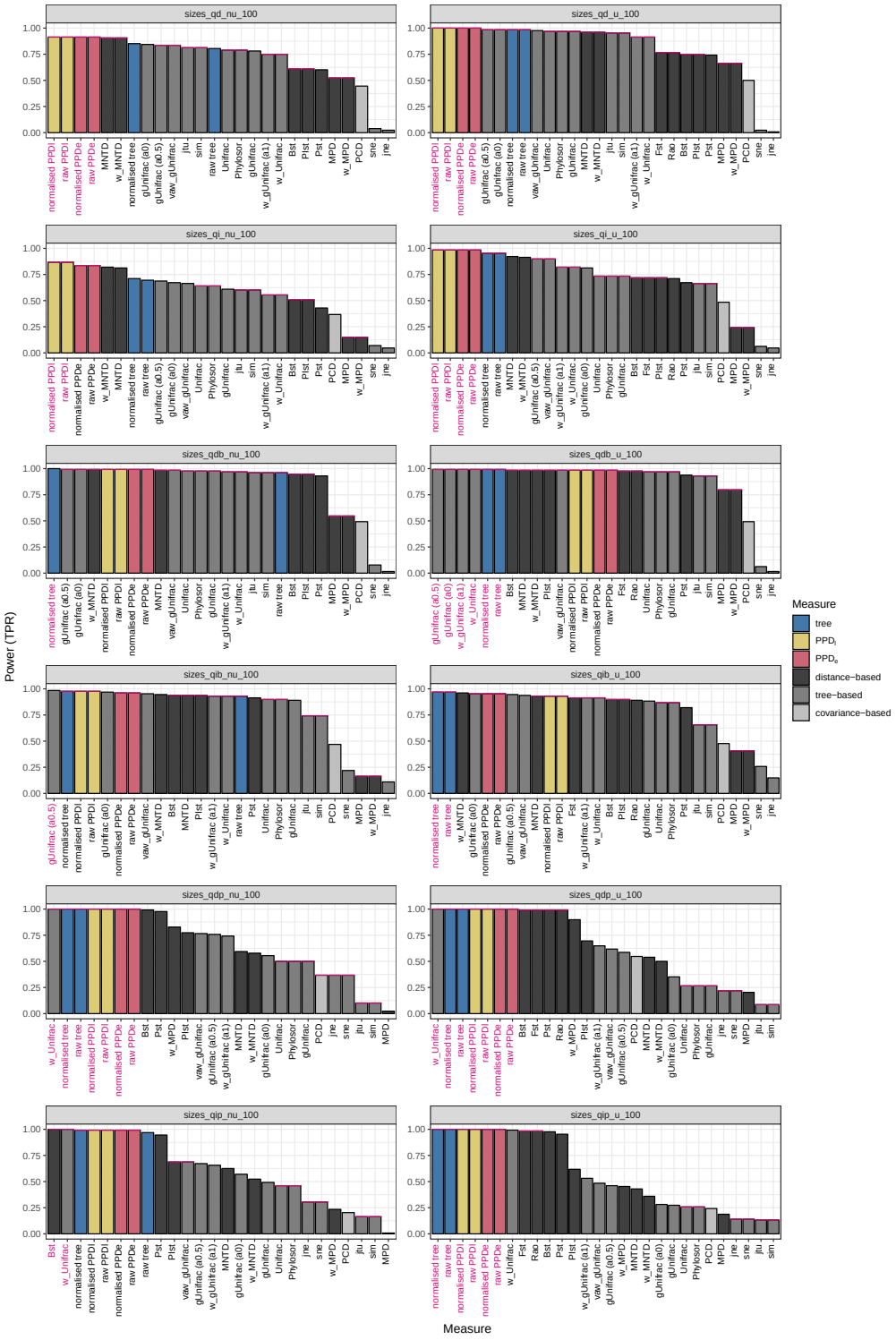


Figure B.33: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 3 (for 100-tip phylogenies): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

B.3.5 Experiment 4: Varying evolutionary rates

B.3.5.1 Quantitative partitioning strategies

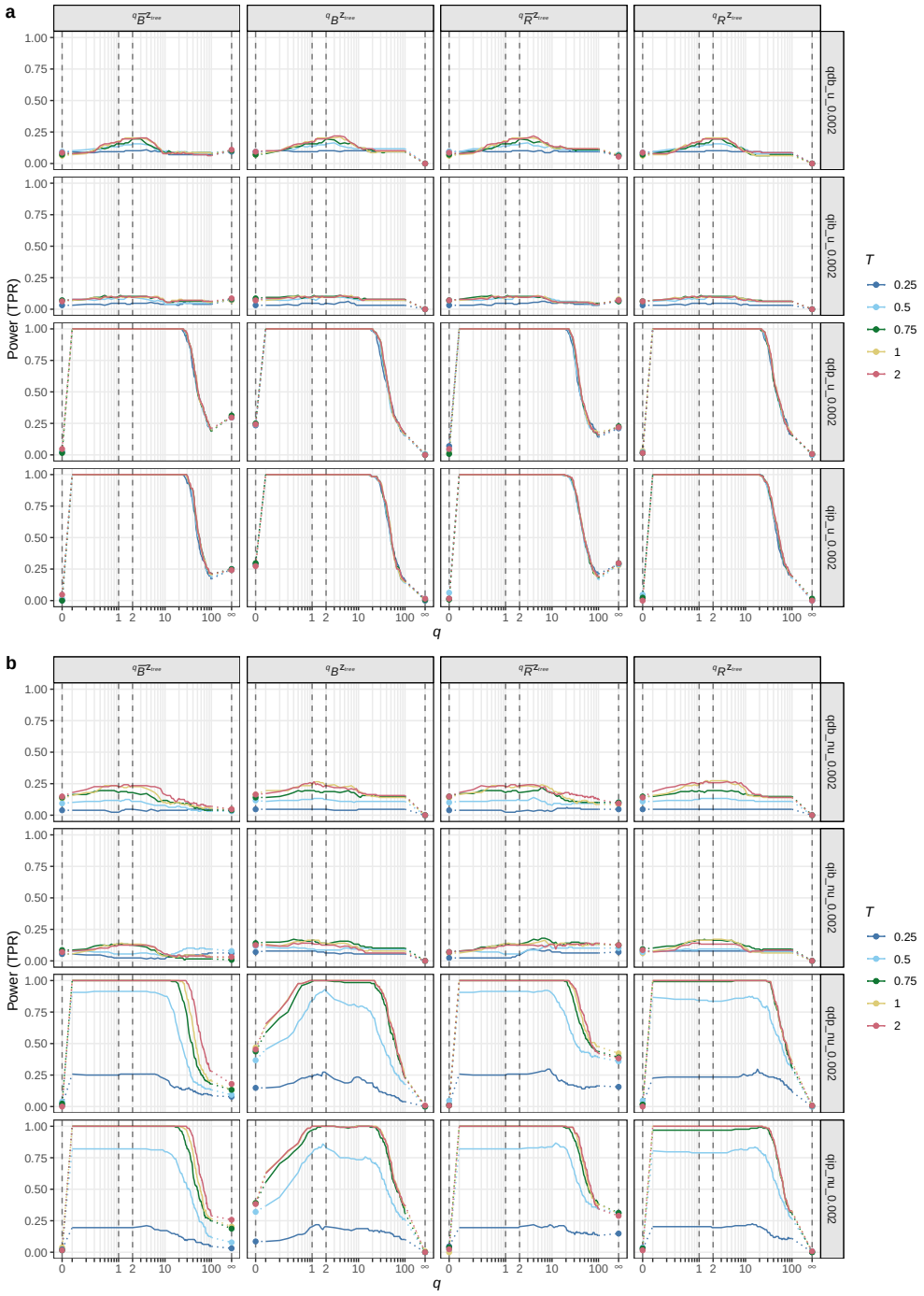


Figure B.34: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 4 (where the quantitative rate is 0.002): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

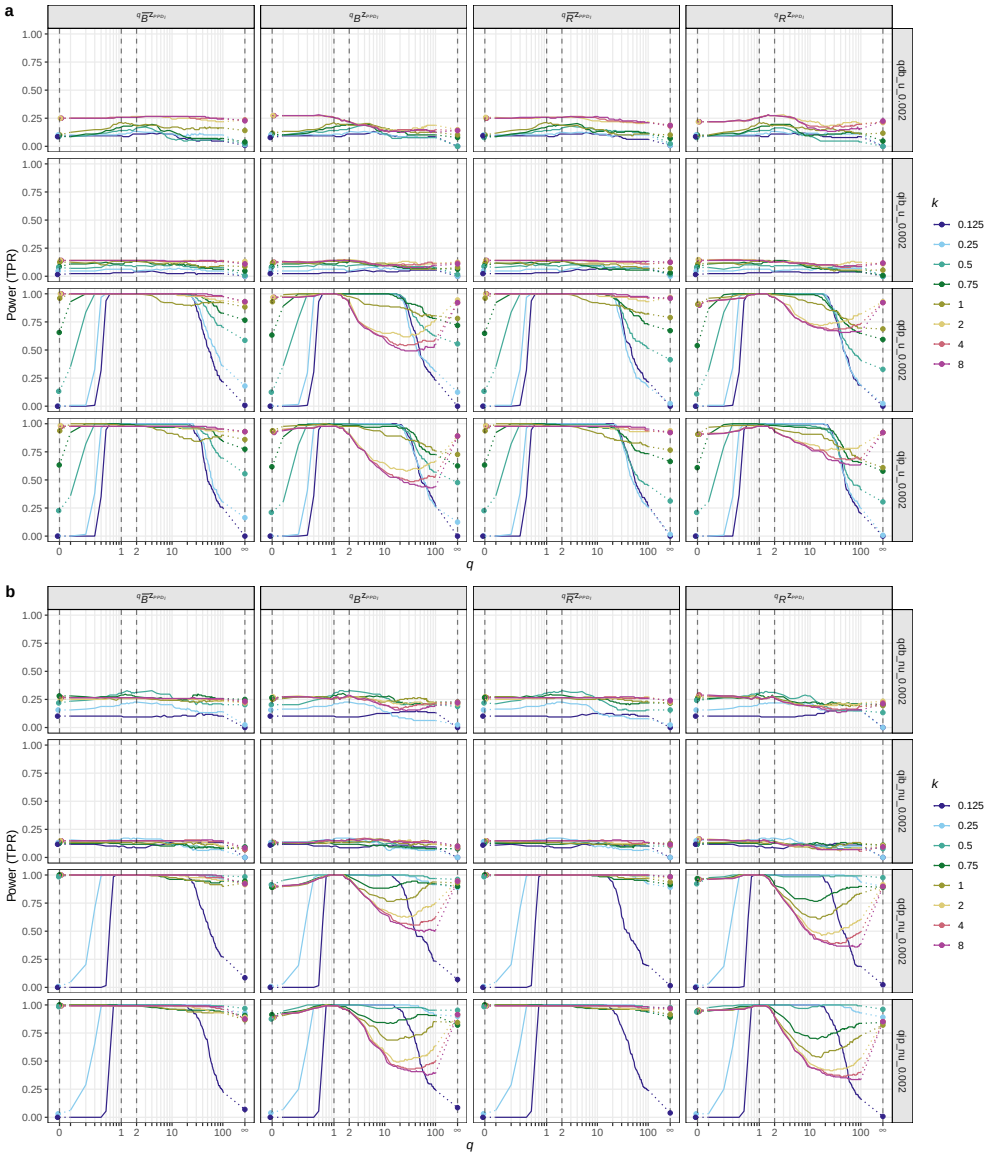


Figure B.35: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 4 (where the quantitative rate is 0.002): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

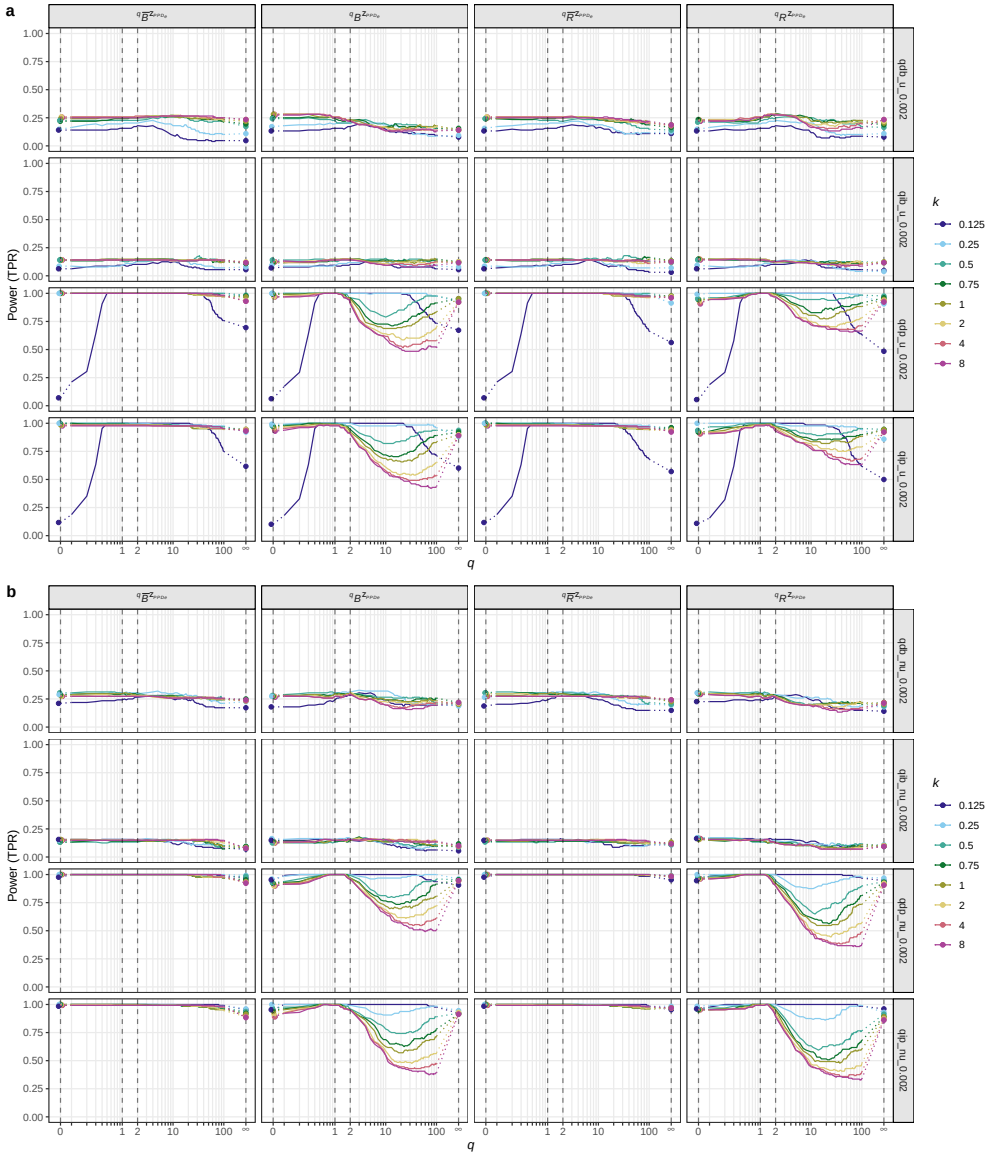


Figure B.36: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 4 (where the quantitative rate is 0.002): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

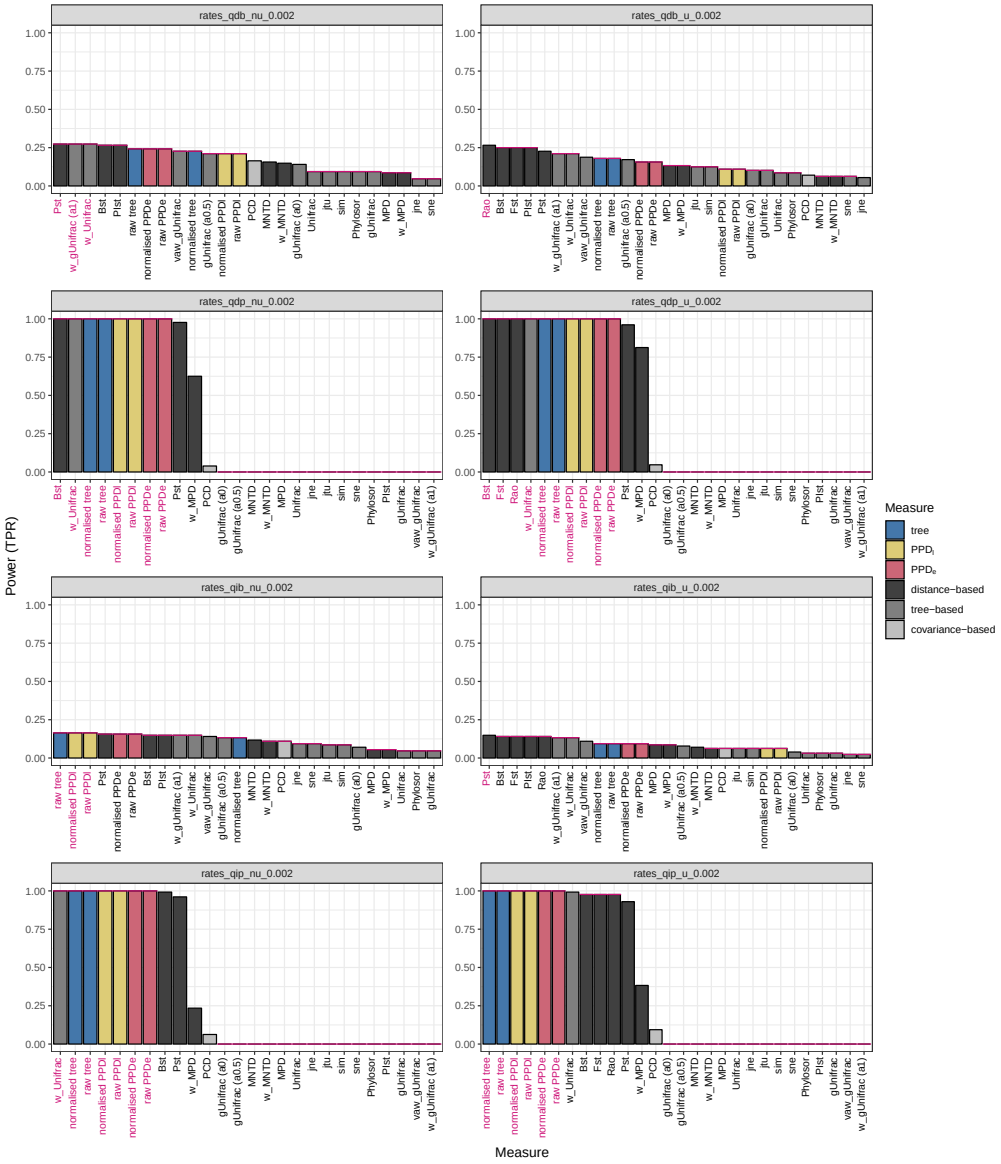


Figure B.37: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 4 (where the quantitative rate is 0.002): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

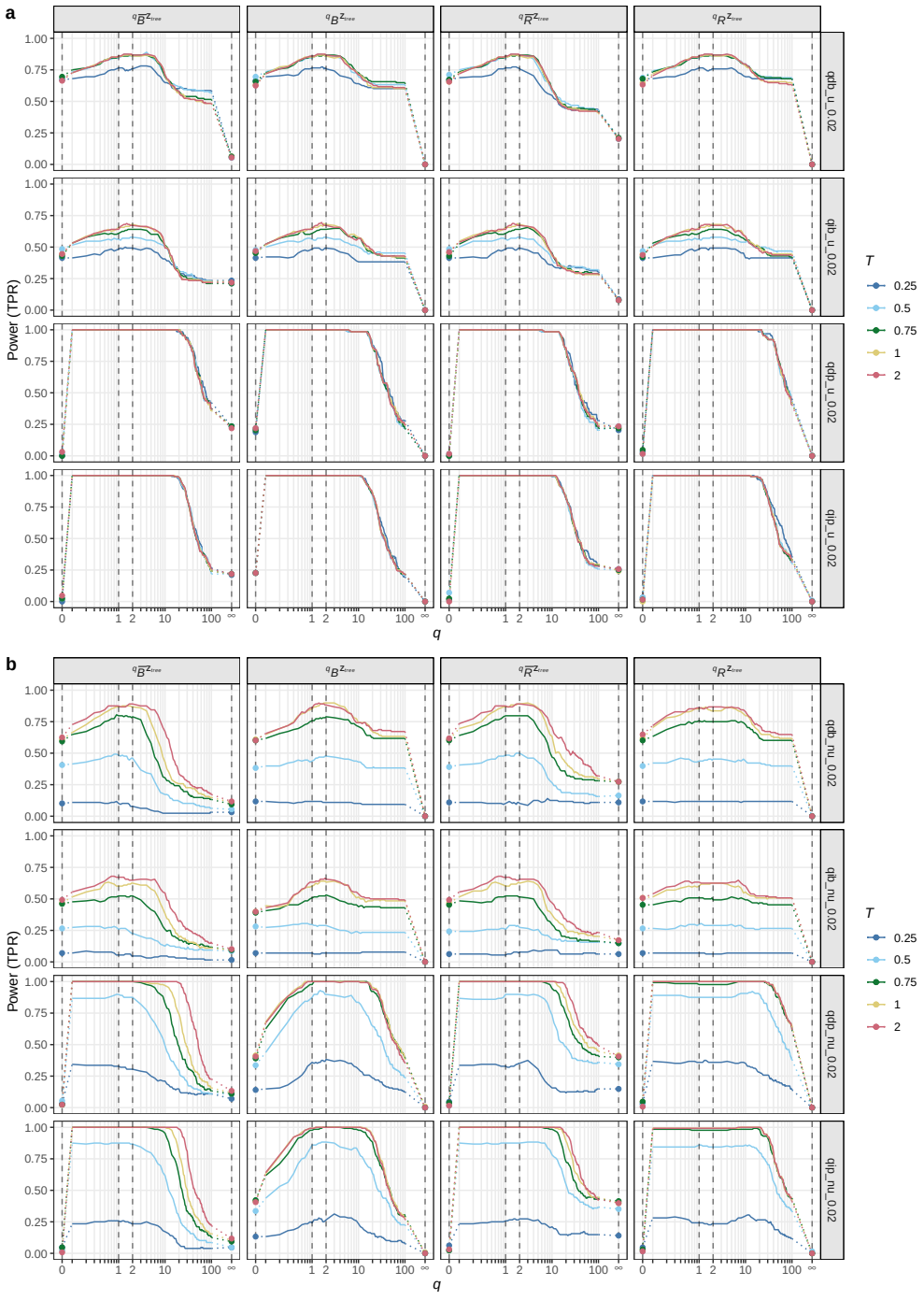


Figure B.38: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 4 (where the quantitative rate is 0.02): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

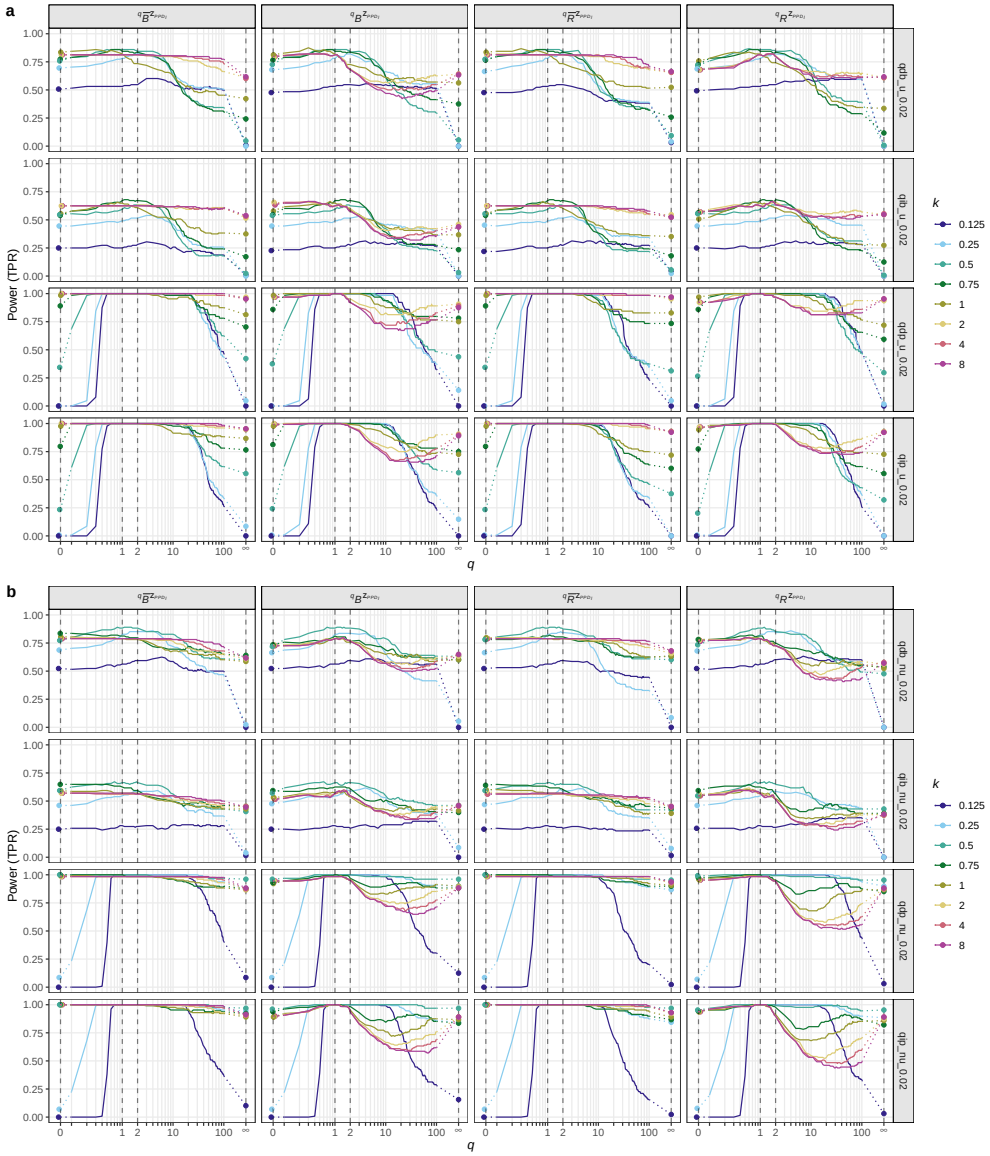


Figure B.39: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 4 (where the quantitative rate is 0.02): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

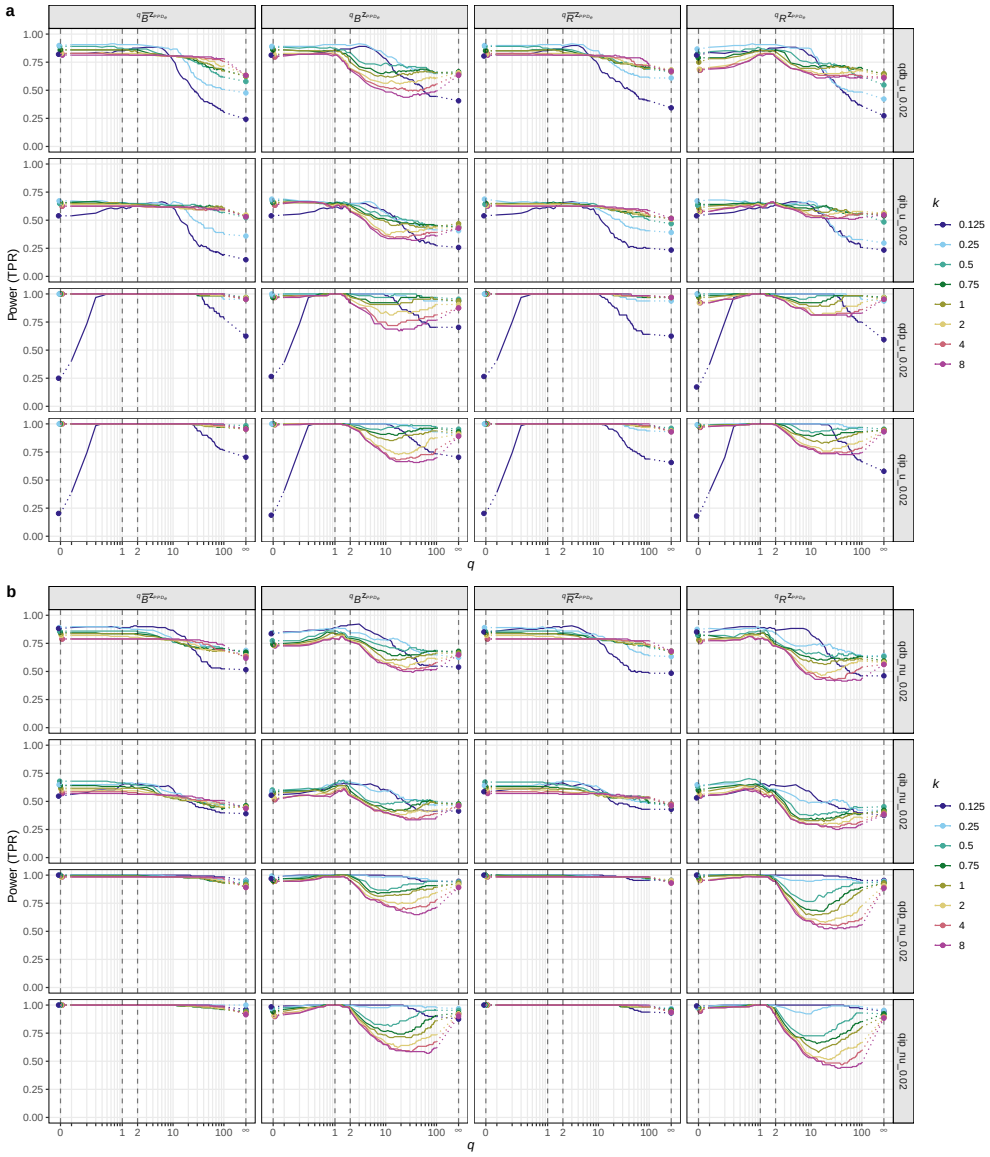


Figure B.40: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 4 (where the quantitative rate is 0.02): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

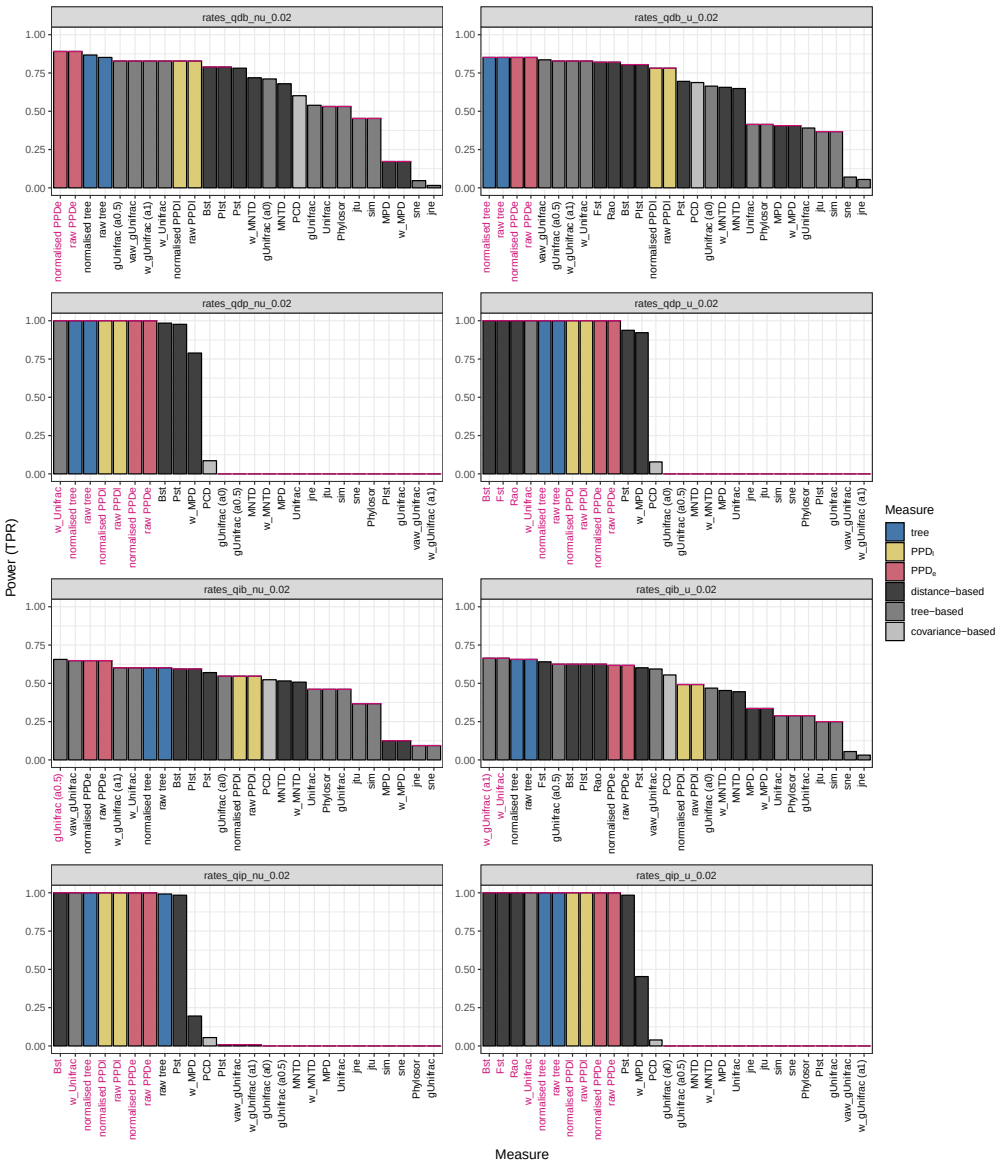


Figure B.41: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 4 (where the quantitative rate is 0.02): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

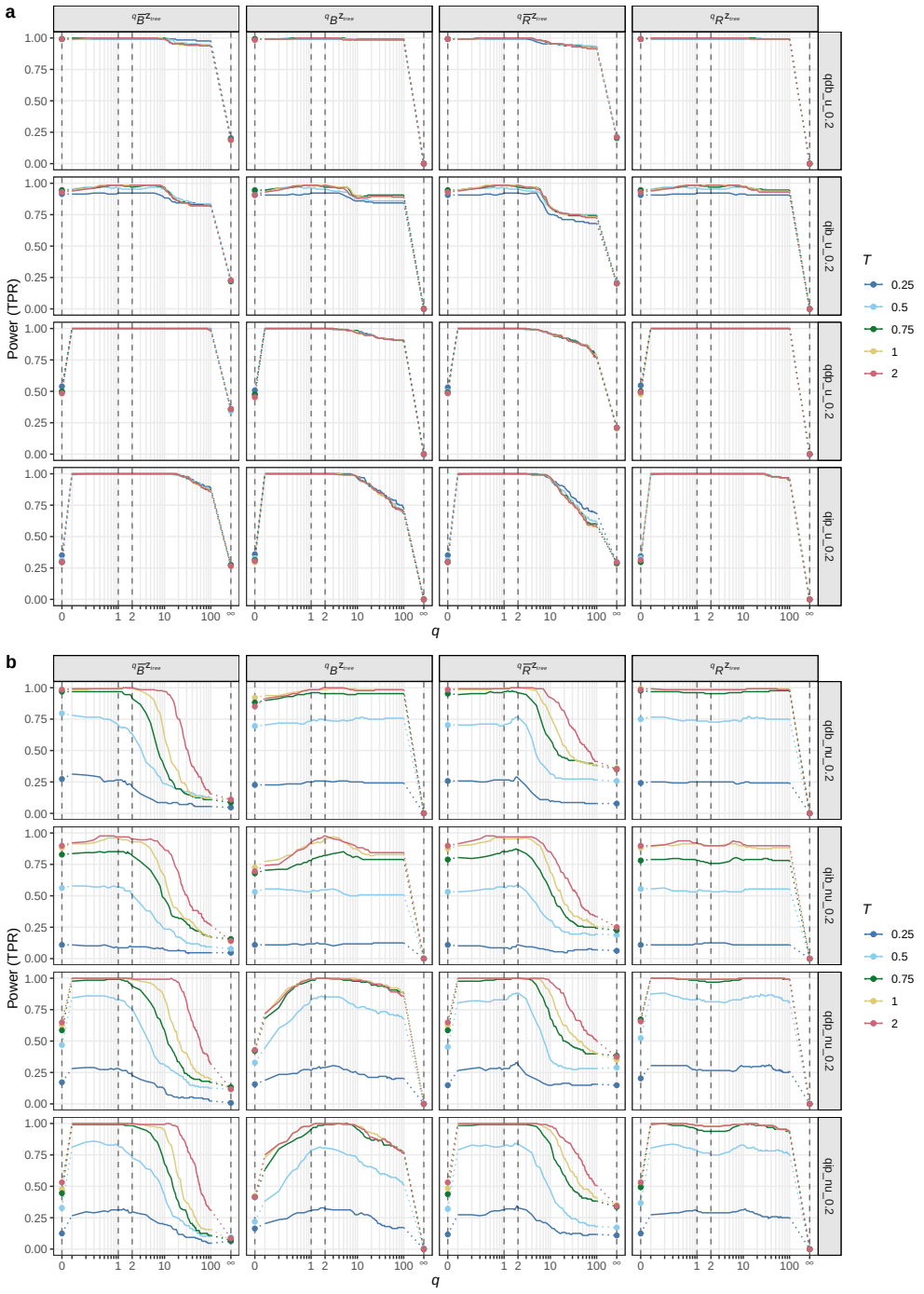


Figure B.42: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 4 (where the quantitative rate is 0.2): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

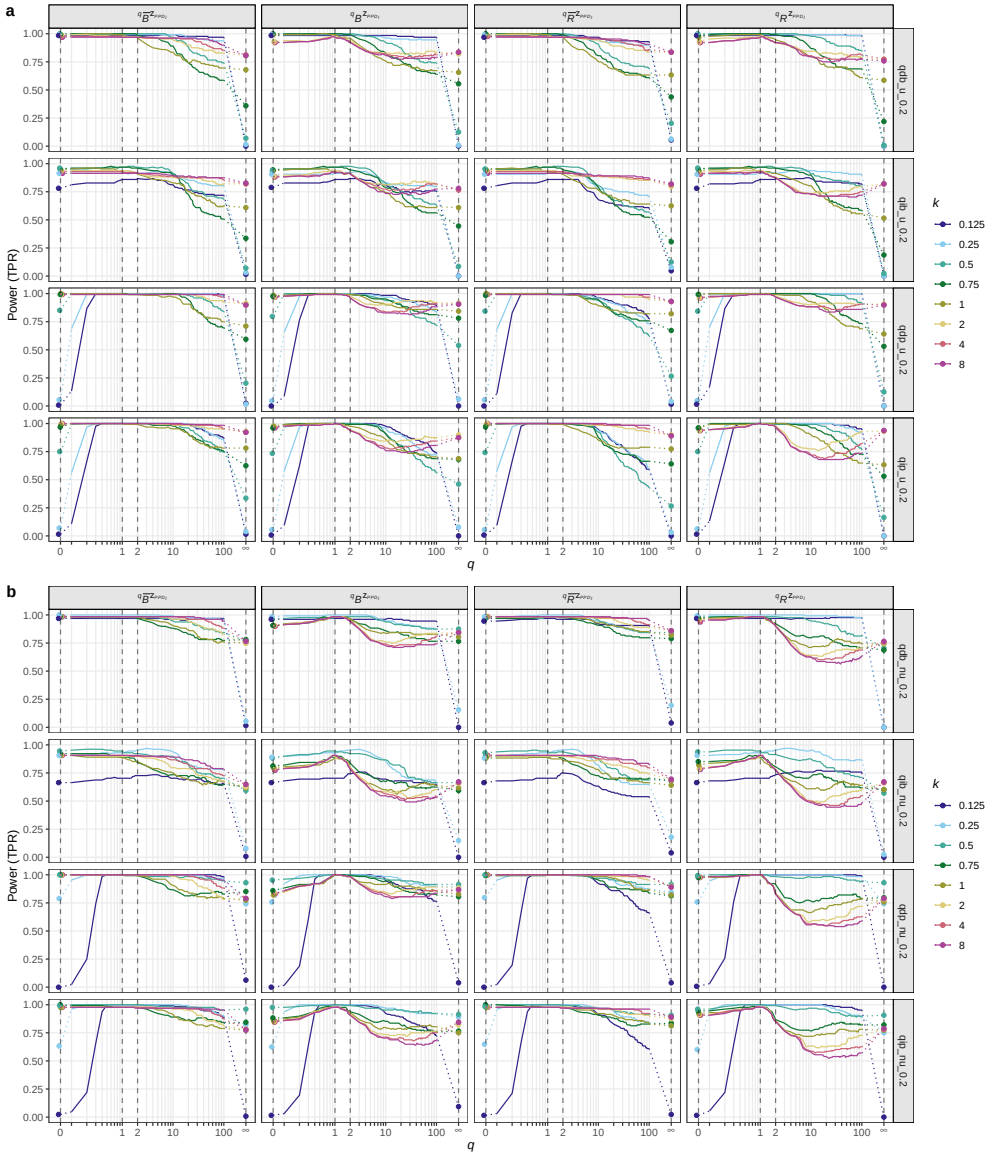


Figure B.43: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 4 (where the quantitative rate is 0.2): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

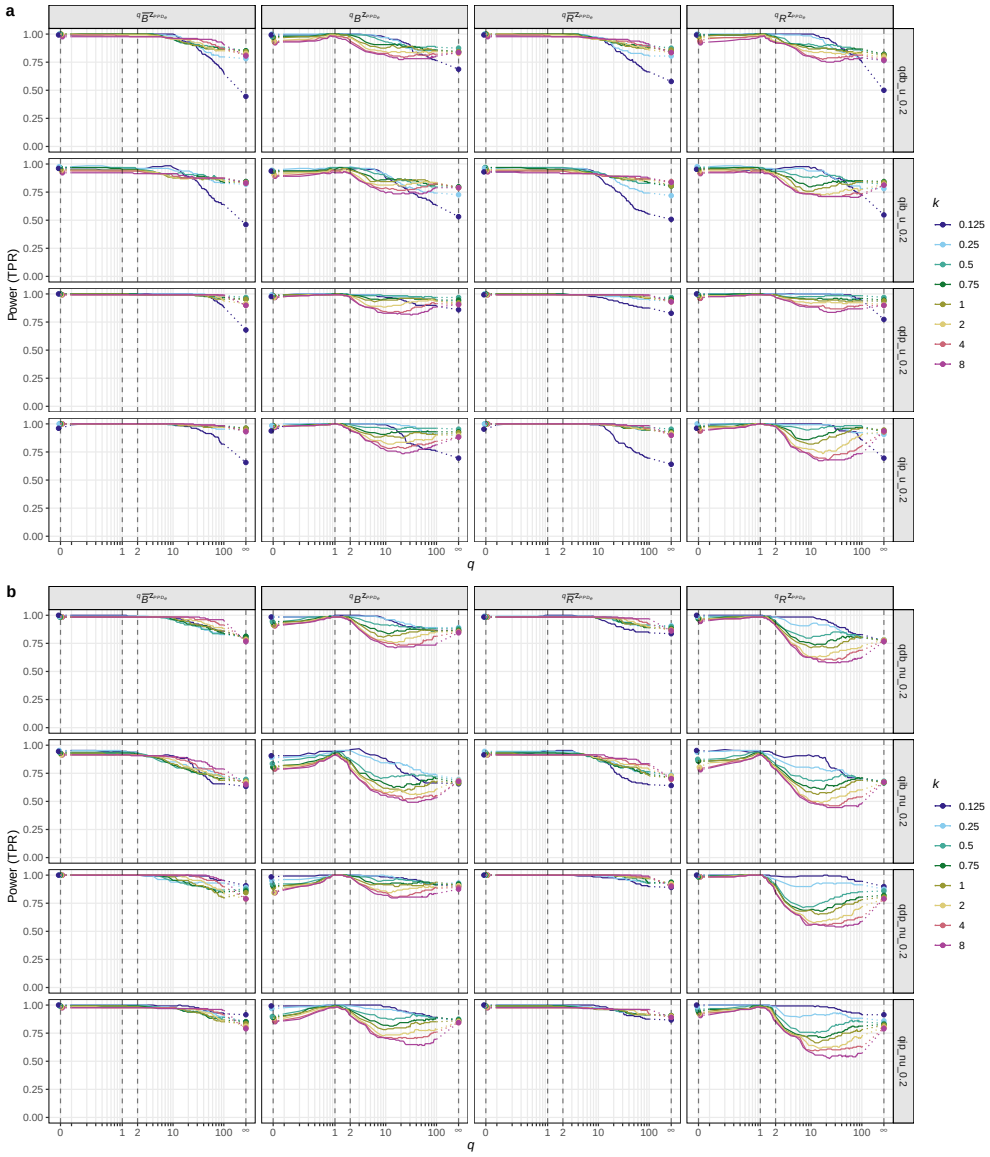


Figure B.44: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 4 (where the quantitative rate is 0.2): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

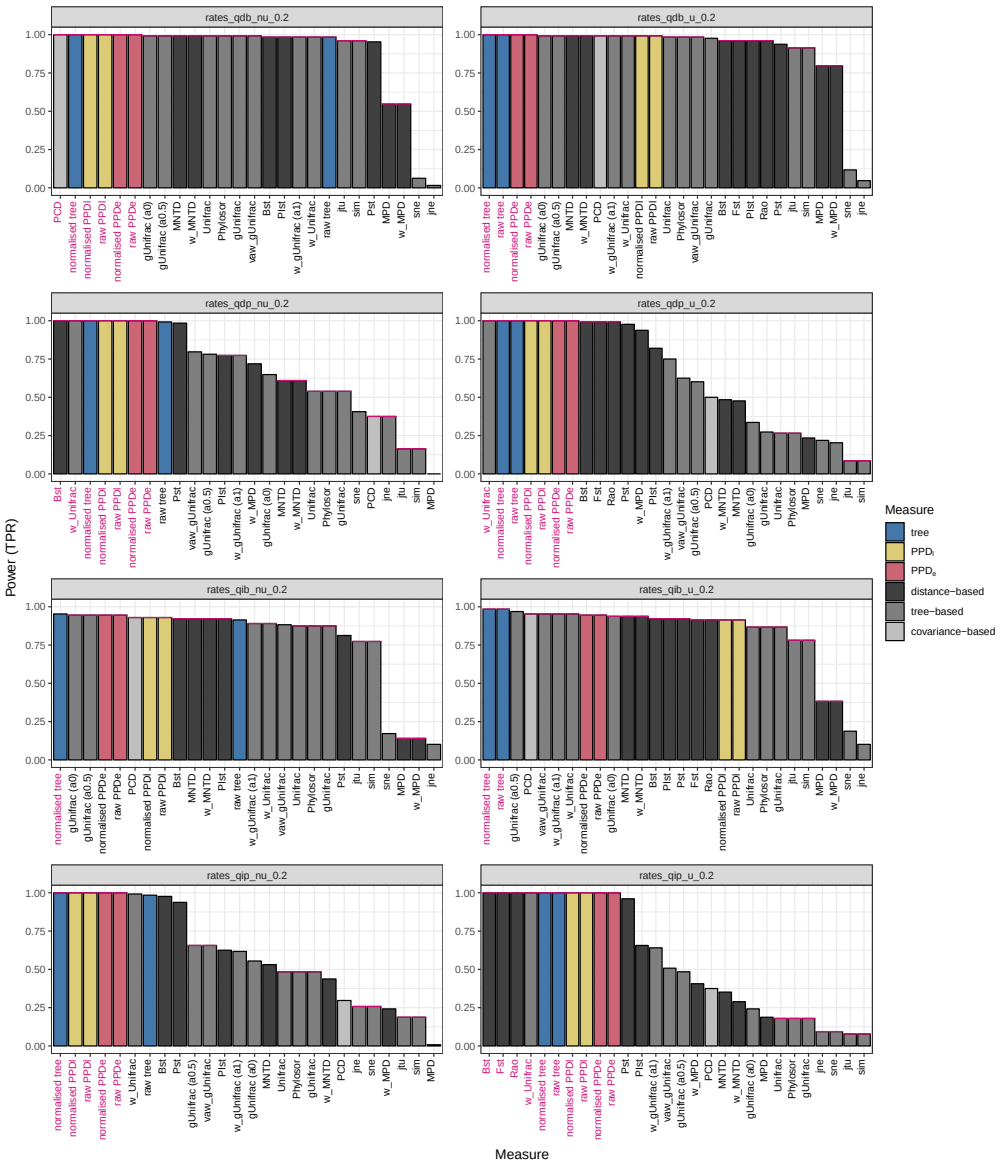


Figure B.45: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 4 (where the quantitative rate is 0.2): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

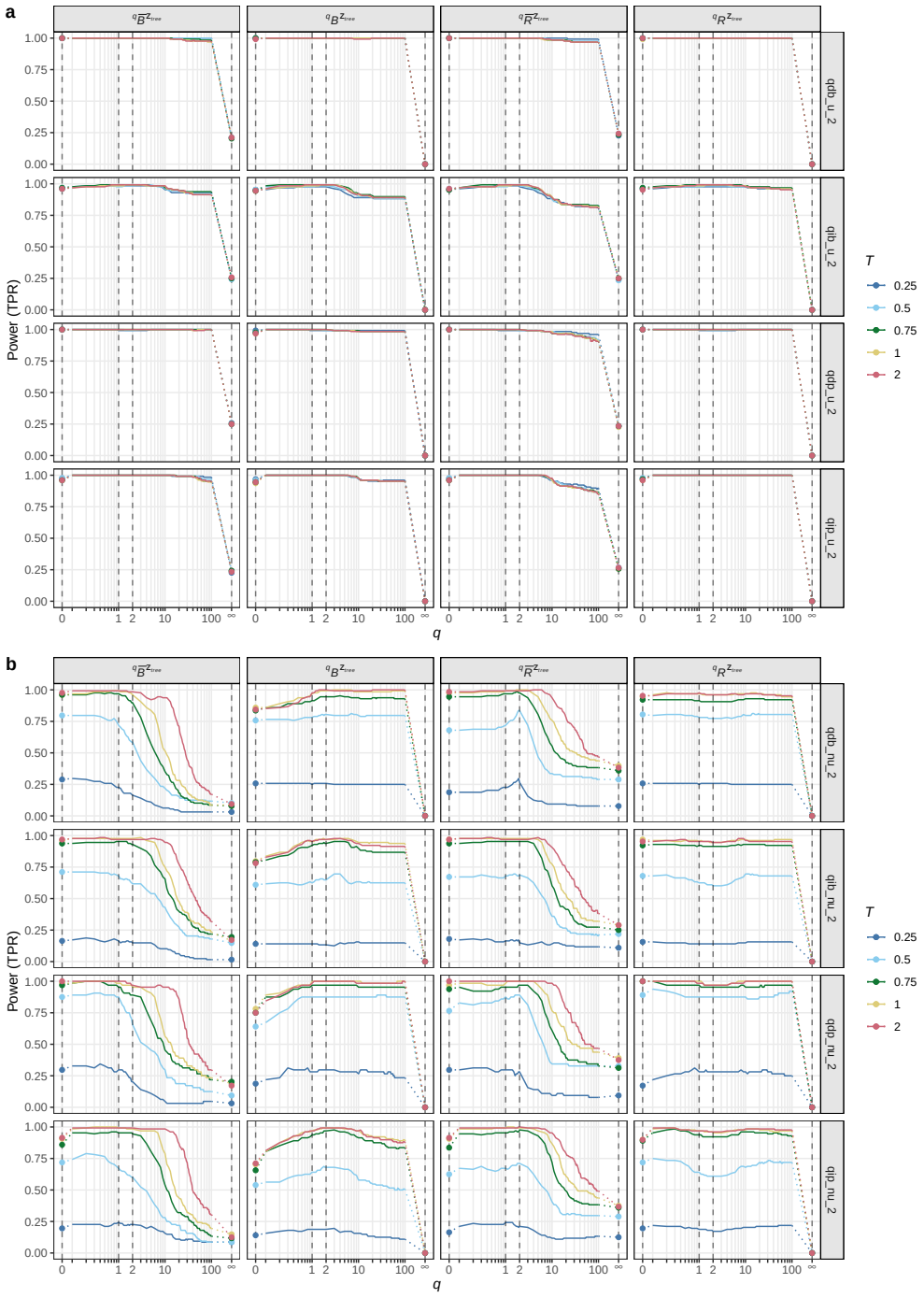


Figure B.46: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 4 (where the quantitative rate is 2): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

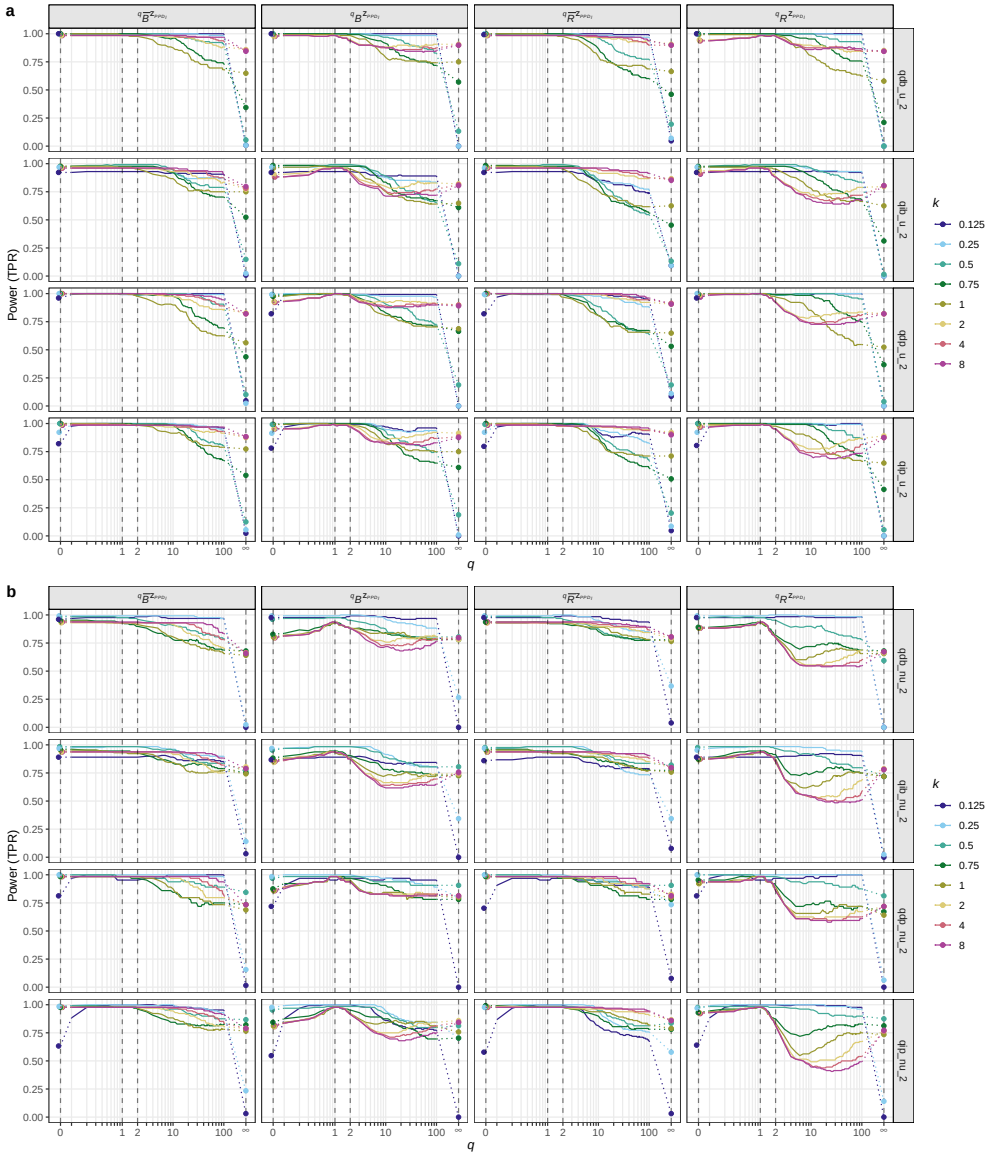


Figure B.47: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 4 (where the quantitative rate is 2): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

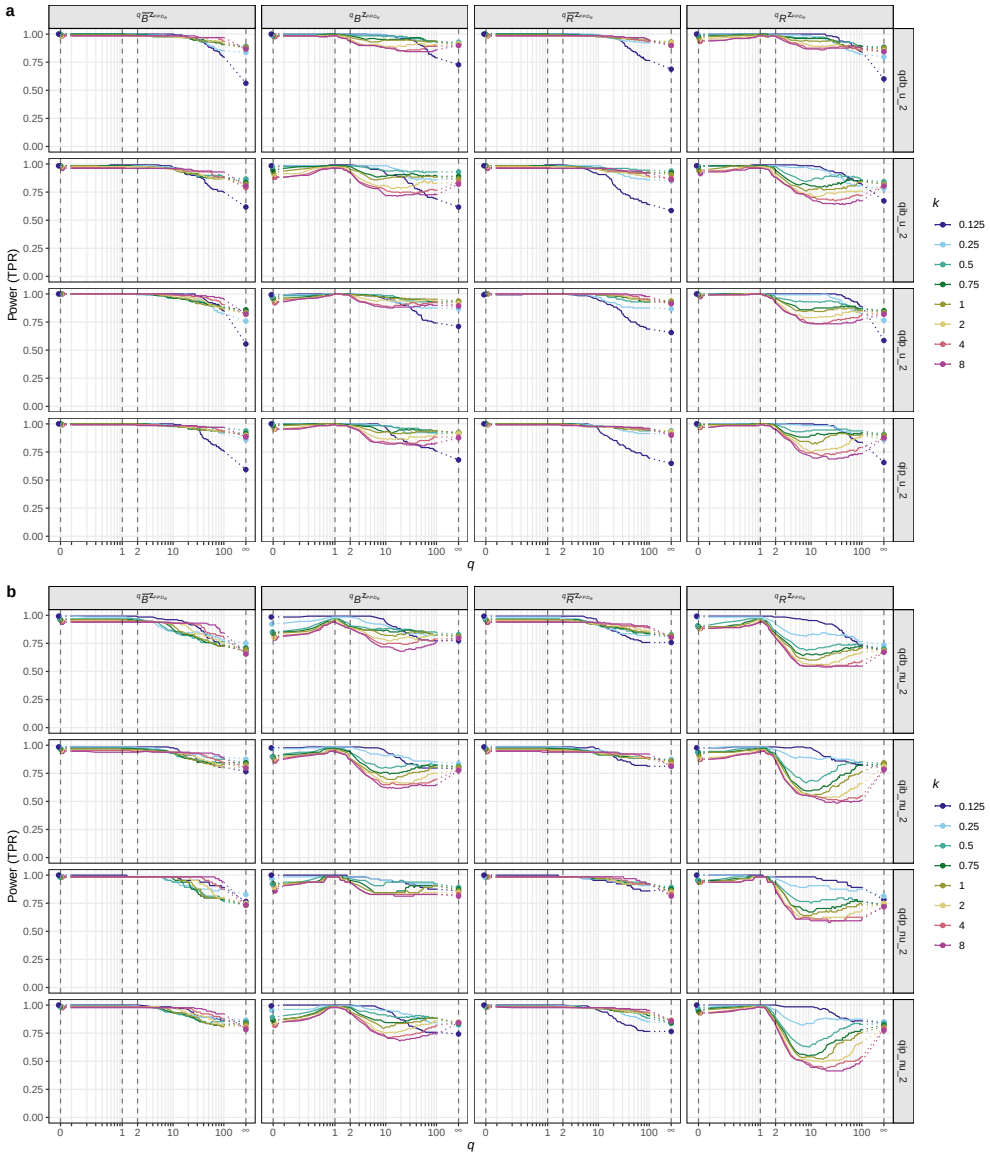


Figure B.48: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 4 (where the quantitative rate is 2): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with quantitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

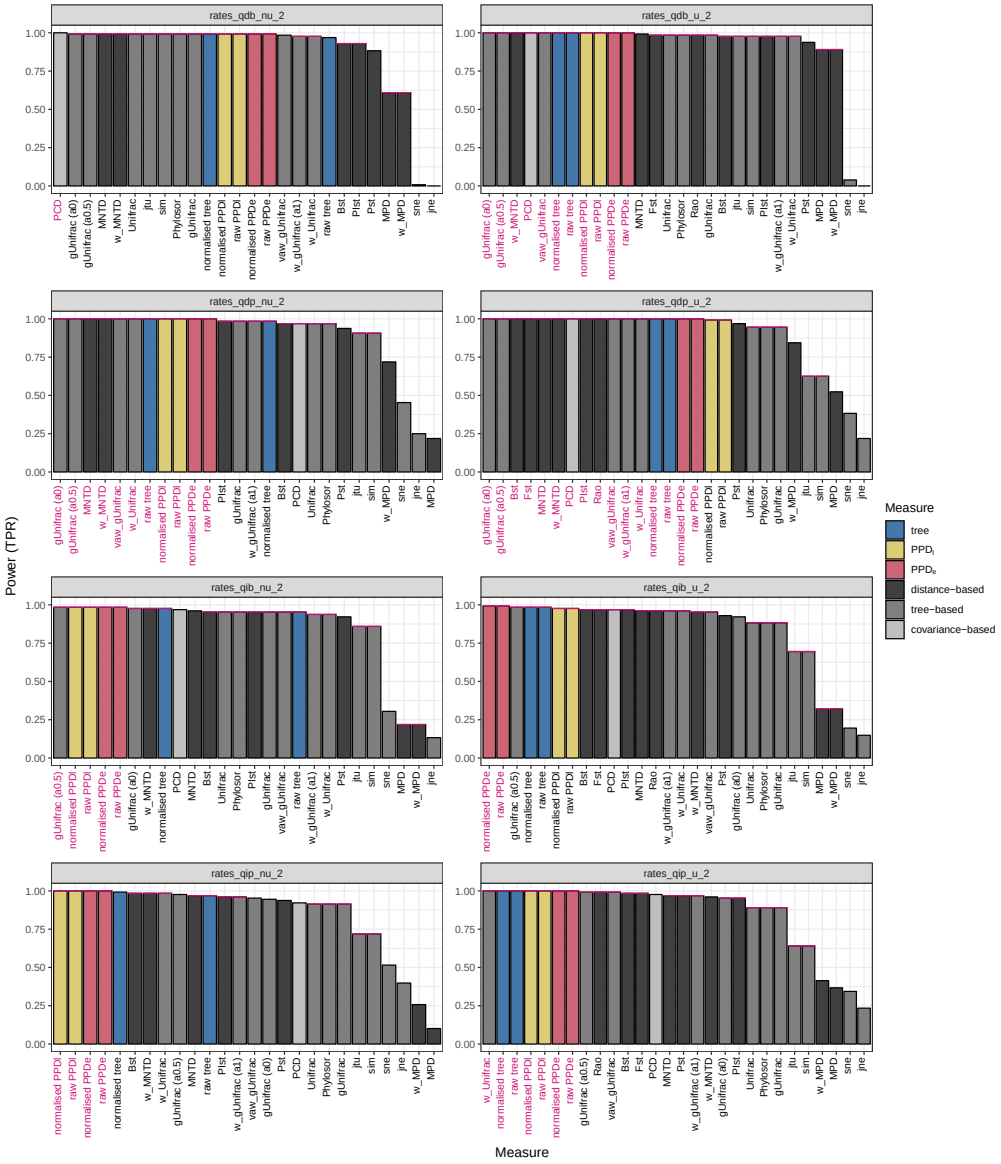


Figure B.49: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 4 (where the quantitative rate is 2): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

B.3.5.2 Qualitative partitioning strategies

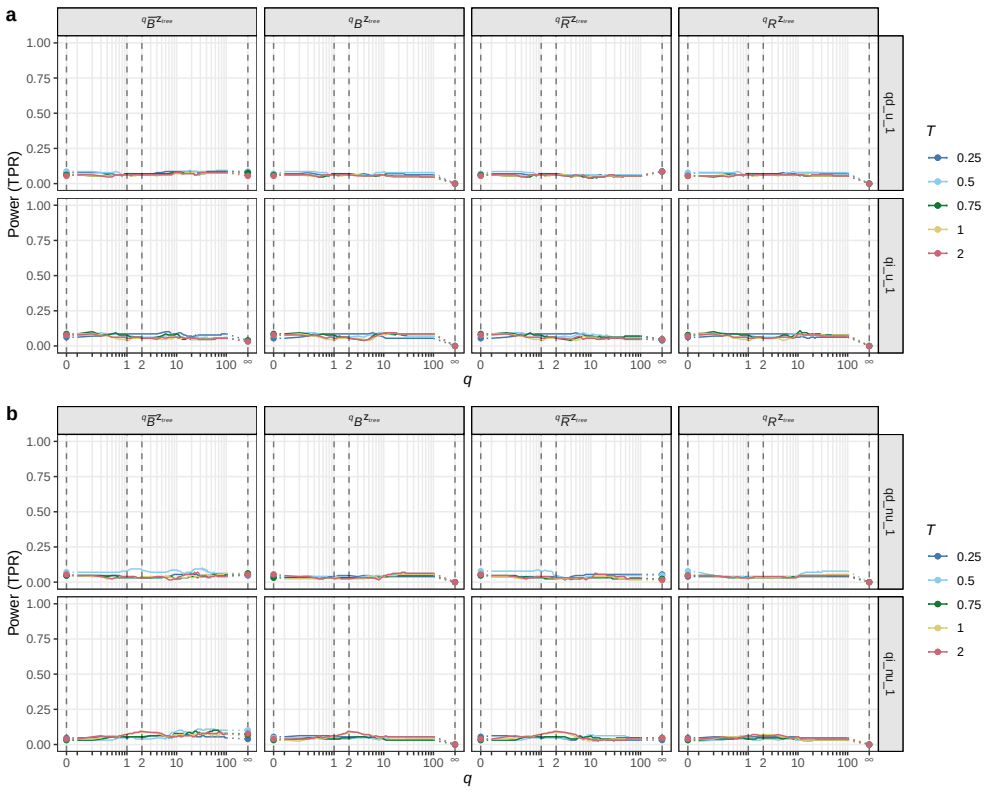


Figure B.50: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 4 (where the qualitative transition rate is 1): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

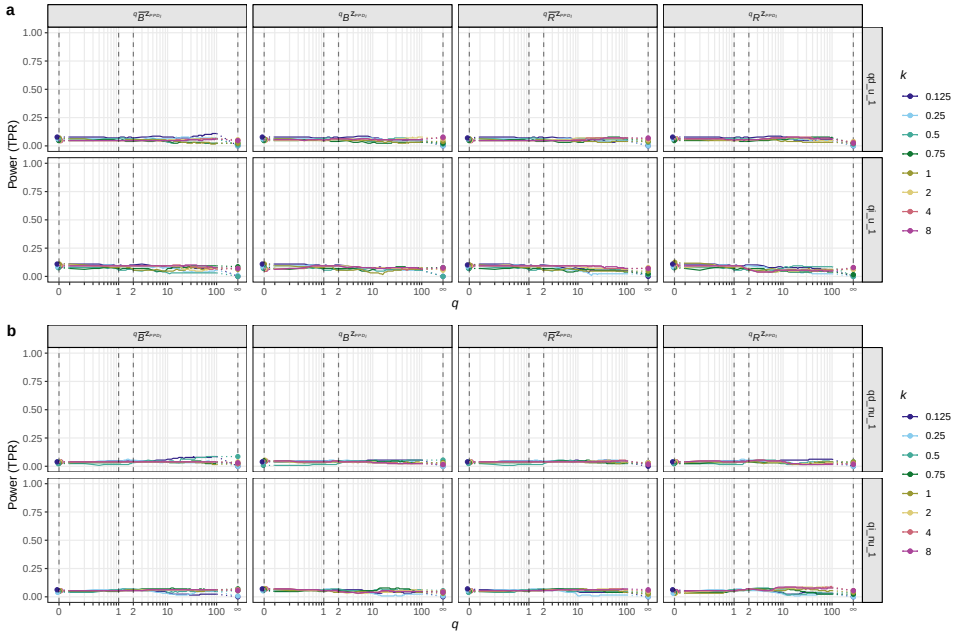


Figure B.51: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 4 (where the qualitative transition rate is 1): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

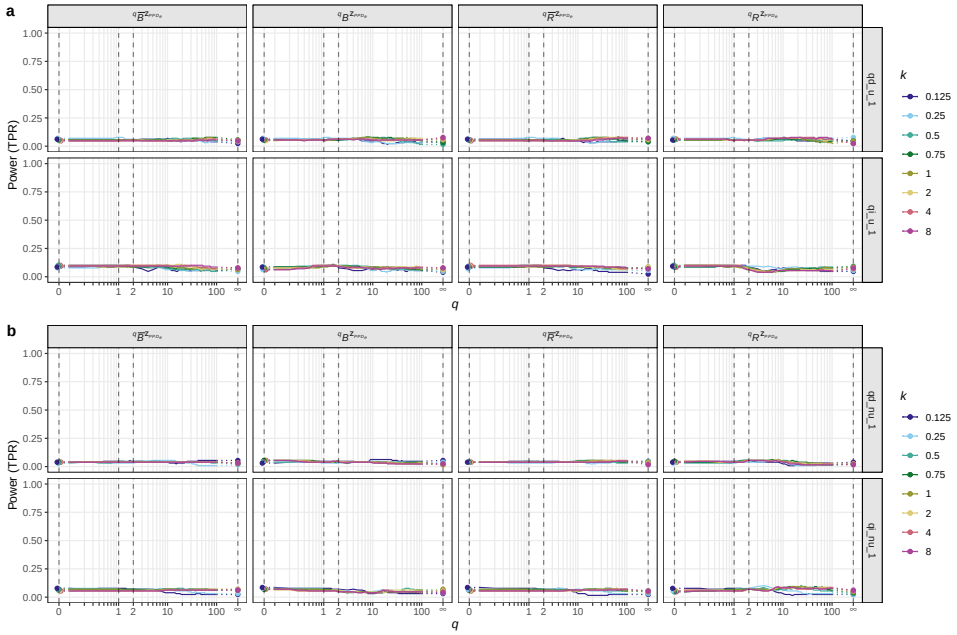


Figure B.52: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 4 (where the qualitative transition rate is 1): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

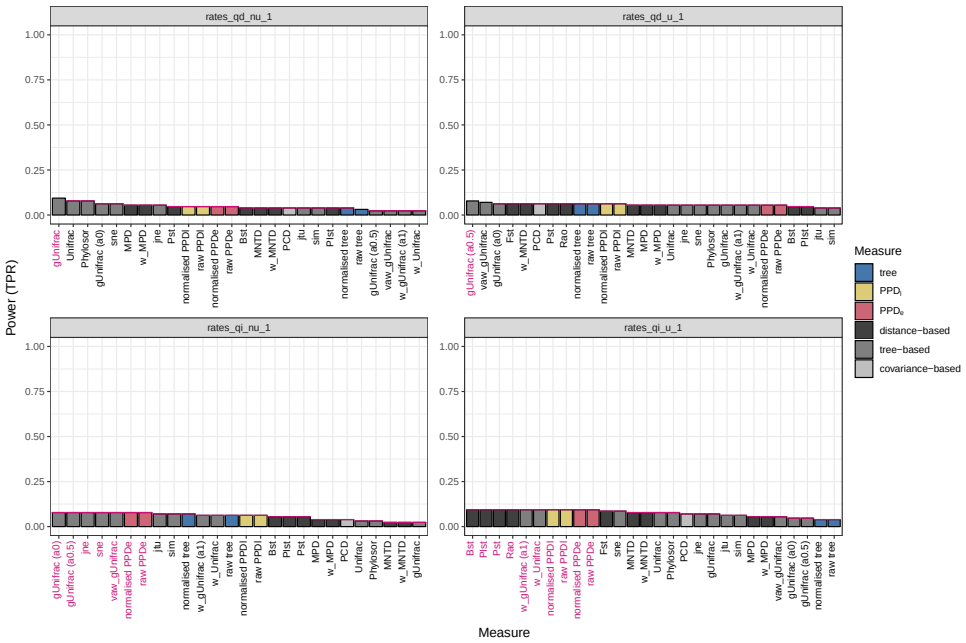


Figure B.53: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 4 (where the qualitative transition rate is 1): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

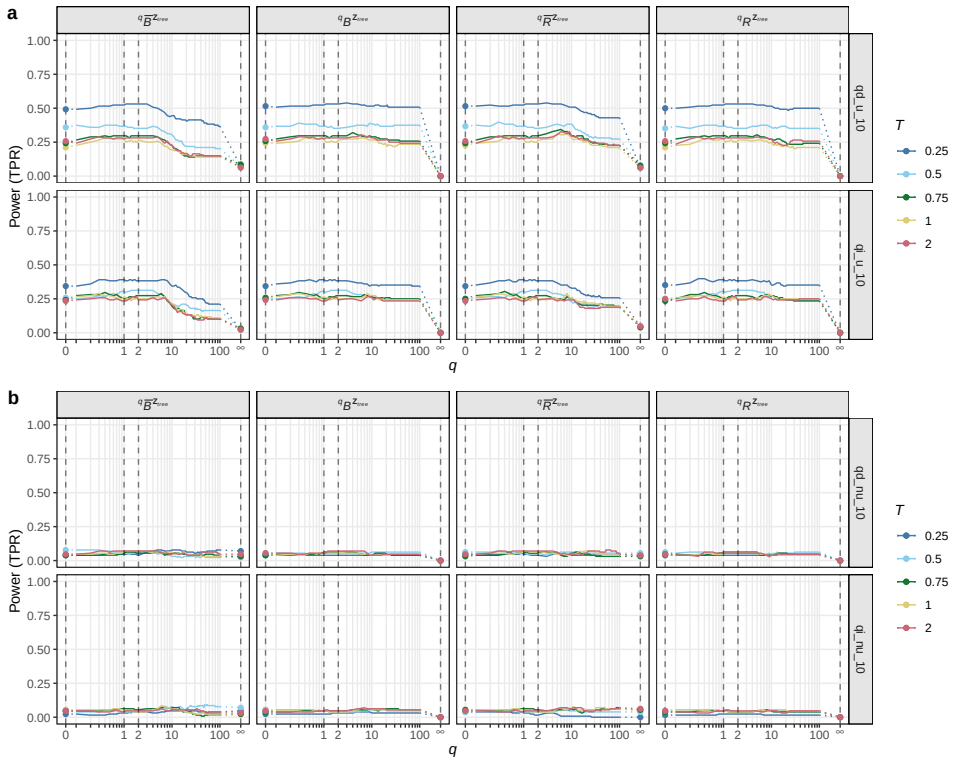


Figure B.54: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 4 (where the qualitative transition rate is 0.1): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

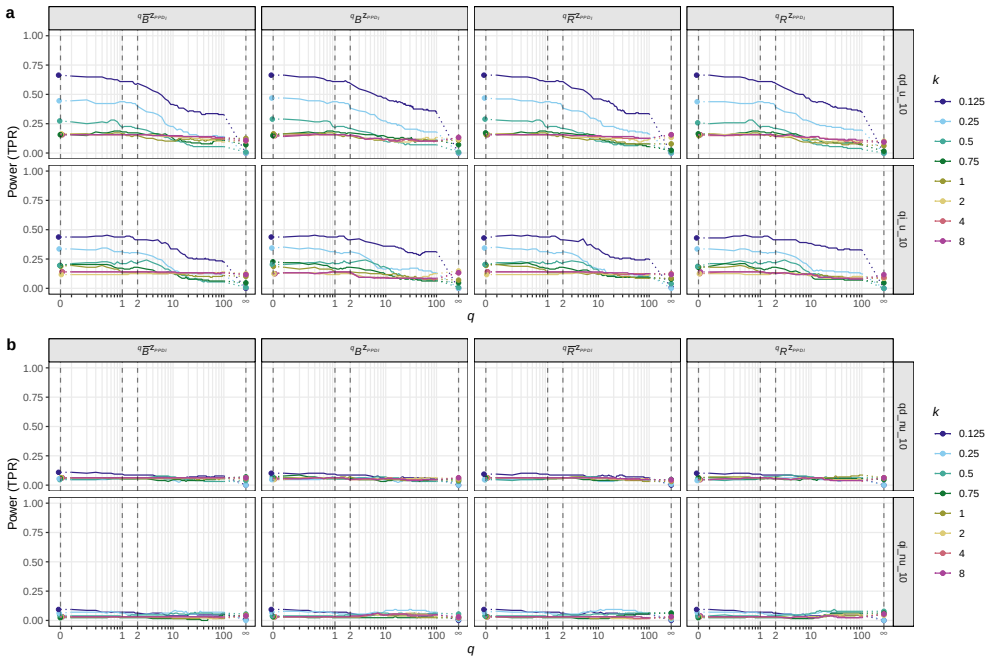


Figure B.55: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 4 (where the qualitative transition rate is 0.1): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

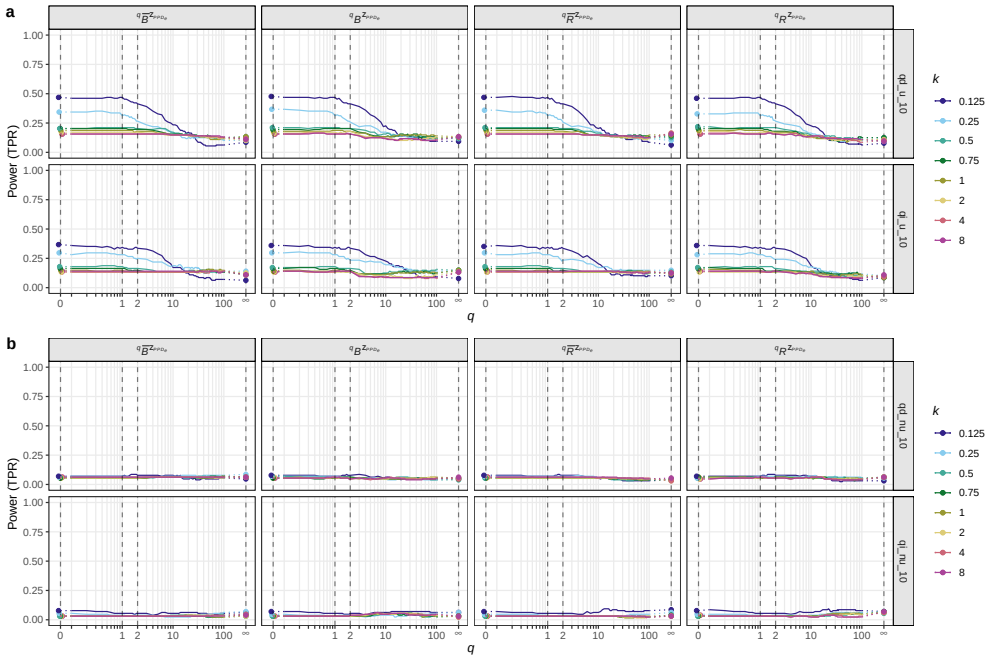


Figure B.56: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 4 (where the qualitative transition rate is 0.1): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

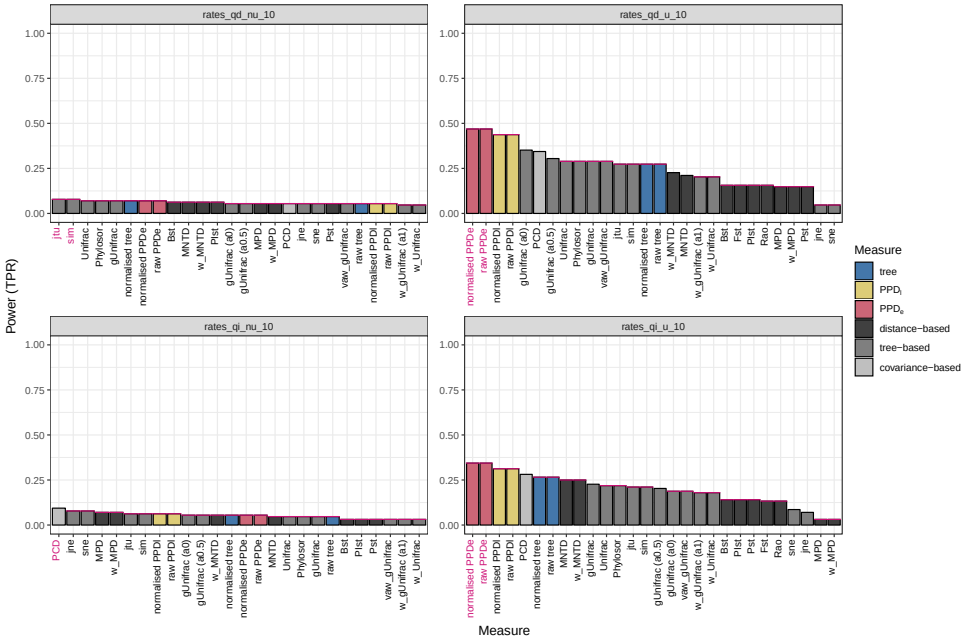


Figure B.57: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 4 (where the qualitative transition rate is 0.1): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

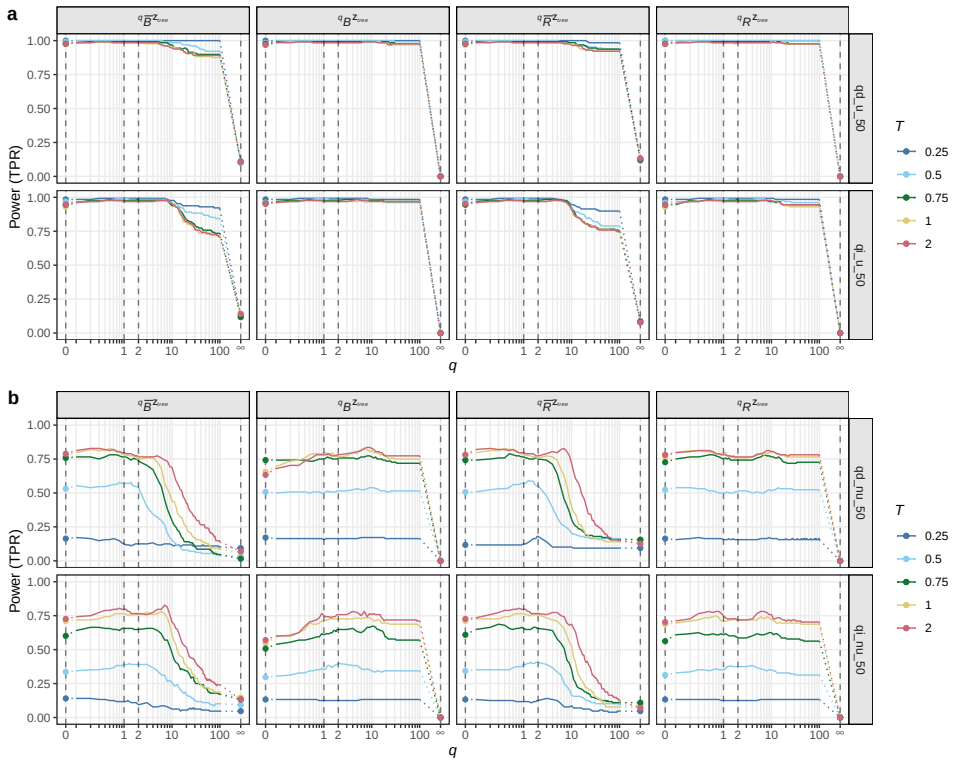


Figure B.58: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 4 (where the qualitative transition rate is 0.02): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

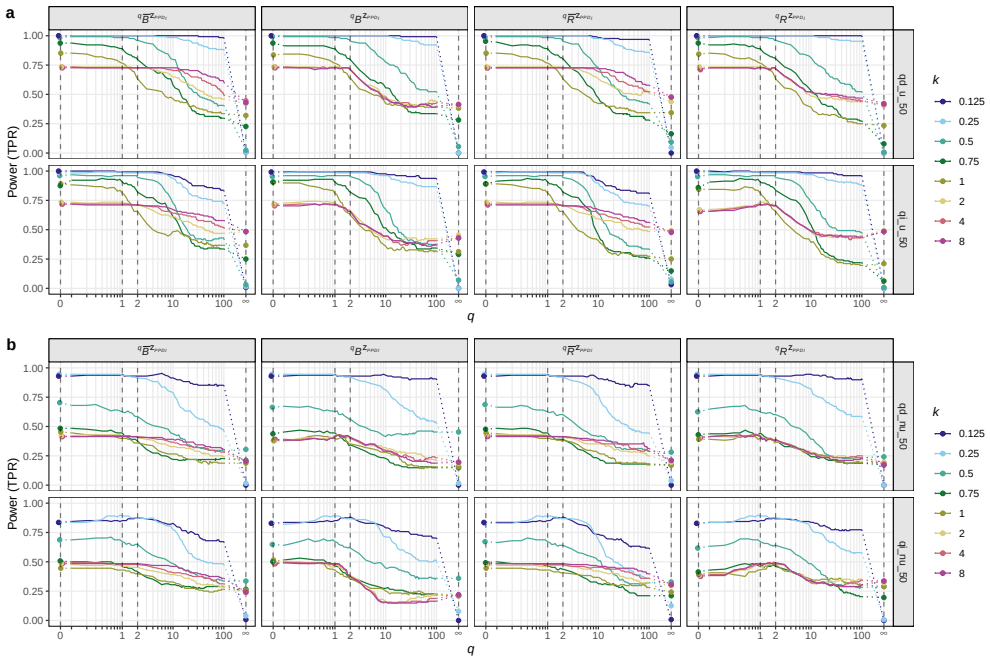


Figure B.59: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 4 (where the qualitative transition rate is 0.02): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

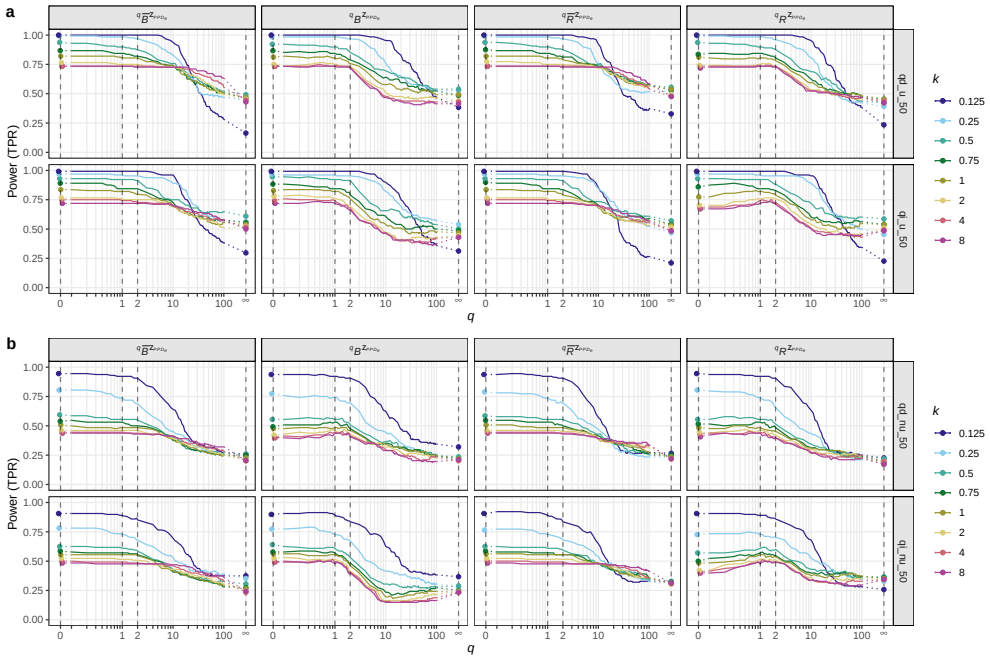


Figure B.60: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 4 (where the qualitative transition rate is 0.02): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

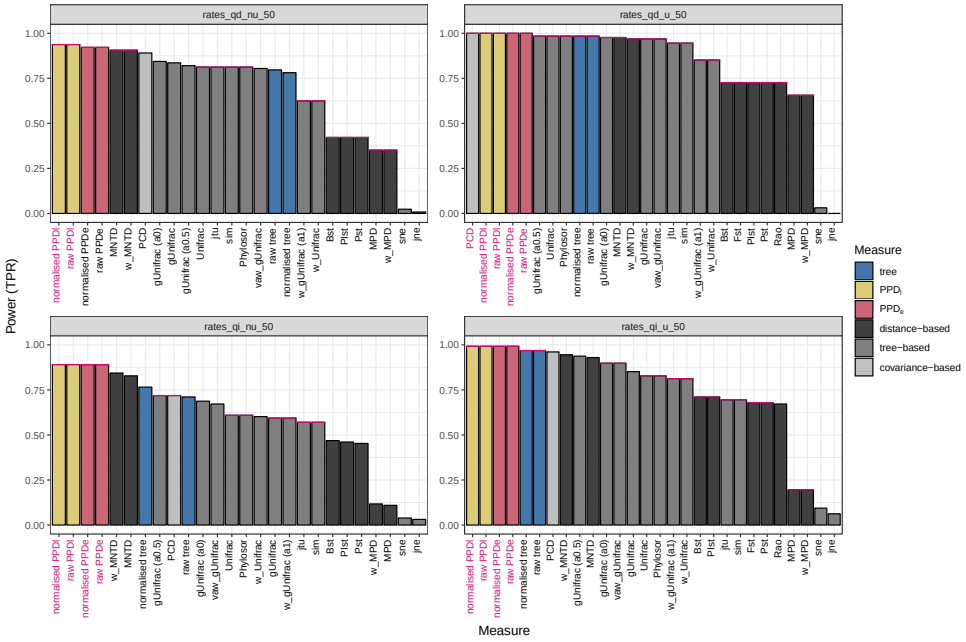


Figure B.61: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 4 (where the qualitative transition rate is 0.02): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

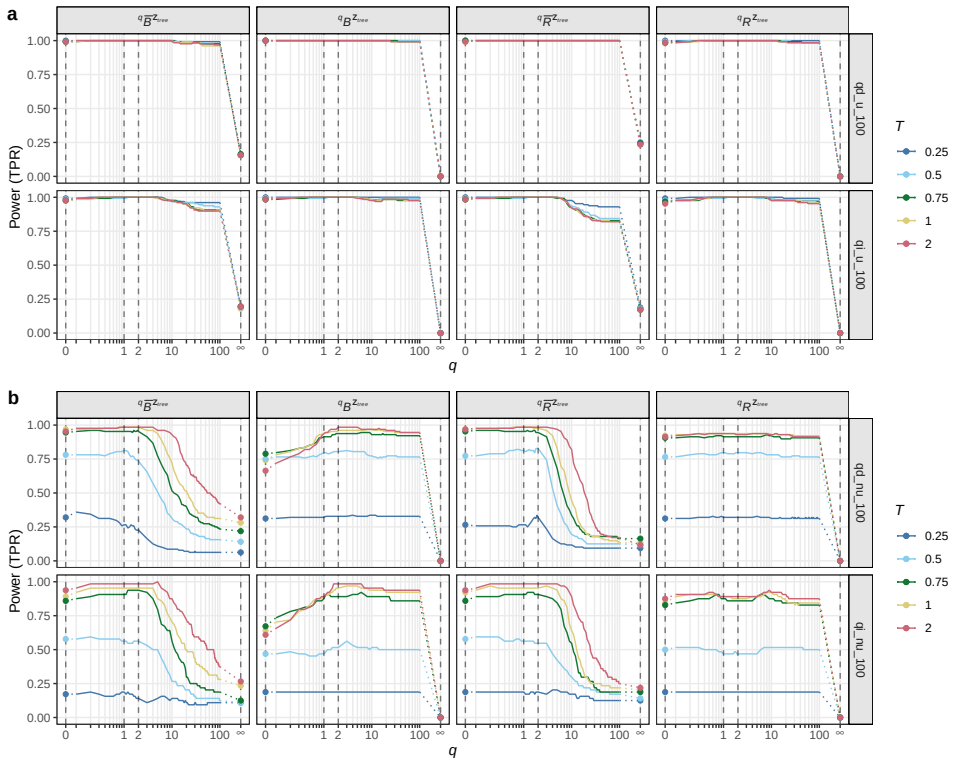


Figure B.62: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 4 (where the qualitative transition rate is 0.01): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

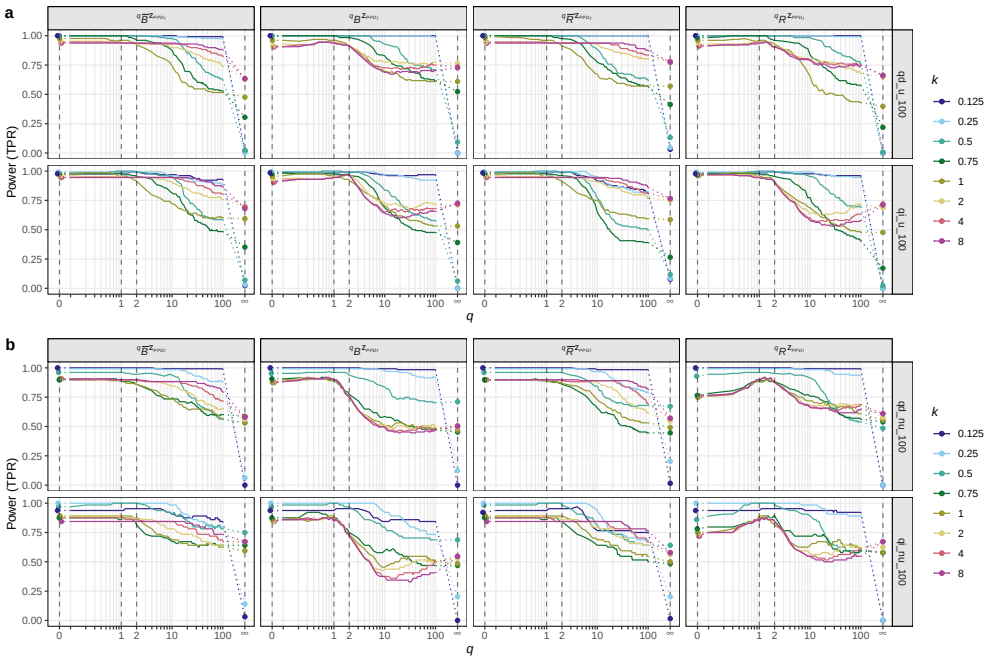


Figure B.63: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 4 (where the qualitative transition rate is 0.01): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

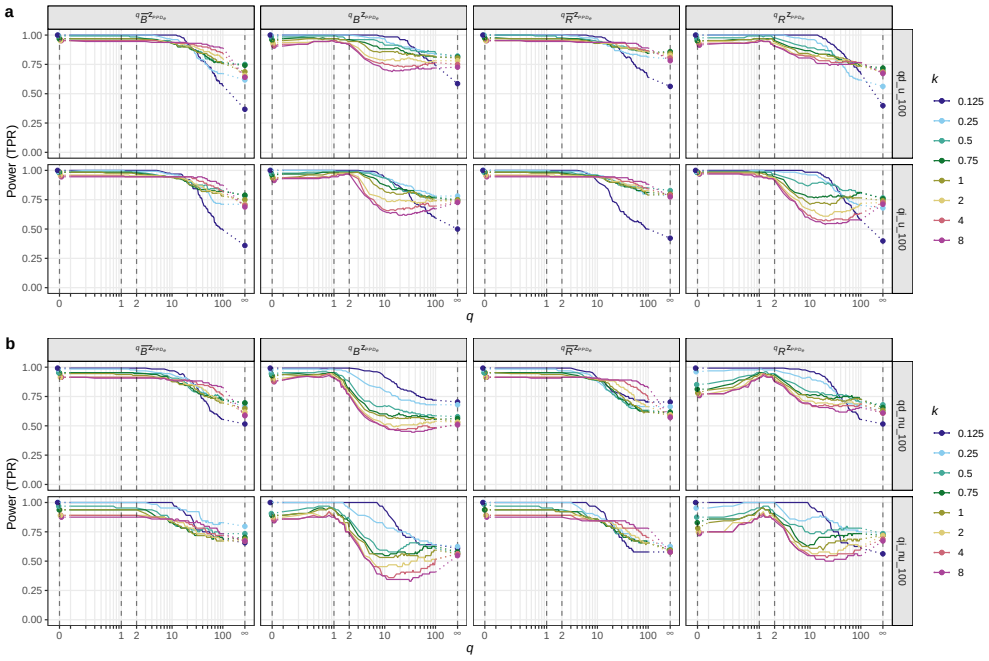


Figure B.64: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 4 (where the qualitative transition rate is 0.01): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes with qualitative traits, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

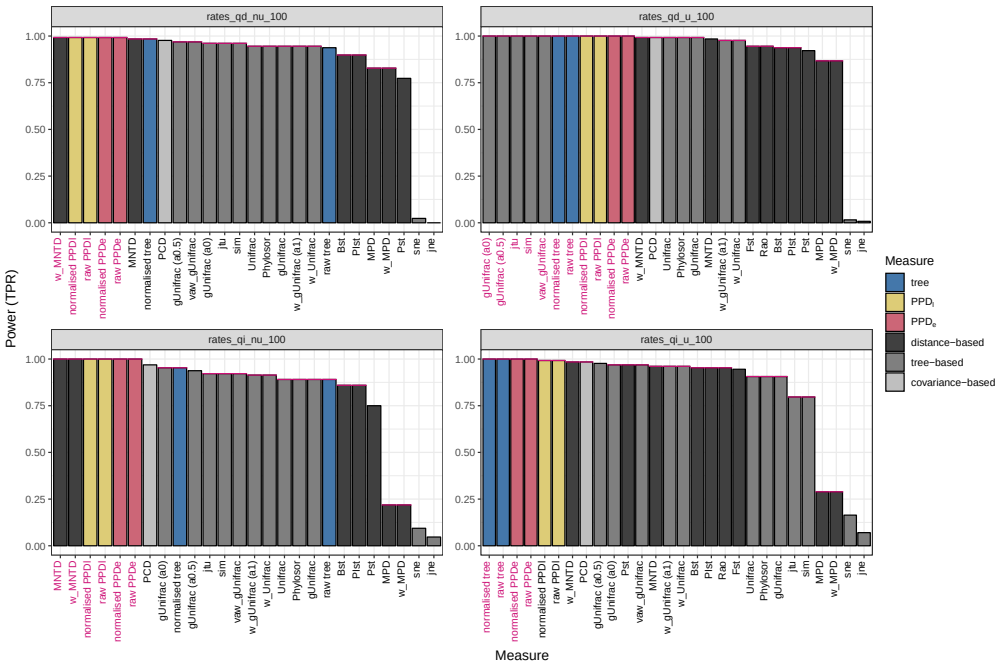


Figure B.65: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 4 (where the qualitative transition rate is 0.01): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

B.3.6 Experiment 5: Multiple subcommunities

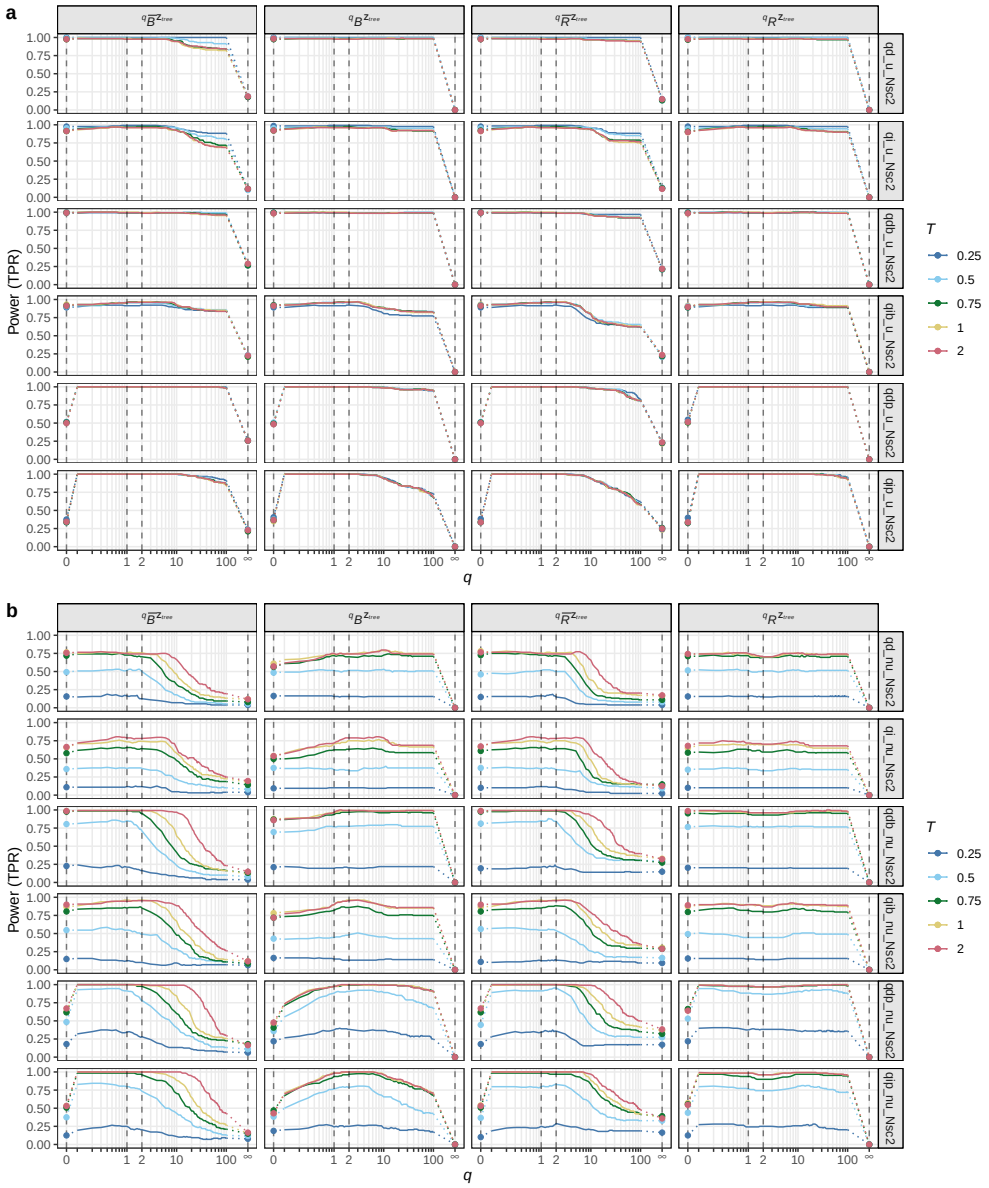


Figure B.66: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 5 (for 2 subcommunities): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

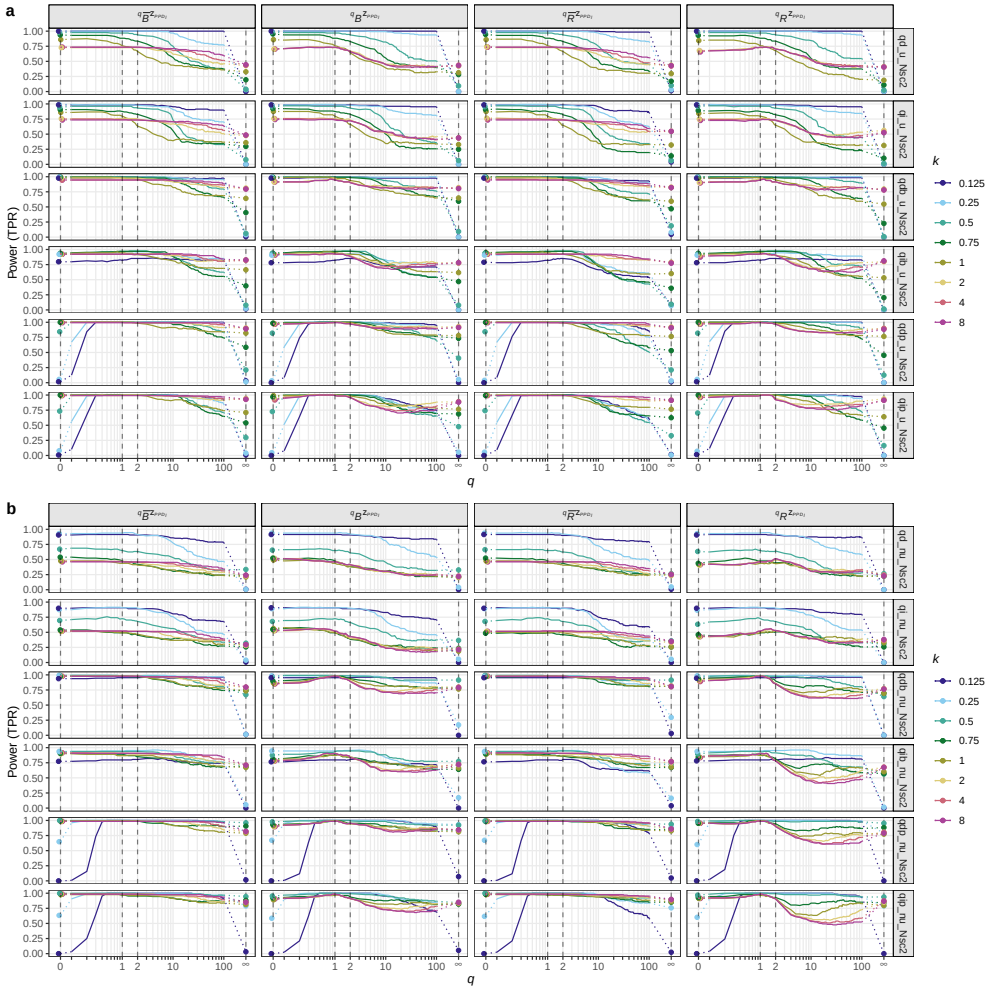


Figure B.67: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 5 (for 2 subcommunities): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

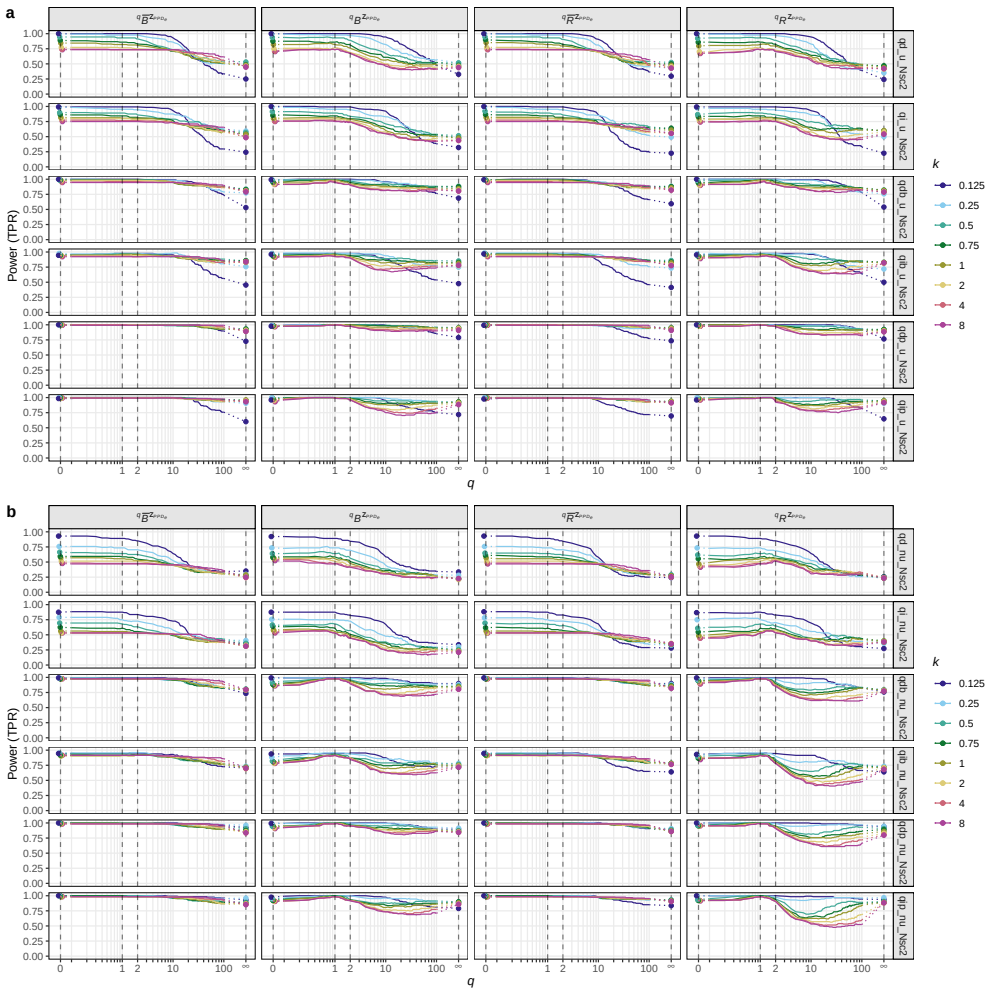


Figure B.68: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 5 (for 2 subcommunities): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

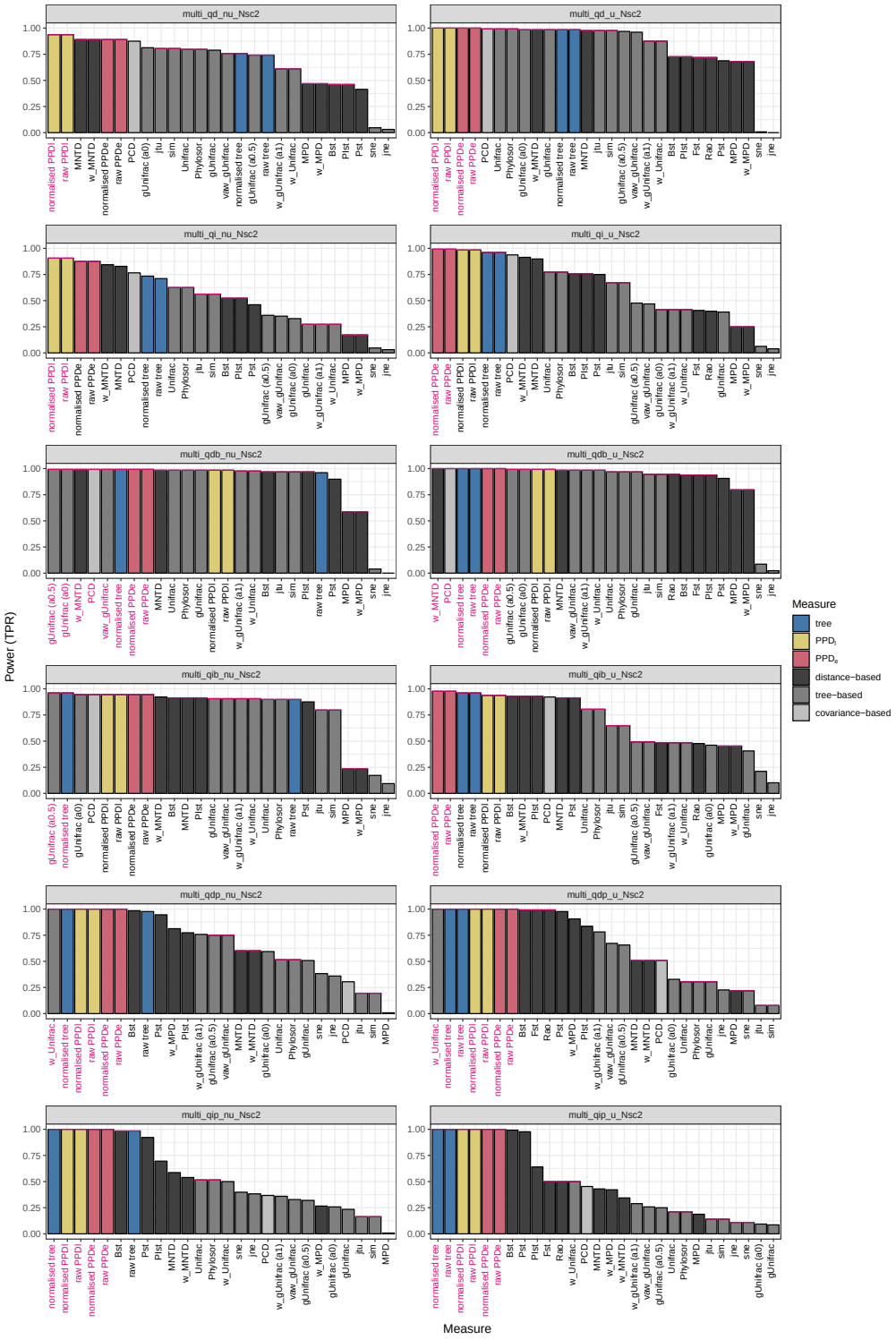


Figure B.69: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 5 (for 2 subcommunities): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

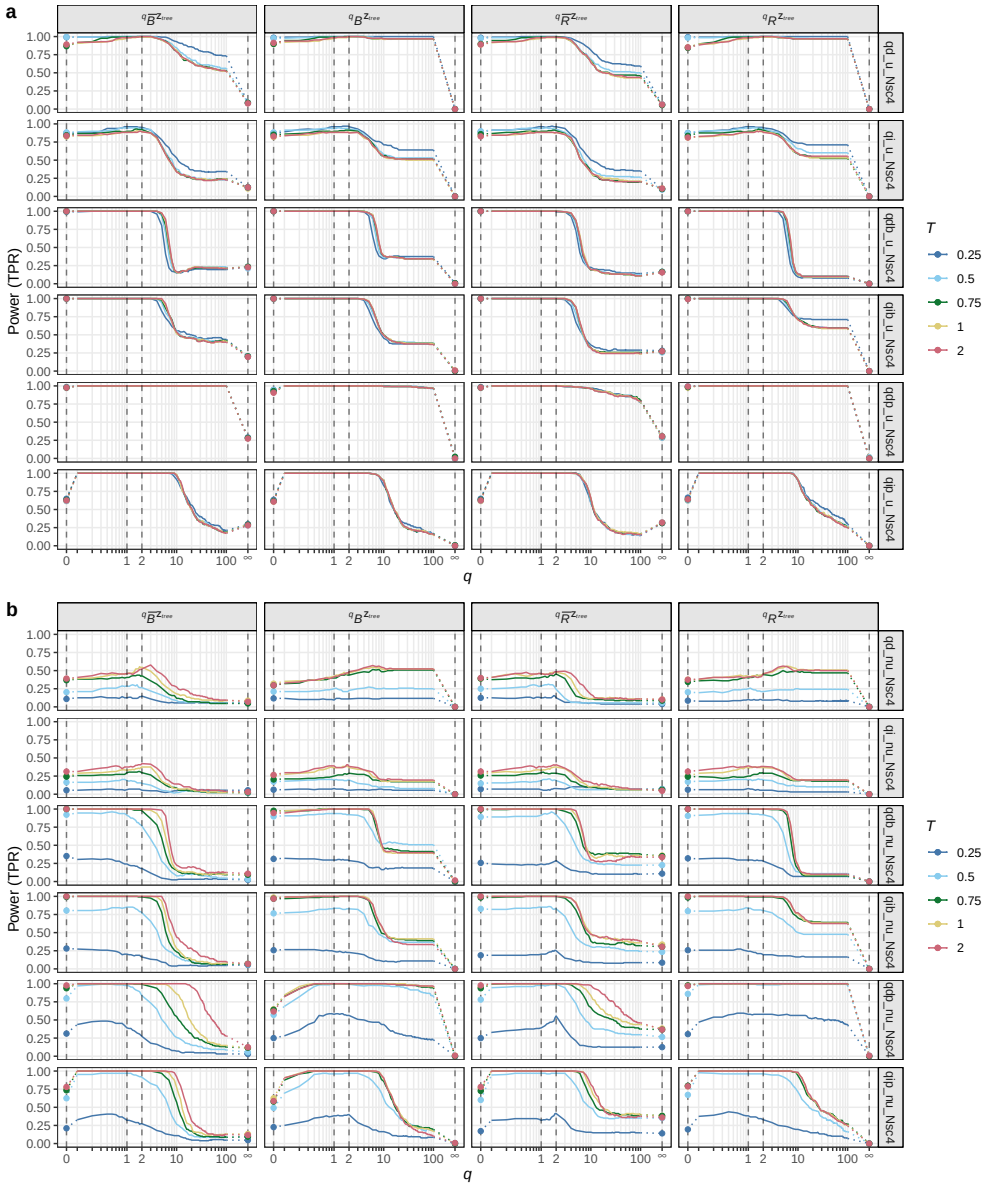


Figure B.70: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 5 (for 4 subcommunities): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

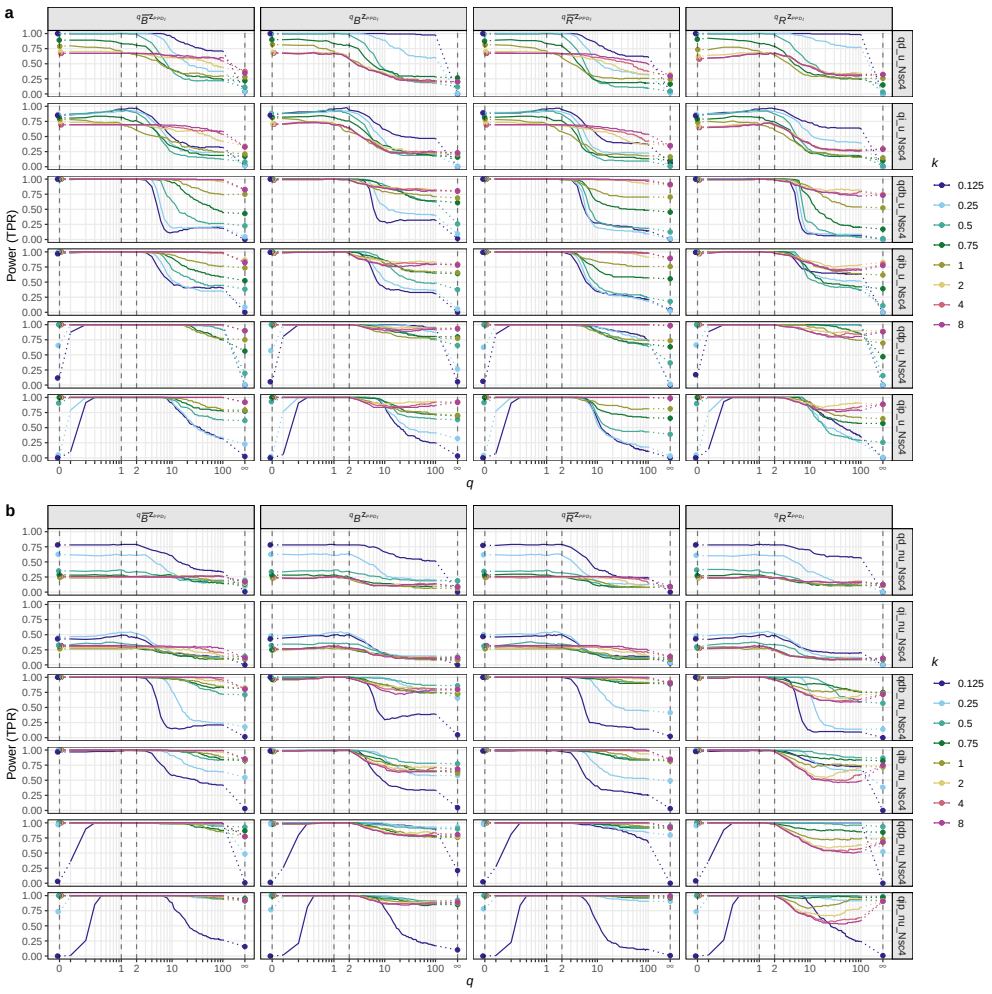


Figure B.71: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 5 (for 4 subcommunities): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

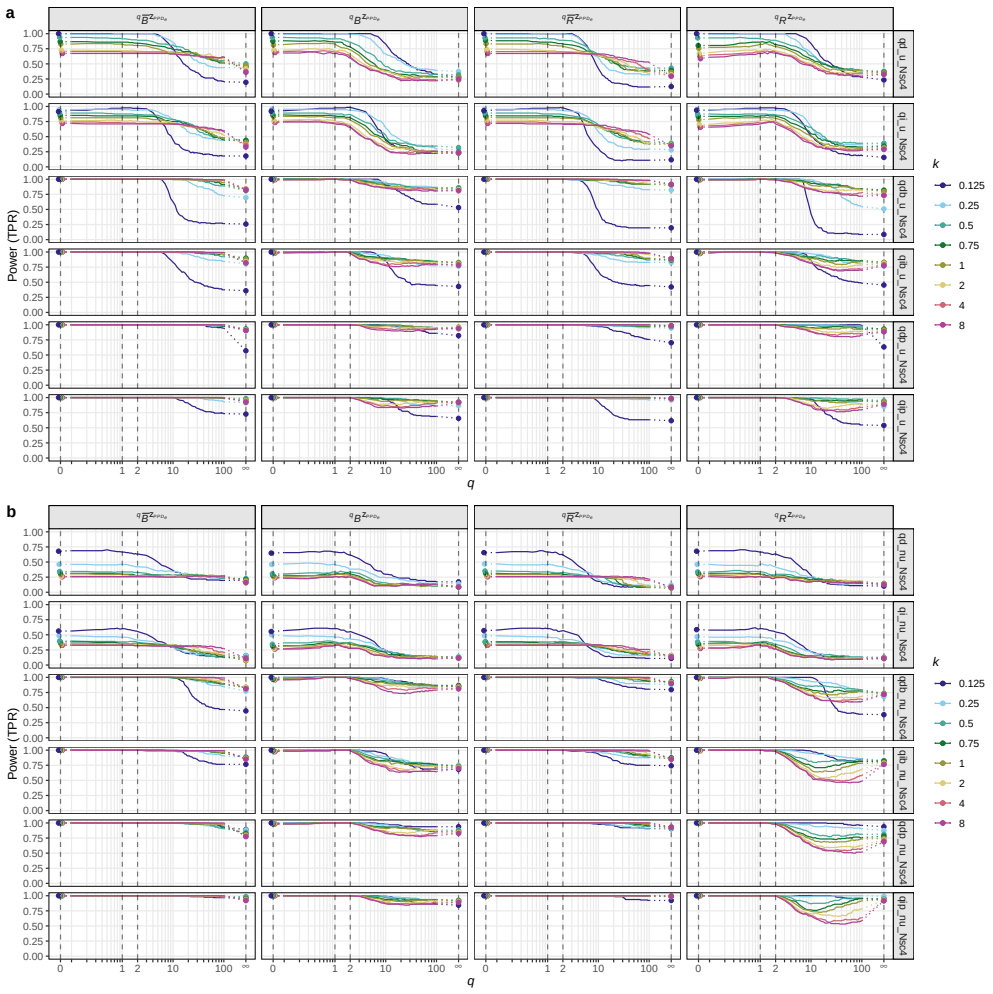


Figure B.72: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 5 (for 4 subcommunities): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

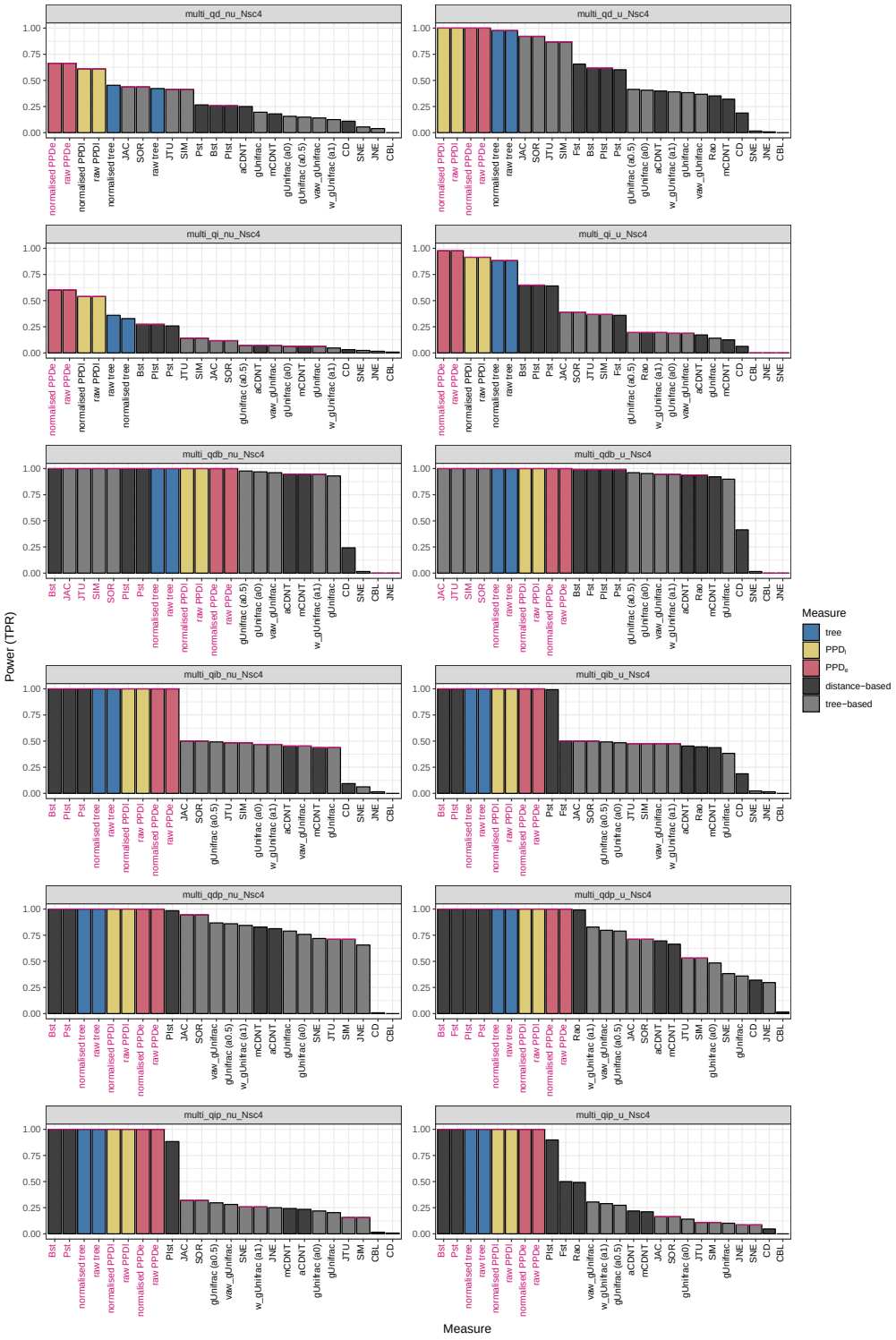


Figure B.73: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 5 (for 4 subcommunities): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

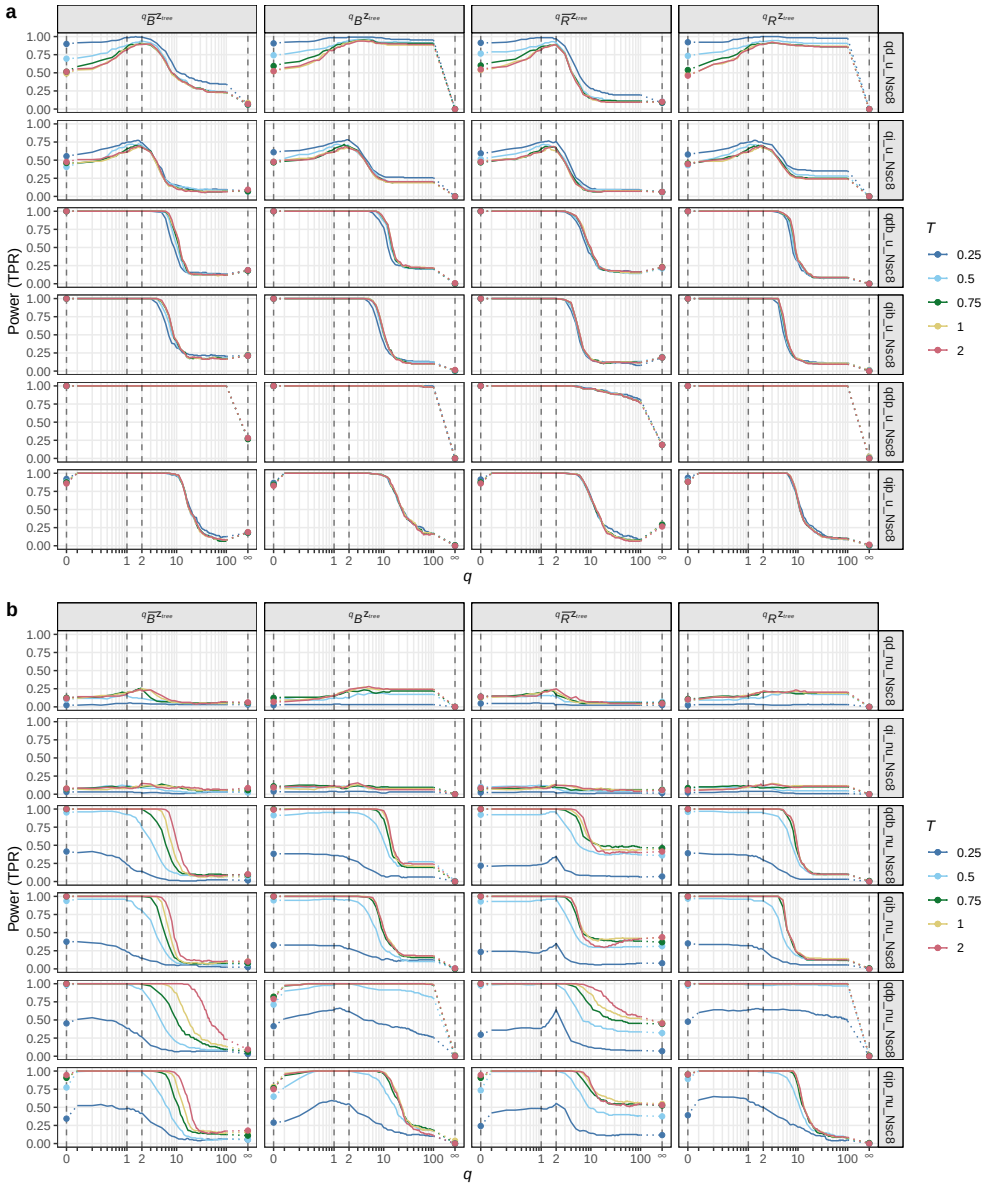


Figure B.74: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{tree}}$, ${}^{q}R^{\mathbf{Z}_{tree}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$, and ${}^{q}B^{\mathbf{Z}_{tree}}$, in Experiment 5 (for 8 subcommunities): $\mathbf{Z}_{tree}$ denotes that similarity was calculated using phylogenetic tree-based methods, for different values of T. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

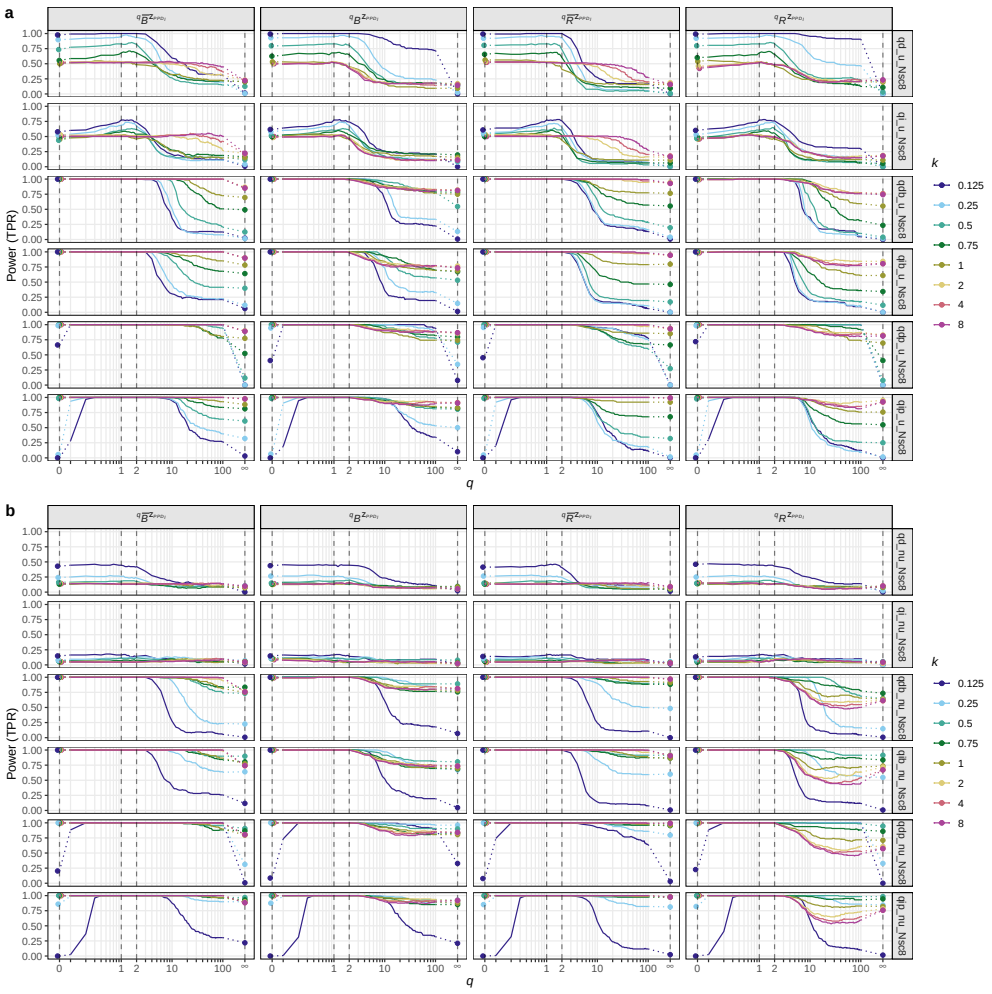


Figure B.75: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_l}}$, ${}^{q}R^{\mathbf{Z}_{PPD_l}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_l}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_l}}$, in Experiment 5 (for 8 subcommunities): $\mathbf{Z}_{PPD_l}$ denotes that similarity was calculated as a linear transformation of phylogenetic distance, for different values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

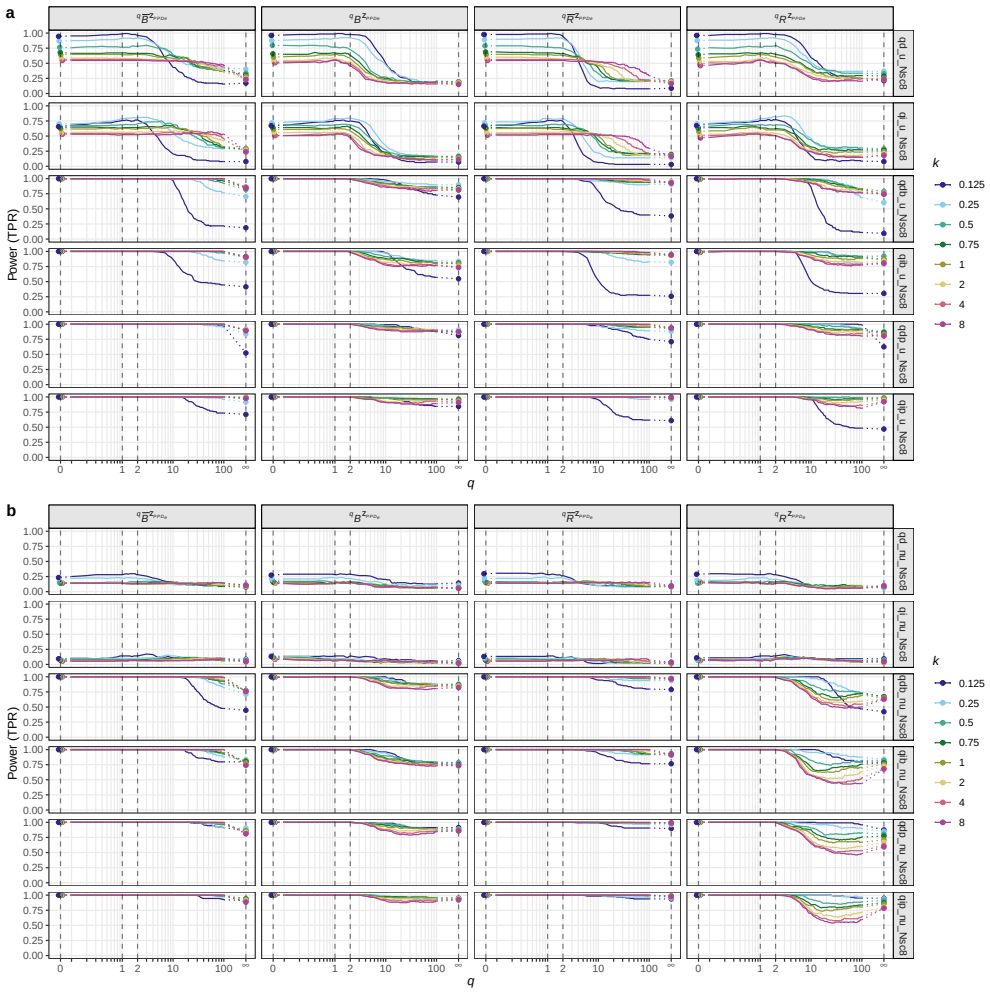


Figure B.76: Plots of power against q, calculated for ${}^{q}\bar{R}^{\mathbf{Z}_{PPD_e}}$, ${}^{q}R^{\mathbf{Z}_{PPD_e}}$, ${}^{q}\bar{B}^{\mathbf{Z}_{PPD_e}}$, and ${}^{q}B^{\mathbf{Z}_{PPD_e}}$, in Experiment 5 (for 8 subcommunities): $\mathbf{Z}_{PPD_e}$ denotes that similarity was calculated as an exponential transformation of phylogenetic distance, for all values of k. Diversity measures were calculated for all subcommunity partition schemes, for (a) non-ultrametric and (b) ultrametric phylogenetic trees.

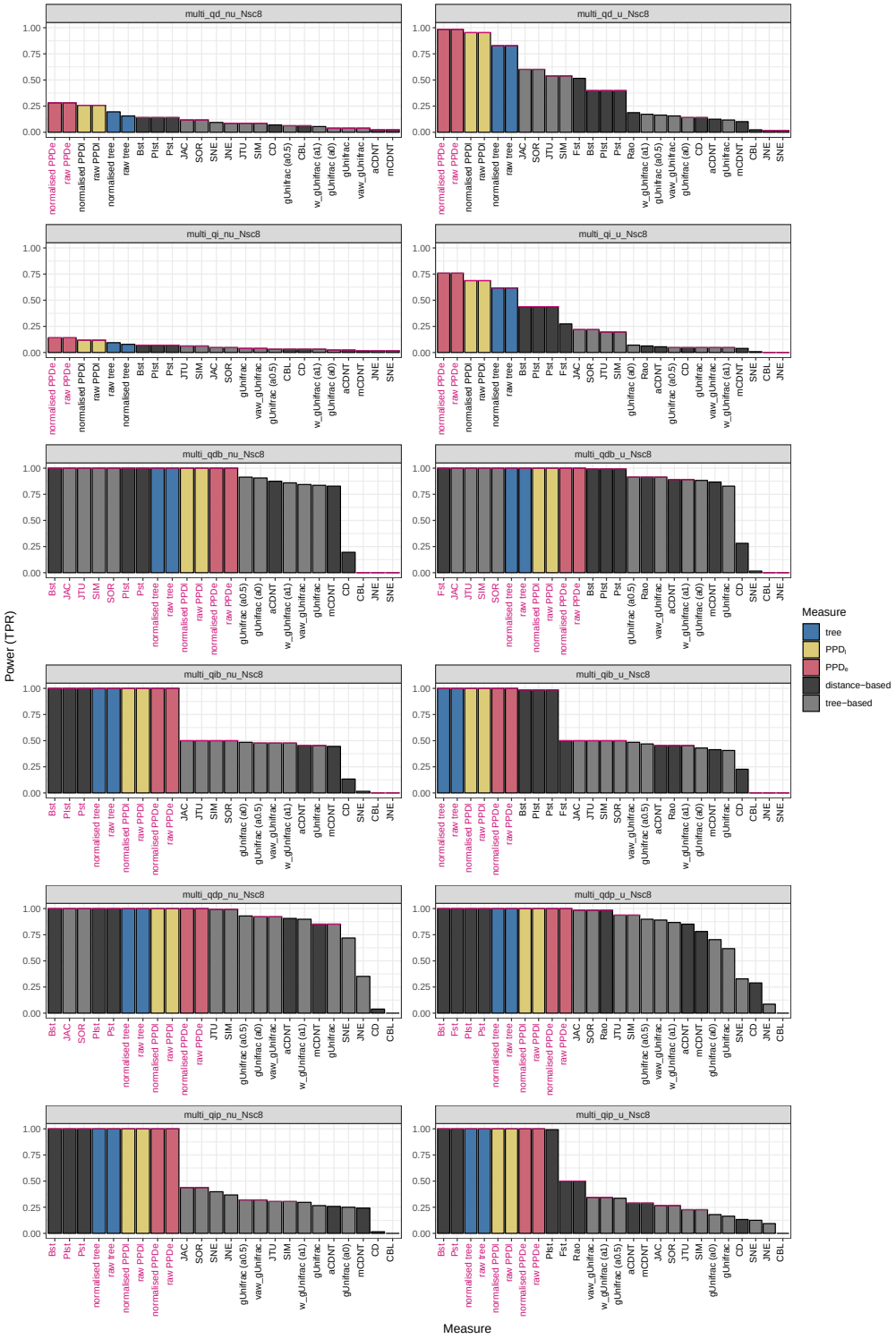


Figure B.77: Bar chart comparing the power of each measure to detect phylogenetic signal in subcommunity structure, in Experiment 5 (for 8 subcommunities): Bars are coloured according to the type of measure, red lines highlight the tops of bars with equal power, and the names of measures with the highest equal values are coloured red.

APPENDIX C

Sampling properties

Table C.1: The effect of subsampling individuals on R^2 values for different measures of diversity and values of q: R^2 values quantify how close the data points are to the fitted regression line.

Measure	Subsample by individual					Subsample by subcommunities				
	50%	**25%**	**12.5%**	**6.25%**	**3.125%**	**50%**	**25%**	**12.5%**	**6.25%**	**3.125%**
${}^{0}\bar{\rho}^{\mathbf{Z}}$	0.759	0.495	0.249	0.101	0.04	0.998	0.995	0.989	0.976	0.959
${}^{1}\bar{\rho}^{\mathbf{Z}}$	0.872	0.669	0.441	0.249	0.121	0.998	0.995	0.988	0.976	0.956
${}^{2}\bar{\rho}^{\mathbf{Z}}$	0.737	0.469	0.273	0.146	0.072	0.981	0.954	0.912	0.857	0.788
${}^{\infty}\bar{\rho}^{\mathbf{Z}}$	0.407	0.178	0.083	0.037	0.015	0.96	0.894	0.795	0.656	0.458
${}^{0}\rho^{\mathbf{Z}}$	0.855	0.57	0.209	0.029	0	0.999	0.995	0.989	0.977	0.965
${}^{1}\rho^{\mathbf{Z}}$	0.872	0.628	0.354	0.163	0.07	0.998	0.994	0.985	0.97	0.946
${}^{2}\rho^{\mathbf{Z}}$	0.727	0.448	0.251	0.127	0.055	0.982	0.956	0.914	0.861	0.796
${}^{\infty}\rho^{\mathbf{Z}}$	0.4	0.177	0.084	0.036	0.012	0.961	0.897	0.799	0.659	0.473
${}^{0}\bar{\beta}^{\mathbf{Z}}$	0.579	0.357	0.282	0.222	0.145	0.999	0.998	0.995	0.989	0.98
${}^{1}\bar{\beta}^{\mathbf{Z}}$	0.91	0.73	0.459	0.259	0.113	0.971	0.949	0.892	0.835	0.768
${}^{2}\bar{\beta}^{\mathbf{Z}}$	0.859	0.582	0.335	0.185	0.107	0.918	0.861	0.752	0.688	0.627
${}^{\infty}\bar{\beta}^{\mathbf{Z}}$	0.485	0.214	0.1	0.048	0.023	0.768	0.54	0.364	0.235	0.144
${}^{0}\beta^{\mathbf{Z}}$	0.826	0.306	0.022	0.001	0.008	0.999	0.995	0.988	0.974	0.958
${}^{1}\beta^{\mathbf{Z}}$	0.966	0.871	0.685	0.398	0.094	0.963	0.962	0.926	0.792	0.671
${}^{2}\beta^{\mathbf{Z}}$	0.954	0.803	0.517	0.231	0.08	0.907	0.894	0.822	0.638	0.524
${}^{\infty}\beta^{\mathbf{Z}}$	0.47	0.2	0.087	0.036	0.014	0.773	0.545	0.367	0.209	0.09
${}^{0}\gamma^{\mathbf{Z}}$	0.627	0.367	0.213	0.119	0.06	0.877	0.741	0.611	0.532	0.466
${}^{1}\gamma^{\mathbf{Z}}$	0.918	0.777	0.571	0.364	0.184	0.99	0.979	0.957	0.939	0.89
${}^{2}\gamma^{\mathbf{Z}}$	0.836	0.625	0.422	0.257	0.123	0.999	0.997	0.992	0.983	0.955
${}^{\infty}\gamma^{\mathbf{Z}}$	0.227	0.139	0.167	0.147	0.102	0.965	0.897	0.741	0.656	0.547
${}^{0}\bar{\alpha}^{\mathbf{Z}}$	0.839	0.69	0.545	0.404	0.302	1	1	1	1	1
${}^{1}\bar{\alpha}^{\mathbf{Z}}$	0.871	0.673	0.435	0.214	0.074	1	1	1	1	1
${}^{2}\bar{\alpha}^{\mathbf{Z}}$	0.877	0.675	0.413	0.174	0.036	1	1	1	1	1
${}^{\infty}\bar{\alpha}^{\mathbf{Z}}$	0.809	0.582	0.336	0.14	0.041	1	1	1	1	1
${}^{0}\alpha^{\mathbf{Z}}$	0.842	0.65	0.441	0.246	0.105	0.999	0.997	0.993	0.985	0.976
${}^{1}\alpha^{\mathbf{Z}}$	0.909	0.755	0.541	0.314	0.138	1	0.999	0.998	0.995	0.993
${}^{2}\alpha^{\mathbf{Z}}$	0.908	0.751	0.531	0.304	0.131	1	0.999	0.999	0.997	0.996
${}^{\infty}\alpha^{\mathbf{Z}}$	0.862	0.691	0.5	0.278	0.122	1	1	0.999	0.997	0.996

Table C.2: The effect of subsampling individuals on AUC values for different measures of diversity and values of q: AUC vaules quantify how well each measure is able to identify the top 5% of the most interesting results (for ${}^{q}\rho_j^{\mathbf{I}}$ and ${}^{q}\bar{\rho}_j^{\mathbf{I}}$ that is the lowest 5%). Perfect accuracy is achieved when AUC = 1 and accuracy is no better than chance when AUC = 0.5.

Measure	Subsample by individual					Subsample by subcommunities				
	50%	**25%**	**12.5%**	**6.25%**	**3.125%**	**50%**	**25%**	**12.5%**	**6.25%**	**3.125%**
${}^{0}\bar{\rho}_j^{\mathbf{I}}$	0.924	0.827	0.684	0.576	0.533	1	0.999	0.998	0.996	0.992
${}^{1}\bar{\rho}_j^{\mathbf{I}}$	0.949	0.875	0.795	0.718	0.661	1	0.999	0.996	0.99	0.977
${}^{2}\bar{\rho}_j^{\mathbf{I}}$	0.945	0.862	0.784	0.716	0.661	0.999	0.997	0.992	0.985	0.972
${}^{\infty}\bar{\rho}_j^{\mathbf{I}}$	0.932	0.836	0.747	0.673	0.623	0.996	0.987	0.975	0.951	0.897
${}^{0}\rho_j^{\mathbf{I}}$	0.954	0.839	0.657	0.504	0.419	1	1	1	0.999	0.997
${}^{1}\rho_j^{\mathbf{I}}$	0.969	0.881	0.763	0.66	0.592	1	0.999	0.998	0.996	0.995
${}^{2}\rho_j^{\mathbf{I}}$	0.953	0.866	0.779	0.711	0.657	0.999	0.998	0.994	0.989	0.985
${}^{\infty}\rho_j^{\mathbf{I}}$	0.936	0.844	0.759	0.687	0.639	0.996	0.989	0.979	0.956	0.897
${}^{0}\bar{\beta}_j^{\mathbf{I}}$	0.967	0.898	0.813	0.754	0.71	1	1	0.999	0.999	0.995
${}^{1}\bar{\beta}_j^{\mathbf{I}}$	0.996	0.987	0.966	0.927	0.862	1	0.999	0.999	0.998	0.997
${}^{2}\bar{\beta}_j^{\mathbf{I}}$	0.978	0.945	0.9	0.845	0.777	0.998	0.991	0.977	0.962	0.959
${}^{\infty}\bar{\beta}_j^{\mathbf{I}}$	0.824	0.721	0.657	0.617	0.586	0.991	0.973	0.941	0.875	0.785
${}^{0}\beta_j^{\mathbf{I}}$	0.995	0.975	0.898	0.742	0.579	1	1	0.999	0.998	0.994
${}^{1}\beta_j^{\mathbf{I}}$	0.993	0.974	0.924	0.842	0.757	1	0.999	0.998	0.997	0.994
${}^{2}\beta_j^{\mathbf{I}}$	0.967	0.923	0.862	0.786	0.716	0.996	0.985	0.964	0.945	0.925
${}^{\infty}\beta_j^{\mathbf{I}}$	0.813	0.706	0.64	0.6	0.578	0.99	0.971	0.929	0.851	0.707
${}^{0}\gamma_j^{\mathbf{I}}$	0.893	0.813	0.747	0.685	0.63	0.992	0.973	0.931	0.864	0.839
${}^{1}\gamma_j^{\mathbf{I}}$	0.993	0.979	0.947	0.889	0.808	1	0.999	0.999	0.996	0.976
${}^{2}\gamma_j^{\mathbf{I}}$	0.991	0.976	0.944	0.891	0.822	1	1	0.999	0.999	0.993
${}^{\infty}\gamma_j^{\mathbf{I}}$	0.998	0.993	0.979	0.952	0.922	1	1	1	1	1
${}^{0}\bar{\alpha}_j^{\mathbf{I}}$	0.978	0.954	0.928	0.893	0.86	1	1	1	1	1
${}^{1}\bar{\alpha}_j^{\mathbf{I}}$	0.982	0.946	0.885	0.795	0.709	1	1	1	1	1
${}^{2}\bar{\alpha}_j^{\mathbf{I}}$	0.982	0.94	0.853	0.725	0.602	1	1	1	1	1
${}^{\infty}\bar{\alpha}_j^{\mathbf{I}}$	0.974	0.923	0.838	0.726	0.621	1	1	1	1	1
${}^{0}\alpha_j^{\mathbf{I}}$	0.977	0.933	0.87	0.793	0.695	1	1	1	1	1
${}^{1}\alpha_j^{\mathbf{I}}$	0.989	0.964	0.911	0.822	0.702	1	1	1	1	1
${}^{2}\alpha_j^{\mathbf{I}}$	0.991	0.969	0.921	0.831	0.7	1	1	1	1	1
${}^{\infty}\alpha_j^{\mathbf{I}}$	0.986	0.959	0.909	0.829	0.709	1	1	1	1	1

C.1 Subsampling individuals

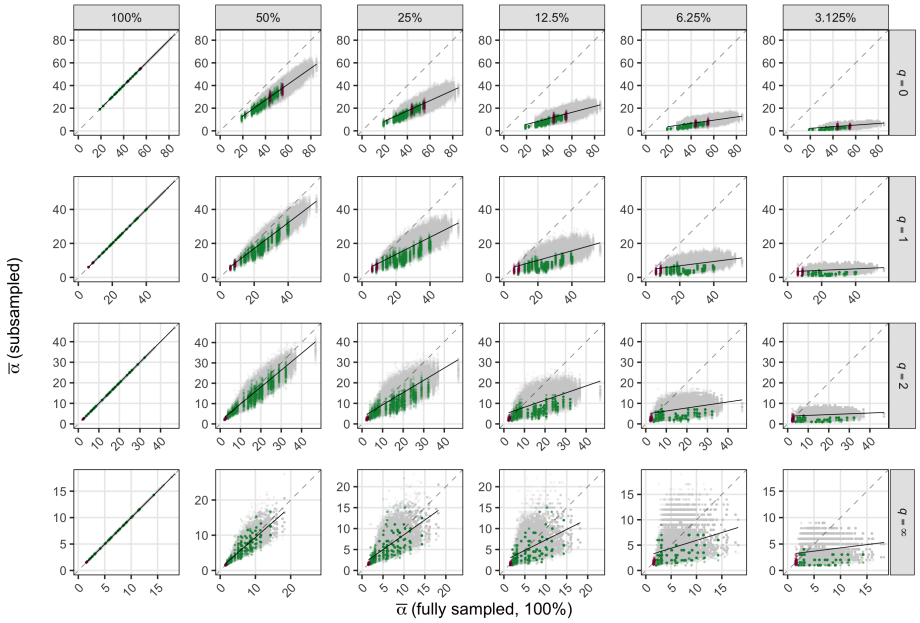


Figure C.1: The effect of subsampling individuals on the magnitude of $\bar{\alpha}$ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m $\times$ 20 m quadrat in the BCI study site ($N = 1250$). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

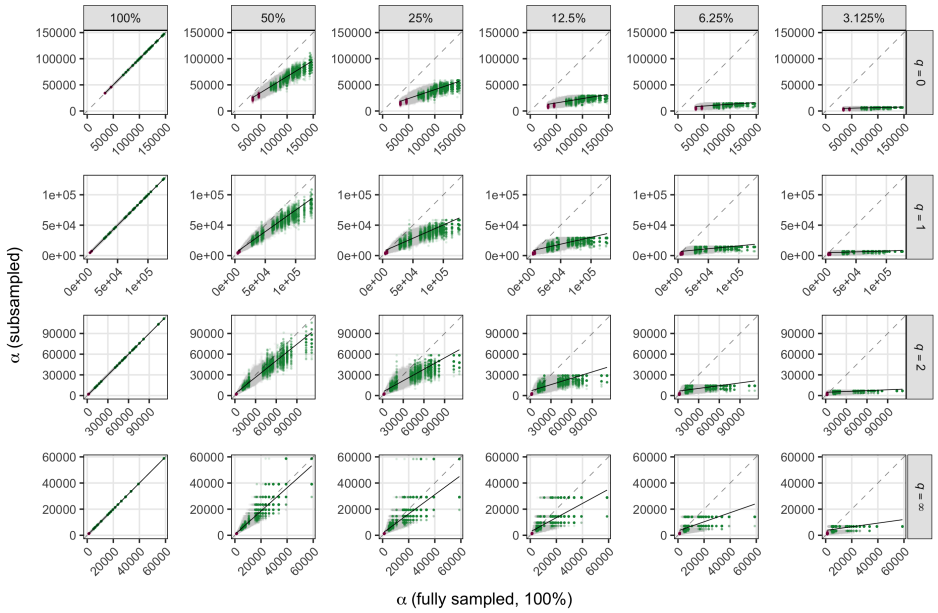


Figure C.2: The effect of subsampling individuals on the magnitude of α diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

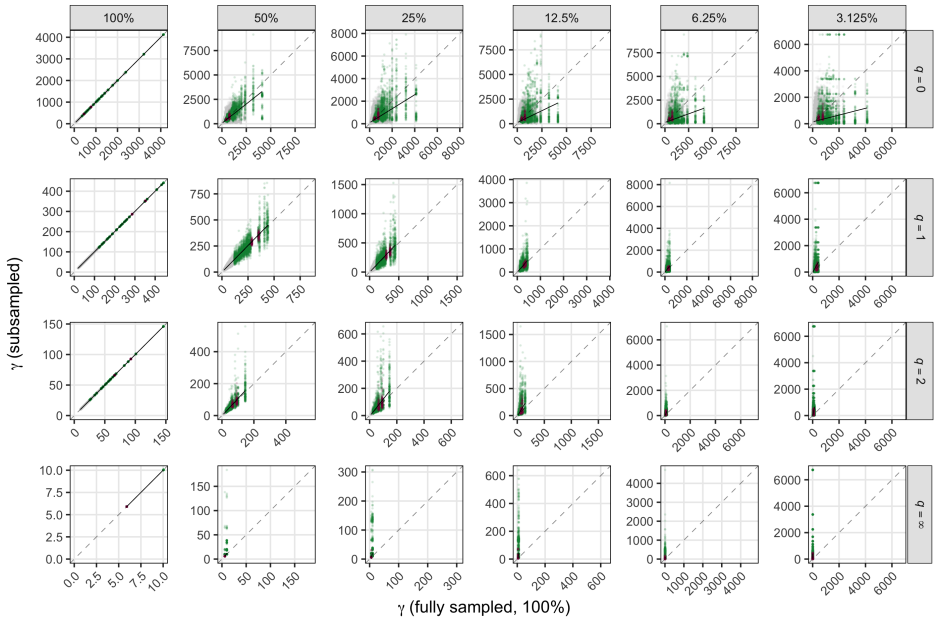


Figure C.3: The effect of subsampling individuals on the magnitude of γ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

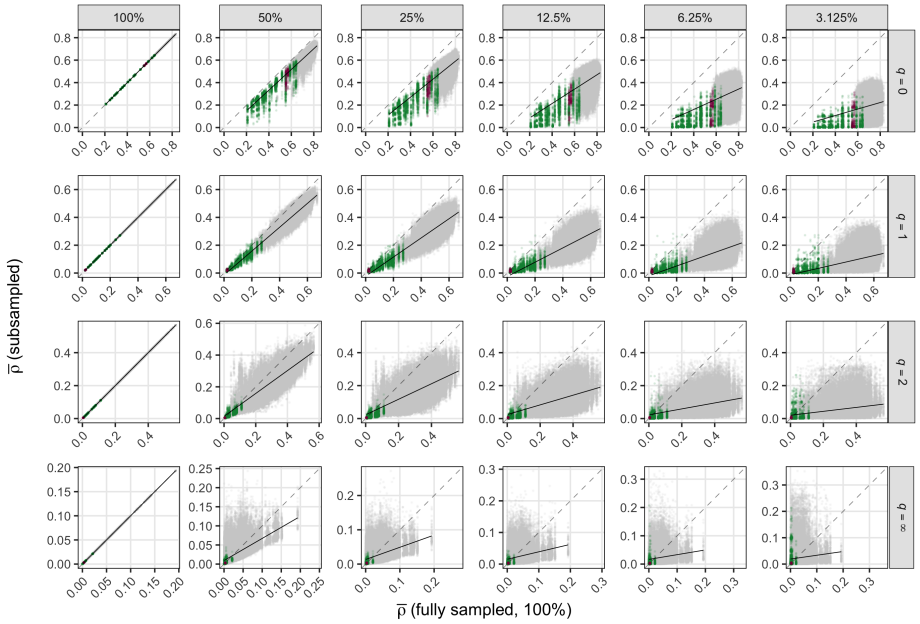


Figure C.4: The effect of subsampling individuals on the magnitude of $\bar{\rho}$ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

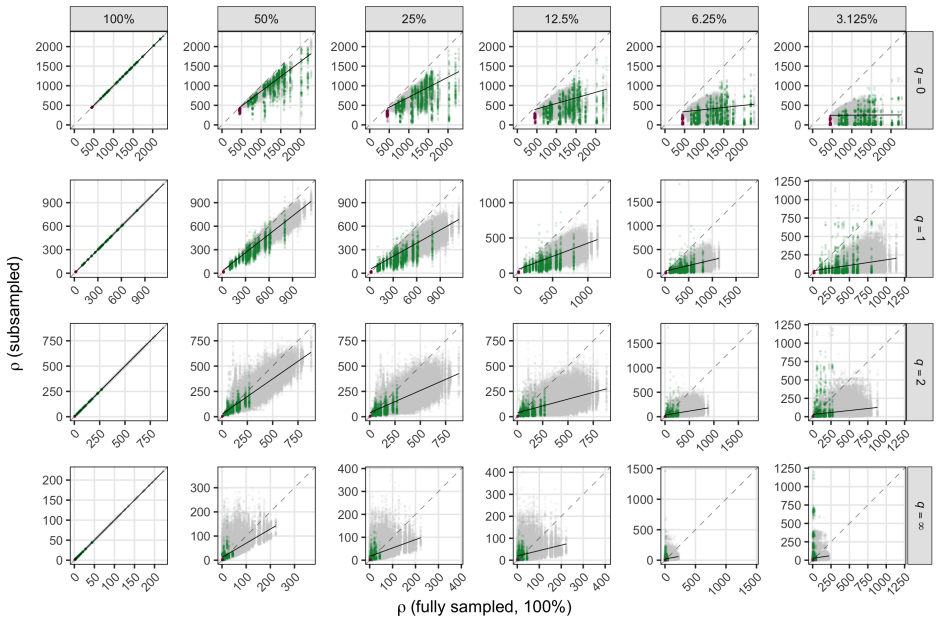


Figure C.5: The effect of subsampling individuals on the magnitude of ρ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

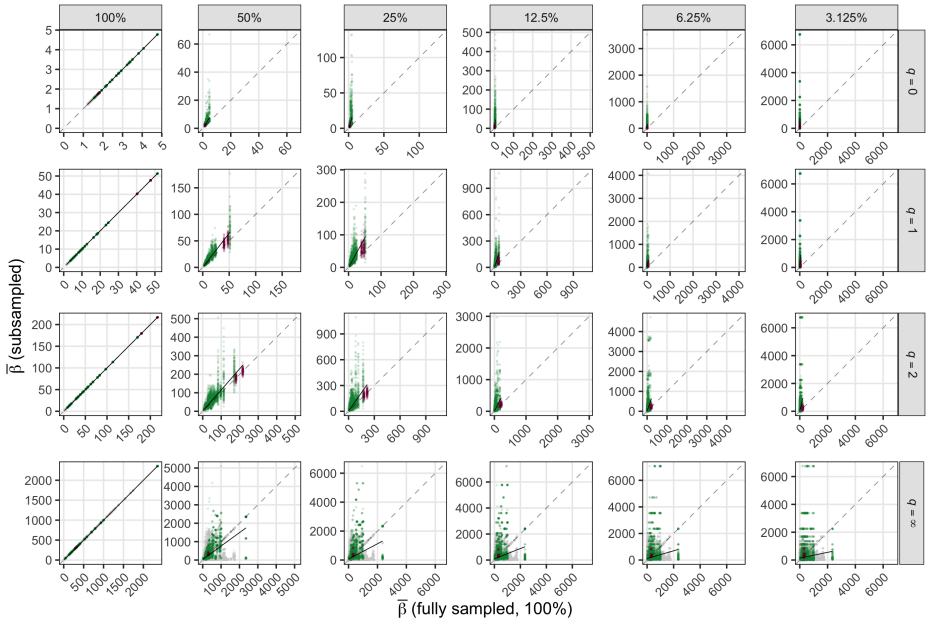


Figure C.6: The effect of subsampling individuals on the magnitude of $\bar{\beta}$ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

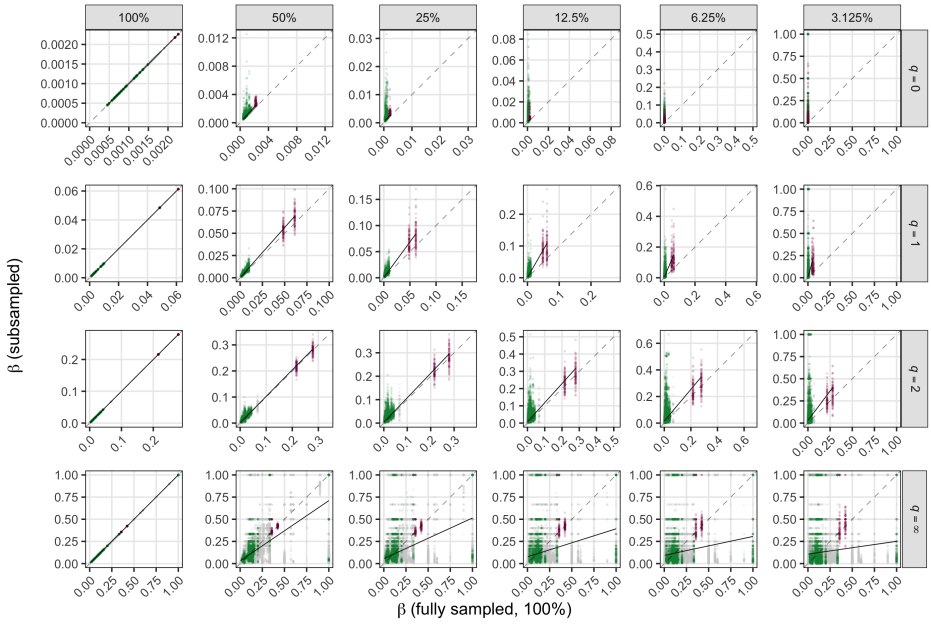


Figure C.7: The effect of subsampling individuals on the magnitude of β diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

C.2 Subsampling subcommunities

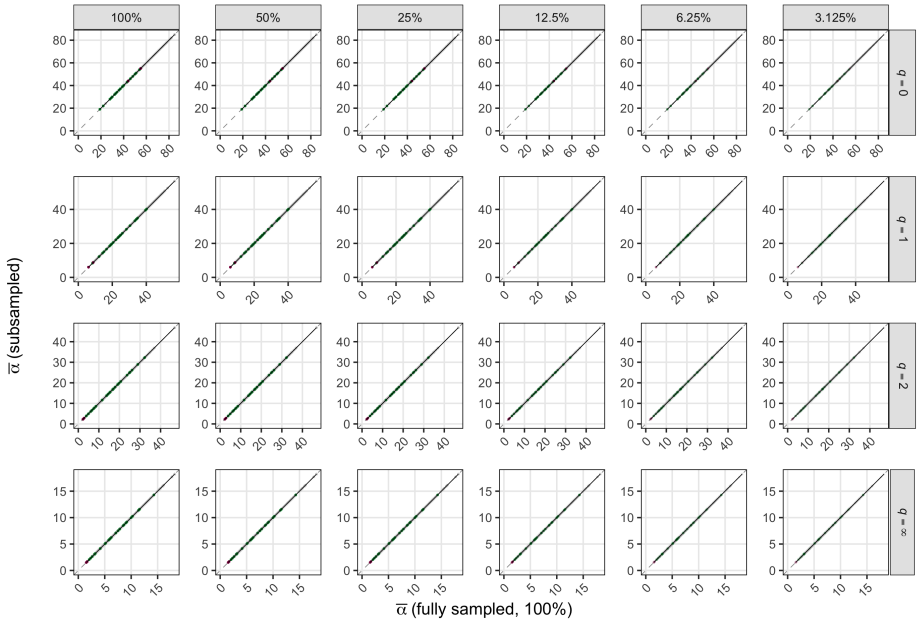


Figure C.8: The effect of subsampling subcommunities on the magnitude of $\bar{\alpha}$ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$ at 100%). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

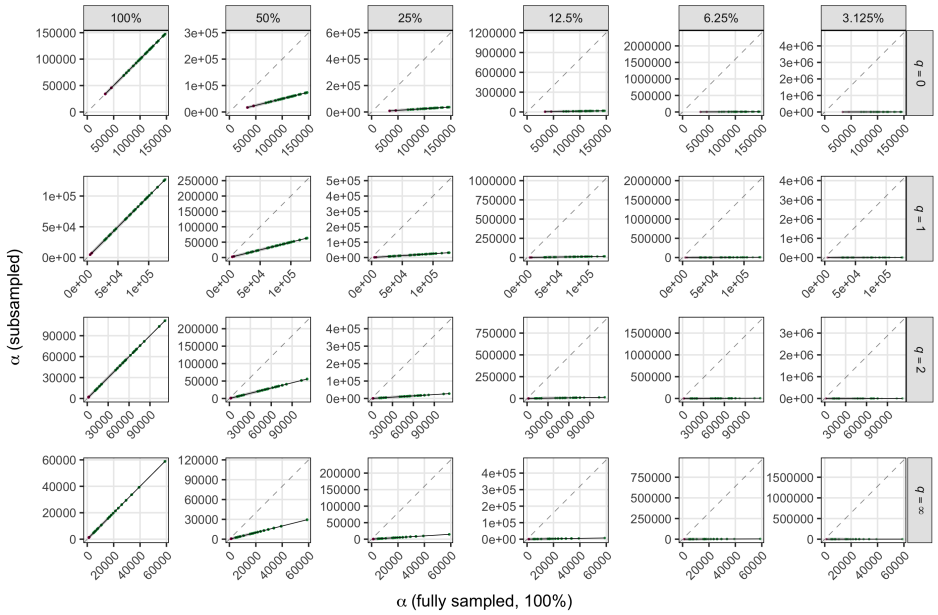


Figure C.9: The effect of subsampling subcommunities on the magnitude of α diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$ at 100%). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

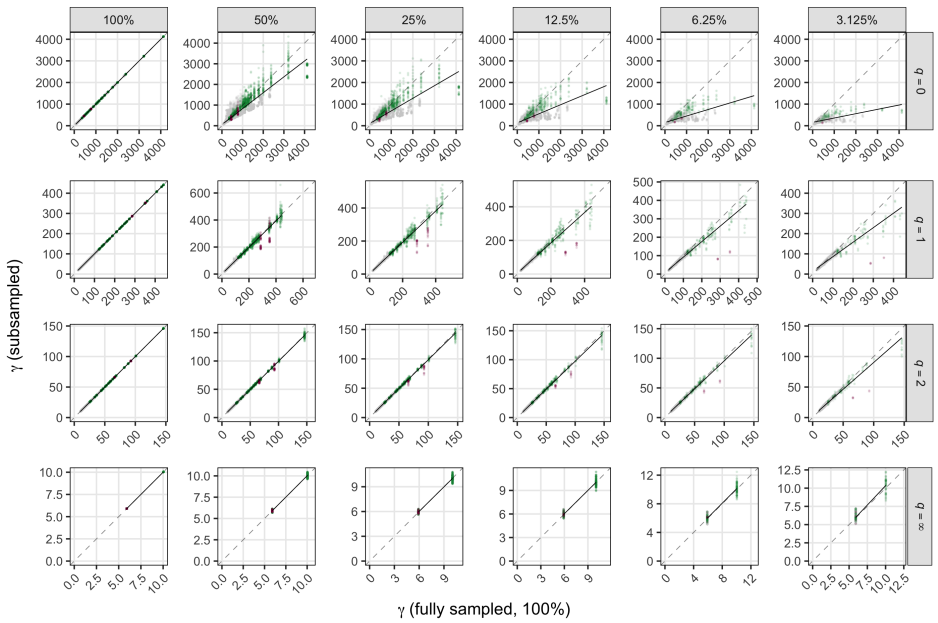


Figure C.10: The effect of subsampling subcommunities on the magnitude of γ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$ at 100%). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

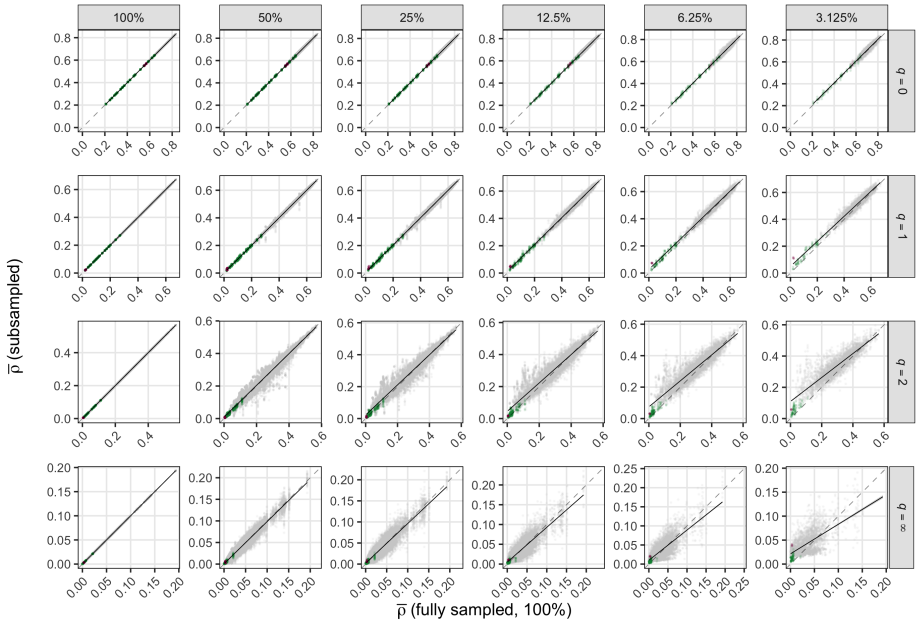


Figure C.11: The effect of subsampling subcommunities on the magnitude of $\bar{\rho}$ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$ at 100%). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

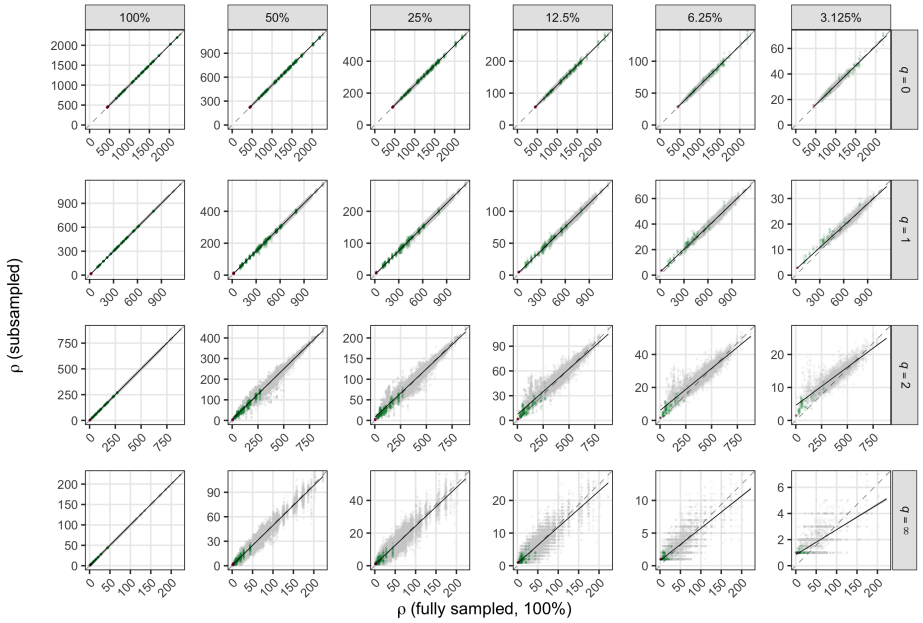


Figure C.12: The effect of subsampling subcommunities on the magnitude of ρ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$ at 100%). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

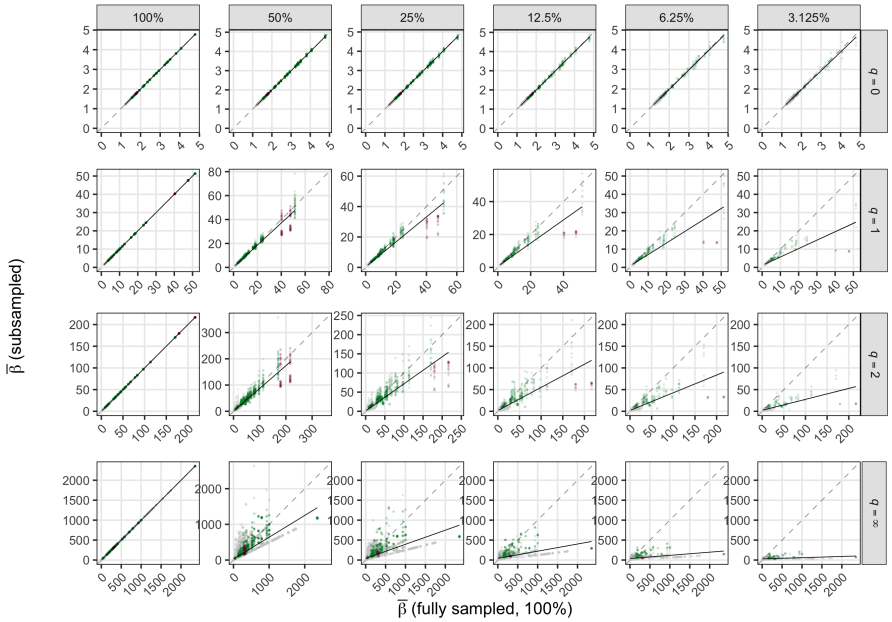


Figure C.13: The effect of subsampling subcommunities on the magnitude of $\bar{\beta}$ diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$ at 100%). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

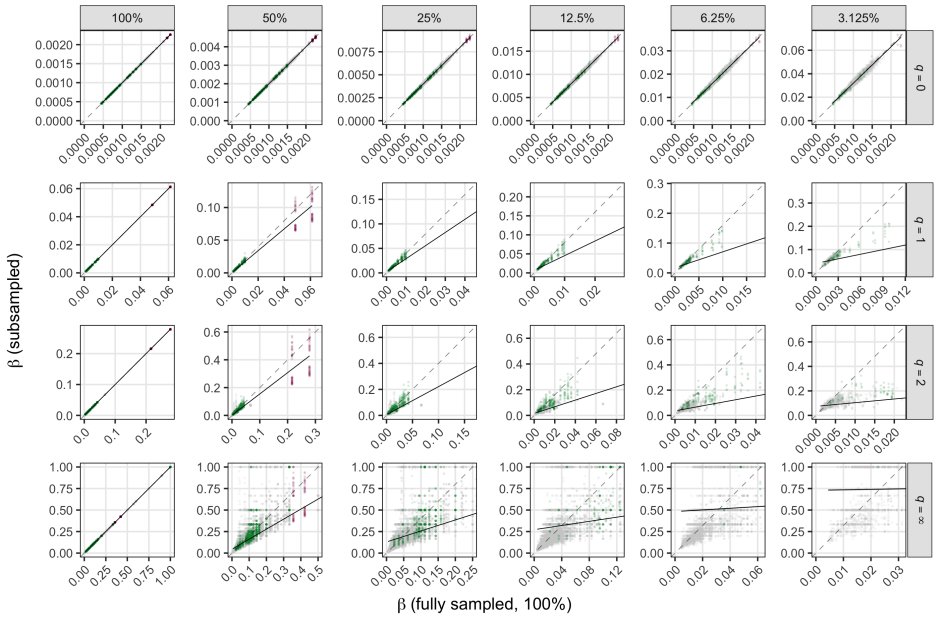


Figure C.14: The effect of subsampling subcommunities on the magnitude of β diversity at different values of q and sampling effort, with highlighted swamp communities: Each point corresponds to a single 20 m × 20 m quadrat in the BCI study site ($N = 1250$ at 100%). Swamp subcommunities are coloured green, the two subcommunities in the top left of the study site are coloured red, and the remaining forest subcommunities are coloured grey. The solid black line shows the linear regression through subsampled values and the grey ribbon (where present) denotes the standard error. Diversity is fully conserved when points follow the dashed grey line.

APPENDIX D

Supplementary case study: Genetic diversity

D.1 Introduction

The following case study mirrors that in Section 4.6, in which phylogenetic diversity measures were used to explore whether or not there was extensive transmission of DT104 between human and animal populations during the epidemics. Here, the same analysis was conducted by investigating the genetic diversity of DT104 isolates directly, using whole genome sequence data.

Acknowledgements: The sequence data used here was kindly provided by Alison Mather.

D.2 Methods

Genetic distance-based methods (analogous to measures of phylogenetic distance-based beta diversity, developed in Section 4.3.3.3) were applied directly to whole genome sequence data to quantify the effective number of epidemics (${}^{q}\bar{B}^{\mathbf{Z}_{GPD_e}}$) over a range of q values, with genetic similarity defined as an exponentially-transformed genetic pairwise distance that scales with k. As k increases, the distance required for sequences to be completely distinct decreases (Table D.1).

The null distribution was simulated at a 95 % confidence interval by randomly relabelling sequences 10000 times, and calculating ${}^{q}\bar{B}^{\mathbf{Z}_{GPD_e}}$ for all values of q.

Table D.1: Methods of calculating pairwise similarity, $Z_{ii'}$, from pairwise genetic distances, $d_{ii'}$: Conversions are shown for both linearly-transformed and exponentially transformed genetic pairwise distances, GPD_l and GPD_e, respectively.

	Transformation	Distance	Similarity
GPD_l	$Z_{ii'} = 1 - (d_{ii'}/kd_{max})$ where $k = 1$	1	0.992
		10	0.925
		100	0.248
	$Z_{ii'} = 1 - (d_{ii'}/seq\_length)$[1]	1	0.999
		10	0.993
		100	0.927
GPD_e	$Z_{ii'} = 2^{-kd_{ii'}}$ where $k = 0.25$	1	0.841
		10	0.177
		100	3×10^{-8}
	$Z_{ii'} = 2^{-kd_{ii'}}$ where $k = 1$	1	0.5
		10	0.001
		100	8×10^{-31}

[1] Nei's (1973) nucleotide diversity: here k is effectively greater than 1, since seq_length (1372 nt) is greater than d_{max} (133 nt).

D.3 Results and discussion

In this case study, the diversity of the isolate samples was examined at a range of sensitivities (by varying k). The calculated effective number of epidemics (${}^{q}\bar{B}^{\mathbf{Z}_{GPD_e}}$) was found to be sensitive to the exact measure of genetic similarity used (Figure D.1b). When $k = 0.25$, results were comparable to those obtained during the phylogenetic study, in that human and animal isolate communities were not well mixed at nearly all scales. However, this was no longer true when $k = 1$. Likewise for GPD_l (linearly-transformed genetic pairwise distance) when $k = 1$ (sensu Nei, 1973).

In conclusion therefore, the epidemics were distinguishable, though the quality of the match to the phylogenetic results was sensitive to how similarity is defined.

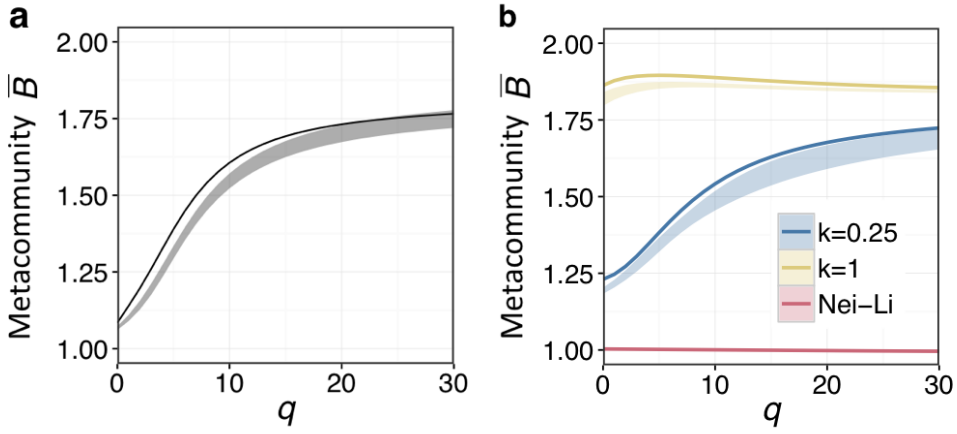


Figure D.1: Plots show comparative plots of phylogenetic and genetic diversity in a representative sample of *Salmonella* DT104: Metacommunity ${}^{q}\bar{B}^{\mathbf{Z}}$ is calculated for $0 \leqslant q \leqslant 30$, yielding (a) the effective number of phylogenies (${}^{q}\bar{B}^{\mathbf{Z}_{tree}}$), and (b) the effective number of distinct genotypes for $Z_{ii'} = 2^{-0.25d}$ (${}^{q}\bar{B}^{\mathbf{Z}_{GPD_e}}$, blue), $Z_{ii'} = 2^{-d}$ (${}^{q}\bar{B}^{\mathbf{Z}_{GPD_e}}$, yellow), and a variant of π (${}^{q}\bar{B}^{\mathbf{Z}_{GPD_l}}$, red). Shaded ribbons simulate the null distribution within a 95 % confidence interval, where $p < 0.05$ for $q > 25$ and $q > 7$, respectively.

APPENDIX E

Supplementary case study: Using diversity-based methods to estimate true epidemic sizes from partially observed outbreaks

E.1 Abstract

Foot and mouth disease virus (FMDV) is a highly contagious viral disease of cloven-hoofed animals (pigs, sheep, cows, *etc.*) that results in a persistent infection and is characterised by high fever, followed by lesions in the tongue, lips, and feet. Although most animals do recover, productivity is greatly reduced, affecting food security and agricultural-based economies. Accurate assessment of epidemic surveillance data is therefore vital, not only to inform a basic understanding of viral demographics, but also to ascertain the efficacy of viral control methods, make outbreak projections, and predict epidemic spread.

Phylodynamic studies of viral evolution, and population genetic models more generally, provide valuable information reflecting viral transmission dynamics. These methods typically use coalescent approaches to provide insight into epidemiological processes such as changes in effective population size, which can be used as an estimate of the true size of an outbreak. However, incomplete surveillance data may cause a mismatch between estimates of incidence observed empirically and those reconstructed from phylodynamic models, with the calculated effective population size often much smaller than the actual size of an epidemic. This means that as sampling resolution decreases, estimates of the true scale of an epidemic worsen just as such estimators become more important.

In this case study, I investigate the potential for new, simpler methods – based on Reeve et al.'s (2016) measures of similarity-sensitive diversity – to examine viral genetic and phylogenetic diversity during the UK 2001 FMDV (foot-and-mouth-disease virus) outbreak and infer the outbreak size from subsampled data more accurately than existing coalescent-based metrics.

E.2 Introduction

This case study builds on previous work by Antonello Di Nardo (2016), who tackled the problem of low reporting rates and non-representative sampling in the control and management of infectious disease. The focus of his work was to define a relationship between the effective population size (the number of individuals in an idealised assemblage with the same value of some epigenetic trait as the population of interest), N_e, estimated by methods of phylodynamic inference, and the infected premises (IPs) observed empirically, N. With data from the exhaustively-sampled UK 2001 FMDV epidemic (whole genome sequences and simulated sequences generated from transmission tree mutations), he generated models to describe the 2001 UK FMDV epidemic. From here, phylogenies were generated and coalescent approaches were used to provide insight into changes in effective population size. Specifically, Bayesian skyline methods were used to reconstruct demographic changes in viral populations, and investigate the effects of subsampling during different stages in the disease outbreak.

Figure E.1 (adapted from Di Nardo, 2016) shows how decreased sampling rate affects how well epidemic demography can be reconstructed from the Bayesian Skyline plot-derived effective population size. These plots show infection prevalence over time (the number of premises still infected each day). The black curve denotes the true infection prevalence of the entire simulated dataset, while the red curves estimate infection prevalence using Bayesian skyline methods at decreasing sampling proportions. These results show that coalescent-based estimators of effective population size break down when sample size is low. In response to this, here, new methods of estimating population size are proposed, based on Reeve et al.'s (2016) framework of similarity-sensitive diversity.

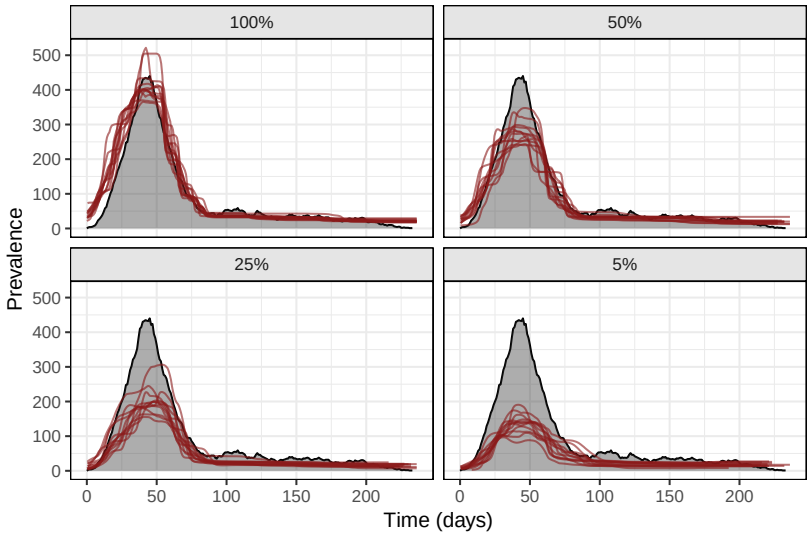


Figure E.1: The effect of decreasing sampling rate on the infection prevalence over time: The black curve denotes the true infection prevalence from complete data and the red curves estimate infection prevalence using Bayesian skyline methods.

Acknowledgements: The R code used to generate FMDV outbreaks was kindly provided by Antonello Di Nardo.

E.3 Methods

Code was provided by Antonello Di Nardo (2016) to simulate the transmission of FMDV in an outbreak. This code was adapted to provide greater variability in demographic structure to test the robustness of the new diversity-based methods (Figure E.2a). For each daughter IP in the outbreak, whole genome sequences were generated alongside time of exposure ($expT$), latency period ($latD$), time of infectiousness ($infT$), and time of removal ($remT$).

The aim of this work was to provide a method of predicting the size of an outbreak from incomplete data, reflecting the fact that people want to know the size of an outbreak while it is still ongoing. To simulate this, outbreaks were truncated by sampling the viral population for a random proportion of time (Figure E.2b).

For simplicity and to avoid errors in phylogenetic inference (also because it is computationally intensive to generate the phylogenies, and previous results in this chapter have shown that genetic and phylogenetic diversities tend to give comparable results), viral genetic diversity

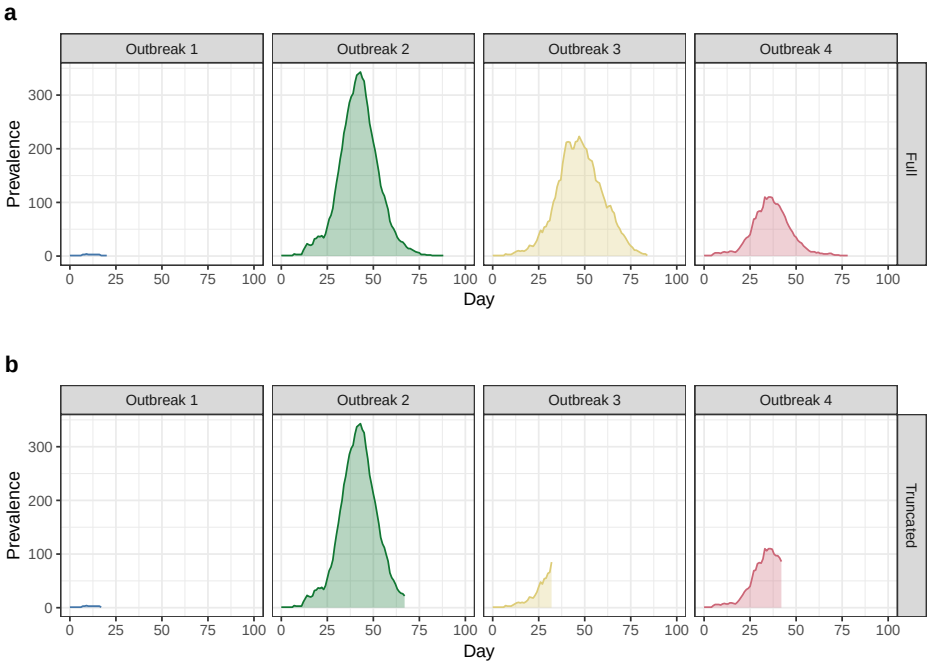


Figure E.2: Plots of viral incidence against time for four randomly generated, *fully sampled* FMDV outbreaks.: incidence is calculated for (a) the entire length of outbreak, and (b) the length of a partial outbreak. The outbreak is sampled at 100% (shaded area).

was assessed directly. To determine the effective number of sequences (or equivalently, the effective number of infected premises) in each outbreak, metacommunity ${}^{q}G^{GPD_l}$ and ${}^{q}G^{GPD_e}$ were calculated for $q \in \{0, 1, 2, \infty\}$ and different sampling rates $\in \{100\%, 25\%, 10\%, 1\%\}$.

Metacommunity diversity was plotted against outbreak size for different values of q and sampling rates. The gradient of the regression line between diversity and true outbreak size produced a scaling factor, which could then be used to recover the true size of the outbreak from any particular diversity measure calculated from a sample.

E.4 Results and discussion

First, the diversity-based estimate of effective population size (y-axis) was compared to the true sample size of truncated outbreaks (x-axis) for full outbreaks (Figure E.3a). It was found that the effective population size was correlated with the true number of infected hosts, where the true outbreak size was substantially, but *consistently*, underestimated (Figure E.4). This relationship was maintained when sampling rate was dropped to 25% of sequences randomly sampled, and even 10%.

Similar results were obtained when assessing truncated outbreaks – partially observed outbreaks ranging from less than 5% observed to fully sampled (Figure E.3b). It was found that the estimated population size was a good fit to the true size of the epidemics, even when less than 5% of the epidemic was observed (Figure E.5).

In conclusion then, these methods accurately predict the true size of the epidemic, even from outbreaks that are only partially observed or that are still ongoing. Given the consistency of these results, these measures – essentially summary statistics that can describe the variability present in a population – appear to be a promising approach to an as yet unsolved problem in practical disease control and could potentially be further improved by combination with other simple summary statistics to fit a more complex model. In this way, diversity-based methods might therefore be used in the future to better inform a robust and general approach to viral demographic inference.

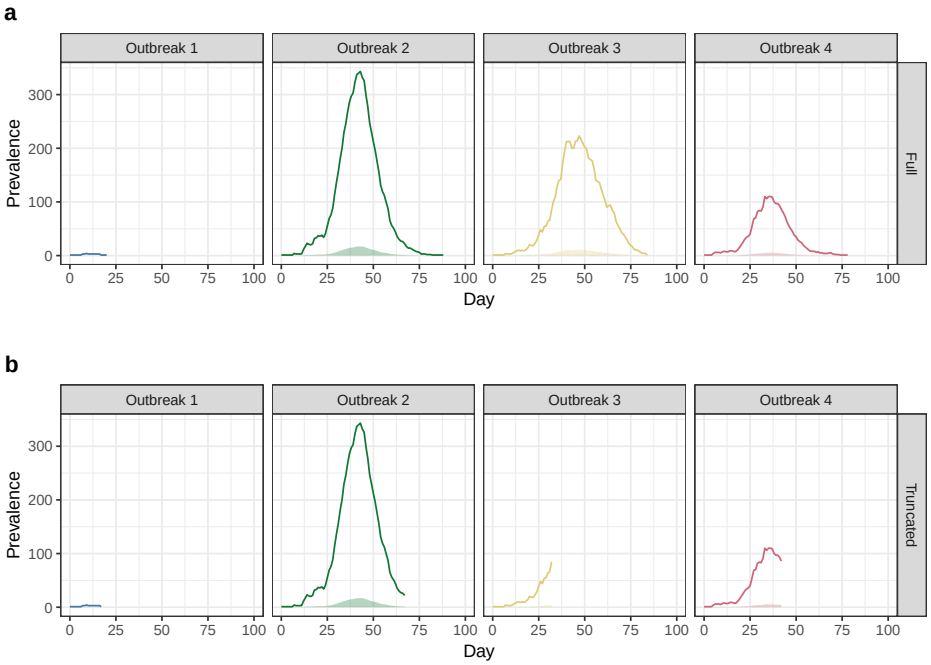


Figure E.3: Plots of viral prevalence against time illustrating *partially sampled* FMDV outbreaks: Prevalence is calculated for (a) the entire length of outbreak, and (b) the length of a partial outbreak. The outbreak is sampled at 5% (shaded area).

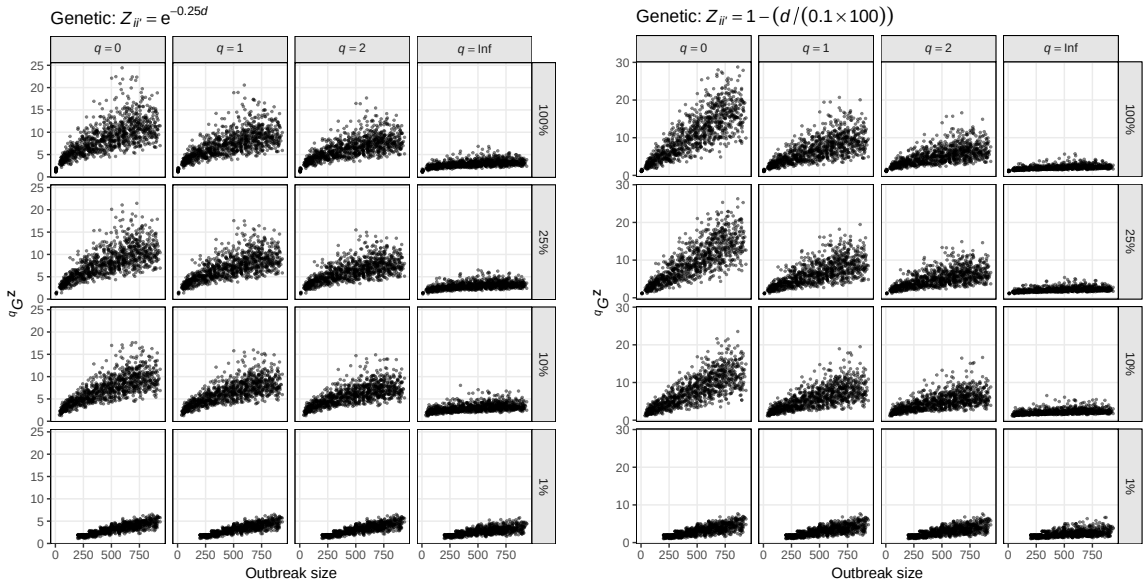


Figure E.4: Plots of metacommunity ${}^qG^Z$ against outbreak size for *full outbreaks*: left - exponential similarity, $Z_{ii'} = e^{-kd}$; right - fixed Nei-Li similarity, $Z_{ii'} = 1-(d/100k)$. Each point describes a different outbreak and each plot displays results for a particular value of q and sampling rate.

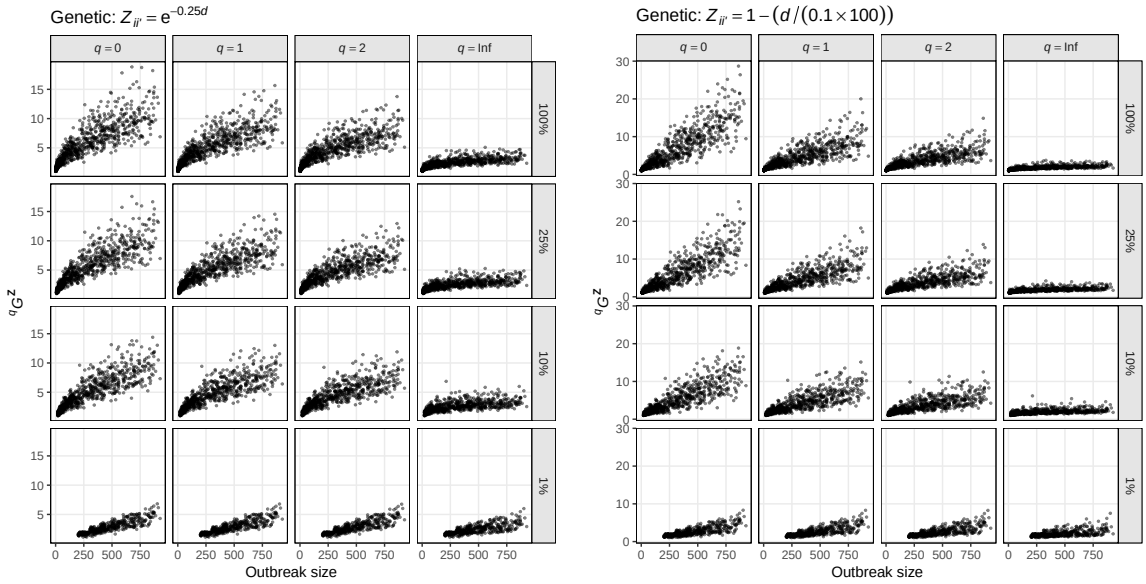


Figure E.5: Plots of metacommunity $^qG^{\mathbf{Z}}$ against outbreak size for *truncated outbreaks*: left - exponential similarity, $Z_{ii'} = e^{-kd}$; right - fixed Nei-Li similarity, $Z_{ii'} = 1 - (d/100k)$. Each point describes a different outbreak and each plot displays results for a particular value of q and sampling rate.

BIBLIOGRAPHY

Bibliography

Adams, D.C., DiBitetti, M.S., Janson, C.H., Slobodkin, L.B. & Valenzuela, N. (1997) An "audience effect" for ecological terminology: Use and misuse of jargon. Oikos **80**, 632–636.

Adelman, M.A. (1969) Comment on the "H" concentration measure as a numbers-equivalent. The Review of Economics and Statistics **51**, 99–101.

Alanis, A.J. (2005) Resistance to antibiotics: Are we in the post-antibiotic era? Archives of Medical Research **36**, 697–705.

Alatalo, R.V. (1981) Problems in the measurement of evenness in ecology. Oikos **37**, 199–204.

Alcaine, S.D., Warnick, L.D. & Wiedmann, M. (2007) Antimicrobial resistance in nontyphoidal Salmonella. Journal of Food Protection **70**, 780–790.

Allan, J. (1975) Components of diversity. Oecologia **18**, 359–367.

Allen, B., Kon, M. & BarYam, Y. (2009) A new phylogenetic diversity measure generalizing the Shannon index and its application to phyllostomid bats. The American Naturalist **174**, 236–243.

Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J., Stegen, J.C. & Swenson, N.G. (2011) Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. Ecology Letters **14**, 19–28.

Andersson, D.I. & Hughes, D. (2010) Antibiotic resistance and its cost: Is it possible to reverse resistance? Nature reviews. Microbiology **8**, 260–271.

Andersson, D.I. & Hughes, D. (2011) Persistence of antibiotic resistance in bacterial populations. FEMS Microbiology Reviews **35**, 901–911.

Bacaro, G., Ricotta, C. & Mazzoleni, S. (2007) Measuring beta-diversity from taxonomic similarity. Journal of Vegetation Science **18**, 793–798.

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

Baselga, A. (2010) Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography **19**, 134–143.

Baselga, A. (2012) The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. Global Ecology and Biogeography **21**, 1223–1232.

Baselga, A. & Orme, C.D.L. (2012) Betapart: An R package for the study of beta diversity. Methods in Ecology and Evolution **3**, 808–812.

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

Bennett, P.M. (2008) Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. British Journal of Pharmacology **153 Suppl**, S347–S357.

Berger, W.H. & Parker, F. (1970) Diversity of planktonic Foraminifera in deep-sea sediments. Science **168**, 1345–1347.

Biro, P. (2013) Are most samples of animals systematically biased? Consistent individual trait differences bias samples despite random sampling. Oecologia **171**, 339–345.

Boltzmann, L. (1866) Über die bedeutung des zweiten hauptsatzes der wärmetheorie.

Boyd, D., Peters, G.A., Cloeckaert, A., Boumedine, K.S., Chaslus-Dancla, E., Imberechts, H. & Mulvey, M.R. (2001) Complete nucleotide sequence of a 43-kilobase genomic island associated with the multidrug resistance region of Salmonella enterica serovar typhimurium DT104 and its identification in phage type DT120 and serovar agona. American Society for Microbiology **183**, 5725–5732.

Boyd, D.a., Peters, G.a., Ng, L.K. & Mulvey, M.R. (2000) Partial characterization of a genomic island associated with the multidrug resistance region of Salmonella enterica Typhymurium DT104. FEMS Microbiology Letters **189**, 285–291.

Bray, J. & Curtis, J. (1957) An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs **27**, 325–349.

Bryant, J.A., Lamanna, C., Morlon, H., Kerkhoff, A.J., Enquist, B.J. & Green, J.L. (2008) Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. Proceedings of the National Academy of Sciences of the United States of America **105**, 11505–11.

Cadotte, M.W., Carscadden, K. & Mirotchnick, N. (2011) Beyond species: Functional diversity and the maintenance of ecological processes and services. Journal of Applied Ecology **48**, 1079–1087.

Carattoli, A. (2003) Plasmid-mediated antimicrobial resistance in Salmonella enterica. Current Issues Molecular Biology **5**, 113–122.

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

Carter, A.J., Heinsohn, R., Goldizen, A.W. & Biro, P.A. (2012) Boldness, trappability and sampling bias in wild lizards. Animal Behaviour **83**, 1051–1058.

Chamberlain, S., Carvalheiro, L., Elle, E. & Vamosi, J.C. (2014) Running headline: Tree shape and networks Phylogenetic tree shape and the structure of mutualistic networks. Journal of Ecology **102**, 1234–1243.

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

Chao, A., Chiu, C.H. & Jost, L. (2010) Phylogenetic diversity measures based on Hill numbers. Philosophical Transactions of the Royal Society B: Biological Sciences **365**, 3599–3609.

Chao, A., Chiu, C.H.H., Hsieh, T.C. & Inouye, B.D. (2012) Proposing a resolution to debates on diversity partitioning. Ecology **93**, 2037–51.

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

Chave, J., Chust, G. & Thébaud, C. (2007) The importance of phylogenetic structure in biodiversity studies. Scaling biodiversity (eds. D. Storch, P. Marquet & J. Braun), pp. 151–167, Cambridge University Press, Cambridge.

Chen, J. (2012) GUniFrac: Generalized UniFrac distances.

Chen, J., Bittinger, K., Charlson, E.S., Hoffmann, C., Lewis, J., Wu, G.D., Collman, R.G., Bushman, F.D. & Li, H. (2012) Associating microbiome composition with environmental covariates using generalized UniFrac distances. Bioinformatics **28**, 2106–2113.

Clarke, K.R. & Warwick, R.M. (1998) A taxonomic distinctness index and its statistical properties. Journal of Applied Ecology **35**, 523–531.

Colwell, R.K. & Coddington, J.A. (1994) Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society B **345**, 101–118.

Condit, R. (1998) Tropical forest census plots. Springer.

Condit, R., Aguilar, S., R., P., Lao, S., Hubbell, S.P. & Foster, R.B. (2017a) BCI 50-ha Plot Taxonomy as of 2017.

Condit, R., Chisholm, R.A. & Hubbell, S.P. (2012a) Thirty years of forest census at Barro Colorado and the importance of immigration in maintaining diversity. PLoS ONE **7**, 1–6.

Condit, R., Hubbell, S.P. & Foster, R.B. (1996) Changes in tree species abundance in a Neotropical forest: Impact of climate change. Journal of Tropical Eco **12**, 231–256.

Condit, R., Lao, S., Pérez, R., Dolins, S.B., Foster, R. & S Hubbell (2012b) Barro Colorado Forest Census Plot Data (Version 2012).

Condit, R., Pérez, R., Lao, S., Aguilar, S. & Hubbell, S.P. (2017b) Demographic trends and climate over 35 years in the Barro Colorado 50 ha plot. Forest Ecosystems **4**, 17.

Culmsee, H. & Leuschner, C. (2013) Consistent patterns of elevational change in tree taxonomic and phylogenetic diversity across Malesian mountain forests. Journal of Biogeography **40**, 1997–2010.

Dalton, H. (1920) The measurement of the inequality of incomes. The Economic Journal .

De Bello, F., Lavergne, S., Meynard, C.N., Lepš, J. & Thuiller, W. (2010) The partitioning of diversity: Showing Theseus a way out of the labyrinth. Journal of Vegetation Science **21**, 992–1000.

Desrochers, R.E. & Anand, M. (2003) The use of taxonomic diversity indices in the assessment of perturbed community recovery. Advances in Ecological Sciences **18**, 111–120.

Di Nardo, A. (2016) Phylodynamic modelling of foot-and-mouth disease virus sequence data. Ph.D. thesis.

Dice, L.R.L. (1945) Measures of the amount of ecologic association between species. Ecology **26**, 297–302.

Doss, S.a. (1994) Chromosomally-mediated antibiotic resistance and virulence. Journal of Medical Microbiology **40**, 305–306.

Duvick, D. (1984) Genetic diversity in major farm crops on the farm and in reserve. Economic Botany **38**, 161–178.

Eastman, J.M., Paine, C.E.T. & Hardy, O.J. (2011) spacodiR: Structuring of phylogenetic diversity in ecological communities. Bioinformatics **27**, 2437–2438.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. Biological Conservation **61**, 1–10.

Feeley, K.J., Davies, S.J., Perez, R. & Hubbell, S.P. (2011) Directional changes in the species composition of a tropical forest. Ecology **92**, 871–882.

Ferrier, S., Manion, G., Elith, J. & Richardson, K. (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions **13**, 252–264.

Fisher, R.A., Corbet, A.S. & Williams, C.B. (1943) The number of animals in a random sample of an animal population. Journal of Animal Ecology **12**, 42–58.

Forest, F., Grenyer, R., Rouget, M., Davies, T.J., Cowling, R.M., Faith, D.P., Balmford, A., Manning, J.C., Proche, S., van der Bank, M., Reeves, G., Hedderson, T.a.J. & Savolainen, V. (2007) Preserving the evolutionary potential of floras in biodiversity hotspots. Nature **445**, 757–60.

Ganeshaiah, K.N., Chandrashekara, K. & Kumar, A.R.V. (1997) Avalanche index: A new measure of biodiversity based on biological heterogeneity of the communities. Current Science **73**, 128–133.

Gini, C. (1912) Contributo allo Studio delle Distribuzioni e delle Relazioni Statistich. Variabilita e Mutuabilita, C. Cuppini, Bologna.

Goberna, M. & Verdú, M. (2016) Predicting microbial traits with phylogenies. The ISME journal **10**, 959–967.

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

Greenberg, J. (1956) The measurement of linguistic diversity. Language **32**, 109–115.

Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R., Finlay, E.K., Jianlin, H., Groeneveld, E. & Weigend, S. (2010) Genetic diversity in farm animals - A review. Animal Genetics **41 Suppl 1**, 6–31.

Hajjar, R., Jarvis, D.I. & Gemmill-Herren, B. (2008) The utility of crop genetic diversity in maintaining ecosystem services. Agriculture, Ecosystems & Environment **123**, 261–270.

Hardy, O.J. & Jost, L. (2008) Interpreting and estimating measures of community phylogenetic structuring. Journal of Ecology **96**, 849–852.

Hardy, O.J. & Senterre, B. (2007) Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. Journal of Ecology **95**, 493–506.

Harms, K.E., Condit, R., Hubbell, S.P. & Foster, R.B. (2001) Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. Journal of Ecology **89**, 947–959.

Hart, P.E. (1971) Entropy and other measures of concentration. Journal of the Royal Statistical Society: Series A (General) **134**, 73–85.

Heal, G., Walker, B., Levin, S., Arrow, K., Dasgupta, P., Daily, G., Ehrlich, P., Maler, K.G., Kautsky, N., Lubchenco, J., Schneider, S. & Starrett, D. (2004) Genetic diversity and interdependent crop choices in agriculture. Resource and Energy Economics **26**, 175–184.

Hefley, T.J., Tyre, A.J., Baasch, D.M. & Blankenship, E.E. (2013) Nondetection sampling bias in marked presence-only data. Ecology and Evolution **3**, 5225–5236.

Heip, C. (1974) A new index measuring evenness. Journal of Marine Biological Association of the United Kingdom **54**, 555–557.

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

Help, C.H.R., Herman, P.M.J., Soetaert, K., Heip, C., Herman, P.M.J., Soetaert, H. & Soetaert, K. (1998) Indices of diversity and evenness. Océanis **24**, 61–87.

Hill, M.O. (1973) Diversity and evenness: A unifying notation and its consequences. Ecology **54**, 427–432.

Hirschman, A.O. (1964) The paternity of an index. The American Economic Review **54**, 761.

Horowitz, I. (1970) Employment concentration in the common market: An entropy approach. Journal of the Royal Statistical Society: Series A (General) **133**, 463–479.

Hubbell, S.P., Condit, R. & Foster, R.B. (1990) Presence and absence of density dependence in a neotropical tree community. Philosophical Transactions of the Royal Society B - Biological Sciences **330**, 269–281.

Hubbell, S.P. & Foster, R.B. (1992) Short-term dynamics of a neotropical forest: Why ecological research matters to tropical conservation and management. Oikos **63**, 48–61.

Hubbell, S.P., Foster, R.B., O'Brien, S.T., Harms, K.E., Condit, R., Wechsler, B., Wright, S.J. & DeLao, L. (1999) Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest.

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

Ives, A.R. & Helmus, M.R. (2010) Phylogenetic metrics of community similarity. The American Naturalist **176**, E128–E142.

Jaccard, P. (1901) Étude comparative de la distribution florale dans une portion des Alpes et du Jura. Bulletin de la Société Vaudoise des Sciences Naturelles **37**, 547–579.

Jaccard, P. (1912) The distribution of the flora in the alpine zone. The New Phytologist **XI**, 37–50.

Jost, L. (2006) Entropy and diversity. Oikos **2**, 363–375.

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

Junge, K. (1994) Diversity of ideas about diversity measurement. Scandinavian Journal of Psychology **35**, 16–26.

Jurasinski, G., Retzer, V. & Beierkuhnlein, C. (2009) Inventory, differentiation, and proportional diversity: A consistent terminology for quantifying species diversity. Oecologia **159**, 15–26.

Kanagaraj, R., Wiegand, T., Comita, L.S. & Huth, A. (2011) Tropical tree species assemblages in topographical habitats change in time and with life stage. Journal of Ecology **99**, 1441–1452.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. (2010) Picante: R tools for integrating phylogenies and ecology. Bioinformatics **26**, 1463–1464.

Kembel, S.W. & Hubbell, S.P. (2006) The phylogenetic structure of a neotropical forest tree community. Ecology **87**, 86–99.

Kohn, D.D. & Walsh, D.M. (1994) Plant species richness - The effect of island size and habitat diversity. Journal of Ecology **82**, 367–377.

Koleff, P., Gaston, K.J. & Lennon, J.J. (2003) Measuring beta diversity for presence-absence data. Journal of Animal Ecology **72**, 367–382.

Kullback, S. & Leibler, R. (1951) On information and sufficiency. The Annals of Mathematical Statistics .

Lande, R. (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. Oikos **76**, 5–13.

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

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: A framework for multi-scale community ecology. Ecology Letters **7**, 601–613.

Leinster, T. & Cobbold, C. (2012) Measuring diversity: The importance of species similarity. Ecology **93**, 477–489.

Leonard, R.D. & Jones, George, T. (1989) Quantifying diversity in archaeology. Cambridge University Press.

Leprieur, F., Albouy, C., de Bortoli, J., Cowman, P.F., Bellwood, D.R. & Mouillot, D. (2012) Quantifying phylogenetic beta diversity: Distinguishing between 'true' turnover of lineages and phylogenetic diversity gradients. PLoS ONE **7**.

Levins, R. (1968) Evolution in changing environments, some theoretical explorations. Princeton University Press, no. 2 edn.

Levy, S.B. & Marshall, B. (2004) Antibacterial resistance worldwide: Causes, challenges and responses. Nature medicine **10**, S122–S129.

Lewontin, R.C. (1972) The apportionment of human diversity. Evolutionary Biology **6**, 381–398.

Lieberson, S. (1969) Measuring population diversity. American Sociological Review **34**, 850–862.

Lipsitch, M. & O'Hagan, J.J. (2007) Patterns of antigenic diversity and the mechanisms that maintain them. Journal of the Royal Society, Interface / the Royal Society **4**, 787–802.

Lloyd, M. & Ghelardi, R. (1964) A table for calculating the 'equitability' component of species diversity. The Journal of Animal Ecology **33**, 217–225.

Lozupone, C.A., Hamady, M., Kelley, S.T. & Knight, R. (2007) Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. Applied and Environmental Microbiology **73**, 1576–1585.

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

MacArthur, R.H. (1965) Patterns of species diversity. Biological Reviews **40**, 510–533.

Marcon, E., Scotti, I., Hérault, B., Rossi, V. & Lang, G. (2014) Generalization of the partitioning of shannon diversity. PloS one **9**, 1–8.

Marczewski, E. & Steinhaus, H. (1958) On a certain distance of sets and the corresponding distance of functions. Colloquium Mathematicae **6**, 319–327.

Margalef, D. (1958) Information theory in ecology. General Systems **III**.

Martin, E.J., Blaney, J.M., Siani, M.A., Spellmeyer, D.C., Wong, A.K. & Moos, W.H. (1995) Measuring diversity: Experimental design of combinatorial libraries for drug discovery. Journal of Medicinal Chemistry **38**, 1431–1436.

Mather, A., Matthews, L., Mellor, D., Reeve, R., Denwood, M., Boerlin, P., Reid-Smith, R., Brown, D.J., Coia, J.E., Browning, L., Haydon, D. & Reid, S. (2012) An ecological approach to assessing the epidemiology of antimicrobial resistance in animal and human populations. Proceedings. Biological sciences / The Royal Society **279**, 1630–9.

Mather, A.E., Reid, S.W.J., Maskell, D.J., Parkhill, J., Fookes, M.C. & Harris, S.R. (2014) Distinguishable epidemics within different hosts of the multidrug resistant zoonotic pathogen Salmonella typhimurium. Science **341**, 1514–1517.

May, R.M. (1990) Taxonomy as destiny. Nature **347**, 129–130.

McIntosh, R. (1967) An index of diversity and the relation of certain concepts to diversity. Ecology **48**, 392–404.

McMurdie, P.J. & Holmes, S. (2013) Phyloseq: An R Package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE **8**.

Miriagou, V., Carattoli, A. & Fanning, S. (2006) Antimicrobial resistance islands: Resistance gene clusters in Salmonella chromosome and plasmids. Microbes and Infection **8**, 1923–1930.

Mitchell, S. & Reeve, R. (2017) rdiversity: A package for measuring similarity-sensitive diversity.

Moreno, C.E., Castillo-Campos, G. & Verdú, J.R. (2009) Taxonomic diversity as complementary information to assess plant species diversity in secondary vegetation and primary tropical deciduous forest. Journal of Vegetation Science **20**, 935–943.

Morisita, M. (1959) Measuring of the interspecific association and similarity between communities.

Mouillot, D. & Wilson, J.B. (2002) Can we tell how a community was constructed? A comparison of five evenness indices for their ability to identify theoretical models of community construction. Theoretical Population Biology **61**, 141–51.

Mulvey, M.R., Boyd, D.a., Olson, A.B., Doublet, B. & Cloeckaert, A. (2006) The genetics of Salmonella genomic island 1. Microbes and Infection **8**, 1915–1922.

Mumford, J.A. (2007) Vaccines and viral antigenic diversity. Revue Scientifique Et Technique De L Office International Des Epizooties **26**, 69–90.

Nei, M. (1973) Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences **70**, 3321–3323.

Nei, M. & Li, W.H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences of the United States of America **76**, 5269–5273.

Nipperess, D.A., Faith, D.P., Barton, K., Nipperess, D.A., Faith, D.P. & Barton, K. (2010) Resemblance in phylogenetic diversity among ecological assemblages. Journal of Vegetation Science **21**, 809–820.

Office for National Statistics (2001) 2001 Census: Aggregate data (England and Wales) [computer file]. UK Data Service Census Support. Downloaded from: http://www.neighbourhood.statistics.gov.uk/dissemination and http://infuse.mimas.ac.uk. This data is licensed under the terms of the Open Government Licence [http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2].

Paradis, E., Claude, J. & Strimmer, K. (2004) APE: Analyses of phylogenetics and evolution in R language. Bioinformatics **20**, 289–290.

Patil, G.P. & Taillie, C. (1982) Diversity as a concept and its measurement. Journal of the American Statistical Association **77**, 548–561.

Pavoine, S., Bonsall, M.B., Dupaix, A., Jacob, U. & Ricotta, C. (2017) From phylogenetic to functional originality: Guide through indices and new developments. Ecological Indicators **82**, 196–205.

Peet, R.K. (1974) The measurement of species diversity. Annual Review of Ecology and Systematics **5**, 285–307.

Peet, R.K. (1975) Relative diversity indices. Ecology **56**, 496–498.

Pielou, E. (1967) The use of information theory in the study of the diversity of biological populations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability pp. 163–177.

Plazzotta, G. & Colijn, C. (2017) Phylodynamics without trees: Estimating R0 directly from pathogen sequences. bioRxiv pp. 1–15.

Plotkin, J.B. & Muller-Landau, H.C. (2002) Sampling the species composition of a landscape. Ecology **83**, 3344–3356.

R Core Team (2016) R: A Language and Environment for Statistical Computing.

Rambaut, A., Pybus, O.G., Nelson, M.I., Viboud, C., Taubenberger, J.K. & Holmes, E.C. (2008) The genomic and epidemiological dynamics of human influenza A virus. Nature **453**, 615–619.

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

Reeve, R., Leinster, T., Cobbold, C., Thompson, J., Brummitt, N., Mitchell, S. & Matthews, L. (2016) How to partition diversity. arXiv **1404.6520**, 1–9.

Rényi, A. (1961) On measures of entropy and information. Fourth Berkeley Symposium on Mathematical Statistics and Probability **1**, 547–561.

Revell, L.J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution **3**, 217–223.

Ricklefs, R.E. & Michael, L. (1980) Bias and dispersion of overlap indices: Results of some Monte Carlo simulations. Ecology **61**, 1019–1024.

Ricotta, C. (2004) A recipe for unconventional evenness measures. Acta Biotheoretica **52**, 95–104.

Ricotta, C. (2005) On hierarchical diversity decomposition. Journal of Vegetation Science **16**, 223–226.

Ricotta, C. (2008) Computing additive β -diversity from presence and absence scores: A critique and alternative parameters. Theoretical Population Biology **73**, 244–249.

Ricotta, C. (2010) On beta diversity decomposition: Trouble shared is not trouble halved. Ecology **91**, 1981–83.

Ricotta, C. & Szeidl, L.L. (2006) Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index. Theoretical Population Biology **70**, 237–243.

Routledge, R.D. (1979) Diversity indices: Which ones are admissible? Journal of theoretical biology **76**, 503–15.

Routledge, R.D. (1983) Evenness indices: Are any admissible? Oikos **40**, 149–151.

Sen, A. (1973) Poverty, inequality and unemployment: Some conceptual issues in measurement. Economic and Political Weekly **8**, 1457–1459.

Shannon, C. (1948) A mathematical theory of communication. The Bell System Technical Journal **XXVII**.

Shimatani, K. (2001) On the measurement of species diversity incorporating species differences. Oikos **93**, 135–147.

Simpson, E.H. (1949) Measurement of diversity. Nature **163**, 688–688.

Simpson, G.G. (1943) Mammals and the nature of continents. American Journal of Science **241**, 1–31.

Smith, B. & Wilson, J. (1996) A consumer's guide to evenness indices. Oikos **76**, 70–82.

Smith, W. & Grassle, J.F. (1977) Sampling propeties of a family of diversity measures. Biometrics **33**, 283–292.

Sørensen, T. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity c
København, I kommission hos E. Munksgaard.

Steinmann, K., Eggenberg, S., Wohlgemuth, T., Linder, H.P. & Zimmermann, N.E. (2011) Niches and noise Disentangling habitat diversity and area effect on species diversity. Ecological Complexity **8**, 313–319.

Stirling, A. (2007) A general framework for analysing diversity in science, technology and society. Journal of the Royal Society Interface pp. 707–719.

Swenson, N.N., Erickson, D.D., Mi, X., Bourg, N., Forero-Montaña, J., Ge, X., Howe, R., Lake, J., Liu, X., Ma, K., Pei, N., Thompson, J., Uriarte, M., Wolf, A., Wright, S., Ye, W., Zhang, J., Zimmerman, J., Kress, W., Forero-Montana, J., Ge, X., Howe, R., Lake, J., Liu, X., Ma, K., Pei, N., Thompson, J., Uriarte, M., Wolf, A., Wright, S., Ye, W., Zhang, J., Zimmerman, J. & Kress, W. (2012) Phylogenetic and functional alpha and beta diversity in temperate and tropical tree communities. Ecology **93**, 112–125.

Tamás, J., Podani, J. & Csontos, P. (2001) An extension of presence/absence coefficients to abundance data: A new look at absence. Journal of Vegetation Science **12**, 401–410.

Threlfall, E.J., Frost, J.A., Ward, L.R. & Rowe, B. (1994) Epidemic in cattle and humans of Salmonella Typhimurium DT104 with chromosomally integrated multiple drug resistance. Veterinary Record **134**, 157.

Tóthmérész, B. (1995) Comparison of different methods for diversity ordering. Journal of Vegetation Science **6**, 283–290.

Tsirogiannis, C. & Sandel, B. (2015) PhyloMeasures: A package for computing phylogenetic biodiversity measures and their statistical moments. Ecography pp. 709–714.

Tsirogiannis, C. & Sandel, B. (2016) Fast computations for measures of phylogenetic beta diversity. PLoS ONE **11**, 1–16.

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

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

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

Tuomisto, H. (2011) Commentary: Do we have a consistent terminology for species diversity? Yes, if we choose to use it. Oecologia **167**, 903–911.

Tuomisto, H. (2012) An updated consumer's guide to evenness and related indices. Oikos **121**, 1203–1218.

Vane-Wright, R., Humphries, C. & Williams, P. (1991) What to protect? Systematics and the agony of choice. Biological Conservation **55**, 235–254.

Veech, J.A. & Crist, T.O. (2010) Toward a unified view of diversity partitioning. Ecology **91**, 1988–92.

Veech, J.A., Summerville, K.S., Crist, T.O. & Gering, J.C. (2002) The additive partitioning of species diversity: Recent revival of an old idea. Oikos **99**, 3–9.

Vellend, M. (2001) Do commonly used indices of βdiversity measure species turnover? Journal of Vegetation Science **12**, 545–552.

Vellend, M., Cornwell, W., Magnuson-Ford, K. & Mooers, A. (2007) Measuring phylogenetic biodiversity. Biological Diversity: Frontiers in Measurement and Assessment (eds. A. Magurran & B.E. McGill), chap. 14, Oxford University Press.

Villéger, S. & Mouillot, D. (2008) Additive partitioning of diversity including species differences: A comment on Hardy & Senterre (2007). Journal of Ecology **96**, 845–848.

Warwick, R.M. & Clarke, K.R. (1995) New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. Marine Ecology Progress Series **129**, 301–305.

Webb, C. (2000) Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. The American naturalist **156**, 145–155.

Webb, C.O., Ackerly, D.D. & Kembel, S.W. (2008) Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics **24**, 2098–2100.

Webb, C.O., Ackerly, D.D., Mcpeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. Annual Review of Ecology and Systematics **33**, 475–505.

Weitzman, M. (1993) What to preserve? An application of diversity theory to crane conservation*. The Quarterly Journal of Economics .

Whittaker, R. (1960) Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs **30**, 279–338.

Whittaker, R. (1972) Evolution and measurement of species diversity. Taxon **21**, 213–251.

Wilson, M. & Shmida, A. (1984) Measuring beta diversity with presence-absence data. Journal of Ecology **72**, 1055–1064.

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

World Health Organization (2015) Antimicrobial resistance.

Wright, S. (1951) The genetical structure of populations. Annals of Eugenics **15**, 323–354.

Wright, S. (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution **19**, 395–420.

Yang, Z. & Rannala, B. (2012) Molecular phylogenetics: Principles and practice. Nature Reviews Genetics **44**, 303–314.