Palacký University Olomouc

Faculty of Science

Department of Ecology and Environmental Sciences

Ecological processes in vegetation driven by functional traits and phylogeny

Martin Bitomský

A doctoral Dissertation Presented to

the Faculty of Science

Palacký University Olomouc

In Partial Fulfilment

of the Requirements for the Degree

Philosophiae Doctor in Ecology

Supervisor: prof. MVDr. Emil Tkadlec, CSc.

Olomouc 2020

© Martin Bitomský, 2020

Department of Ecology and Environmental Sciences

Palacký University in Olomouc

ABSTRACT

Phylogenies are a crucial (but still overlooked) source of information in ecology. Accompanied with trait data, phylogenetic information can be used to study the most important ecological questions and processes that have implications for nature conservation and community ecology. This recent scientific field is often called community phylogenetics. In this thesis, I summarise our current knowledge about the concept of phylogenetic diversity, the relationship between phylogenetic and functional diversity, and functional and phylogenetic patterns in herbaceous vegetation in response to productivity. I primarily focus on analytical approaches and also provide an up-to-date critical review of some of limitations and issues with popular methods, while I also highlight some important considerations related to every ecological analysis with a phylogenetic and trait framework. Finally, I encourage to routinely implement phylogenetic information (and phylogenetic comparative methods) in ecology because it is needed due to both biological and statistical reasons.

ABSTRAKT

Fylogeneze jsou stěžejním (ale pořád stále přehlíženým) zdrojem informací v ekologii. Společně s daty o funkčních znacích může být fylogenetická informace použita pro studium těch nejdůležitějších ekologických otázek, které mají důsledky pro ochranu přírody a komunitní ekologii. Tento čerstvý vědecký obor se často označuje jako komunitní fylogenetika. V této práci shrnuji naše dosavadní znalosti ohledně konceptu fylogenetické diverzity, vztahu mezi fylogenetickou a funkční diverzitou, či funkčními a fylogenetickými procesy ve vegataci s ohledem na změny v její produktivitě. Primárně se zaměřuji na analytické přístupy a také předkládám aktuální kritický souhrn limitací a poblémů, které některé populární metody mají, zatímco také zdůrazňuji některé okolnosti vyžadující pozornost v případě všech ekologických analýz s fylogenetickým a funkčním rámcem. Na závěr povzbuzuji k tomu, aby se fylogenetická informace (a fylogenetické komparativní metody) rutinně používala v ekologii, protože je jí potřeba z biologických a statistických důvodů.

CONTENTS

ACKNOWLEDGEMENTS

I have had the good fortune of learning from incredible people. First and foremost, I thank Jan Mládek († 2018) for showing me how to do science. I benefited greatly from discussions with Emil Tkadlec, mostly about statistics. I would like to thank many people who have been collaborating with me during my Ph.D. years, namely Pavla Mládková, Šárka Cimalová, Martin Duchoslav, Hanno Schaefer and Robin Pakeman. I am grateful for the opportunity to study and work at the Department of Ecology and Environmental Sciences of the Palacký University in Olomouc. The R community and many other researchers, who have made community phylogenetics so attractive thanks to their tremendous work, also deserve credits.

My research was supported by the IGA projects of the Palacký University in Olomouc (IGA_PrF_2016_019, IGA_PrF_2017_015, IGA_PrF_2018_020, IGA_PrF_2019_021 and IGA_PrF_2020_020), the Scholarship of the city of Ostrava (Ostrava City Hall) and the LTC18056 project (COST action CA16212; INTER-COST program, Ministry of Education, Youth and Sports).

And last but not least I thank my dog Arnošt, my cat Žofinka, my lovely wife Dominika and my family for keeping me so happy. I dedicate this work to Albert Bitomský.

LIST OF PUBLICATIONS

This doctoral dissertation is based on the following publications (+ my percentage contributions):

1. **Bitomský M (80%)**, Mládek J, & Cimalová Š (2018). Light limitation shapes the community seed mass of annual but not of perennial weeds. *Perspectives in Plant Ecology, Evolution and Systematics* 31: 1–6.
2. **Bitomský M (80%)**, Mládková P, Cimalová Š, & Mládek J (2019). Herbaceous climbers in herbaceous systems are shade-tolerant and magnesium-demanding. *Journal of Vegetation Science* 30: 799–808.
3. **Bitomský M (80%)**, Mládková P, Pakeman RJ, & Duchoslav M (2020). Clade composition of a plant community indicates its phylogenetic diversity. Ecology and Evolution (early view). doi: 10.1002/ece3.6170
4. **Bitomský M (80%)**, Schaefer H, Pakeman RJ, & Duchoslav M et al. (in review). Variation in mycorrhizal status is much larger within than among sites: a phylogenetically-informed assessment across grassland and coastal floras.

CHAPTER ONE

A NEW ERA OF PHYLOGENETICALLY-INFORMED ECOLOGY

Phylogenies have revolutionised research in plant ecology as they help to understand a mechanistic basis of ecological processes at various spatio-temporal scales. Thanks to rapidly increasing availability of synthesis-based phylogenies (i.e. published super-trees; sensu Li et al. 2019), DNA sequences, computing power and bioinformatic tools, community phylogenetics has become one of the leading fields in ecology despite its quite recent origin (the nineties and the beginning of this millennium). The four seminal papers (Faith 1992, Webb et al. 2002, Cavender-Bares et al. 2004, 2009) set the ground for phylogenetically-informed community ecology and nature conservation (Table 1.1). Faith (1992) introduced the concept of phylogenetic diversity (PD) as a natural phylogeny-based measure of biodiversity. This paper is the most-cited research article ever in *Biological Conservation* from more than 7000 papers in that journal spanning more than 40 years (Faith 2018). Webb et al. (2002) wrote a thorough review on how phylogenetic information can connect ecological and evolutionary studies. Cavender-Bares et al. (2004) in their original research article popularised the concept of phylogenetic clustering (co-occurring species sharing the same abiotic environment should be more closely related than expected) and phylogenetic overdispersion (co-occurring species sharing the same abiotic environment should be more distantly related than expected), and tested it on Floridian oak communities. Finally, Cavender-Bares et al. (2009) wrote an influential review where they discuss how phylogenies might help to explain some ecological patterns.

Phylogenetic ecology is closely linked to functional (trait-based) ecology, and both approaches are considered to provide complementary information in several conceptual frameworks (e.g. Cadotte et al. 2013; de Bello et al. 2017, Li et al. 2017, Ovaskainen et al. 2017). Species interactions are based on their phenotypic differences and/or similarities, and that phenotypic variation has an evolutionary basis (Webb et al. 2002, Cavender-Bares et al. 2009). Similarly, the species responses to environmental factors and neutral processes are constrained by phylogenetic relationships as they depend on their traits that might be phylogenetically structured. Thus, it is not surprising that both phylogenetic and functional ecology have influenced each other and often use analogous methods (Swenson 2014).

Table 1.1 Four seminal papers (that have shaped the development of community phylogenetics) with a brief overview and the number of citations according to *Web of Science* (26.2.2020).

Study	Overview	Cited (WoS)
Faith 1992	The author introduced the concept of phylogenetic diversity and proposed that it enables setting of conservation priorities.	2054
Webb et al. 2002	The authors reviewed various approaches how to incorporate phylogenetic information into community ecology.	2118
Cavender-Bares et al. 2004	The authors examined the links phylogenetic relatedness and co-occurrence of oak species in Florida.	476
Cavender-Bares et al. 2009	The authors reviewed some areas of community ecology where phylogenetic information helps to resolve long-standing controversies.	1102

Even though community phylogenetics is a relatively recent scientific field, it is impossible to provide a thorough review of all aspects. Thus, I only attempt to summarise our knowledge about the processes I have studied during my Ph.D. years, namely the estimation of phylogenetic diversity (Chapter Two), the relationship between phylogenetic and functional diversity (Chapter Three), and functional and phylogenetic patterns in vegetation in response to productivity (Chapter Four). All these three chapters have a similar structure. The section *State of the Field* provides an up-to-date summary of the given aspect, the section *Analytical Approaches* describes commonly used methodology, while the section *Challenges* provides an in-depth critical review of known analytical issues and unresolved questions. I also briefly present my work that is related to the topics discussed in this thesis. All my research articles can be found in the *Supplementary Section* (Appendices A–D). In the fifth chapter, I synthesize and propose some outstanding questions I think are worthy to study in the future.

CHAPTER TWO

PHYLOGENETIC DIVERSITY

2.1 State of the field

A simple conceptual example by Vellend et al. (2011) captures the need to consider phylogeny-based measures in community ecology or nature conservation: there are two hypothetical islands, each with only two species of vertebrate animals in equal abundance: two birds in one case and a bird plus a mammal in the other. Both islands have species richness = 2 (for vertebrates) and the same maximal value of species evenness. However, our intuition tells us that a bird plus a mammal represents more biodiversity than does two birds. Probably the first attempt to capture this aspect of biodiversity were the species/genus, species/family ratios (e.g. Elton 1946). A similar approach based on taxonomic relatedness (derived from a hierarchical Linnaean classification with applied taxonomic weights proportional to the level of the taxonomic rank two species hold in common, i.e. genus, family or order) has proven to be useful to estimate biodiversity patterns in fish communities (Warwick & Clarke 1995, Hall & Greenstreet 1998, Campbell et al. 2010). However, with the recent rapid accumulation of DNA sequence data, more and more phylogenies are being constructed based upon sequence comparisons. Therefore, the vast majority of up-to-date PD measures is based on phylogenetic distances (the amount of time since the most common ancestor of a pair of species), which are derived from phylogenetic trees. For the purposes of nature conservation or community ecology, phylogenetic distances can be then used to set conservation priorities or to inform our understanding of the assembly mechanisms that structure communities, respectively.

Evolutionary relationships between species have been proposed as one of the components of biodiversity relevant for nature conservation (Vane-Wright et al. 1991, Faith 1992, Winter et al. 2013). The original motivation to consider PD in nature conservation is simple: species loss is accompanied with the loss of evolutionary information. But why to protect PD per se? Intuitively, phylogenetically diverse sampling units (communities, plots, sites, regions etc.) are considered to have "higher biological value" as they contain more evolutionary history. Additionally, PD has been proposed to be related to processes such as biotic invasion (Winter et al. 2009), ecosystem functioning and services (Faith et al. 2010, Srivastava et al. 2012) or extinction (Purvis et al. 2000). Winter et al. (2013) summarised four central justifications for the conservation of PD: 1) rare species are often the first to become extinct, and phylogenetic distinctiveness is often correlated with rarity (the rarity aspect; Purvis et al. 2000, Redding & Mooers 2006); 2) phylogenetic approach reveals areas with evolutionarily very young or old clades, and phylogenetic measures can reflect processes that are not captured by changes in species richness (the richness aspect; Warwick & Clarke 1995, Forest et al. 2007, Knapp et al. 2008); 3) maximizing PD ensures that a wide variety of forms and functions, described as functional diversity (FD), are present within a species set, which helps to maintain ecological processes at an ecologically relevant timescale (PD as a proxy for FD; Webb et al. 2002, Cavender-Bares et al. 2009, Cadotte et al. 2011); 4) an increase in PD of a community increases the evolutionary potential, i.e. species capacities to evolve in response to environmental changes, which might be an important feature of communities in the face of the current environmental crisis (phylogenetic diversity as a proxy for evolutionary potential; Lavergne et al. 2010, Lankau et al. 2011).

Phylogenies provide an important insight into community ecology as well because they enable to infer mechanisms of community assembly. Species attributes (e.g. niche conservatism) and the strength of species interactions (e.g. competition or facilitation), which structure community composition, can be inferred from phylogenetic distances (Webb et al. 2002, Cavender-Bares et al. 2009). The key assumption is that there is a positive relationship between phylogenetic distance and ecological similarity. For example, the results from a field experiment (Burns & Strauss 2011) suggest that more closely related plants are also more ecologically similar. Species with a recent common evolutionary history will, therefore, likely share similar environmental constraints and strategies for the use of resources, but will also inevitably compete more strongly for those same resources (Kraft et al. 2015). On the other hand, facilitative interactions tend to be more common between distantly related species (Valiente-Banuet & Verdú 2013). Most notable is the idea that habitat filtering results in phylogenetic clustering, while competition results in phylogenetic overdispersion (Cavender-Bares et al. 2004). Many studies use or mention this simplistic framework for testing community assembly processes (743 papers found when I submitted the query “phylogenetic clustering” AND “habitat filtering“, and 1230 papers found when I submitted the query “phylogenetic overdispersion“ AND competition in *Google Scholar* on March 9th 2020).

At this point, it is very important to note that all ideas regarding the role of PD in nature conservation and community ecology described above are still controversial and may be based on implicit assumptions whereby some of them are weakly supported (e.g. Mayfield & Levine 2010, Gerhold et al. 2015, Mazel et al. 2018) or might be biased by methodological decisions (e.g. Park et al. 2018, Jantzen et al. 2019, Li et al. 2019). Understanding of these biases is important for correct and justified use of phylogenies in ecology. Discussion is provided in the *Challenges* sections (2.3, 3.3 and 4.3) of this and the next chapters.

2.2 Analytical approaches

Various sets of PD metrics tend to correlate (Vellend et al. 2011, Swenson 2014), thus, there has been some attempts to develop a unifying framework. Tucker et al. (2017) proposed that all described PD measures reflect three dimensions of phylogenies: richness, divergence and regularity. This classification comprises more than 70 existing metrics and each dimension can be proxied by a leading measure, namely, Faith's PD (Faith 1992) for the richness dimension, mean phylogenetic distance (MPD, Webb et al. 2002) for the divergence dimension, and variation of pairwise distances (VPD, Clarke & Warwick 2001) for the regularity dimension. Obviously, each dimension tries to estimate a different aspect of phylogenetic distances of co-occurring species in a community (see Table 1 in Appendix C). Richness estimates the total evolutionary history (sum), divergence estimates phylogenetic relatedness (mean) and regularity estimates distribution of phylogenetic relatedness (variation). The framework by Tucker et al. (2017) helped to navigate current jungle of various approaches for estimation of alpha (i.e. within sites) and beta (i.e. among sites) PD, except for beta regularity because in that time there was no measure of phylogenetic beta regularity. But Scheiner et al. (2017) filled this gap shortly after. Finally, this framework is also related (to some extent) to Hill numbers (diversity measures that aim to quantify diversity in units of equivalent numbers of equally abundant species) including phylogenetic information (Hill 1973, Chao et al. 2010, 2014).

There is another classification of PD based on two qualitatively different approaches (type I and type II; Vellend et al. 2011). Type I measures begin by calculating an evolutionary distinctness score for all species in a regional phylogeny (species pool) and then calculating some function of these scores (typically the sum) for particular focal subsets of species to yield an estimate of PD. Type I measures are sometimes called EDGE-like (Isaac & Pearse 2018), where the EDGE stands for Evolutionary Distinct Globally Endangered (Isaac et al. 2007), and have been widely implemented in conservation decisioning. Type II metrics start with a local phylogeny (or possibly a regional phylogeny), and for a focal subset of species they depend only on properties of the subset phylogeny (i.e. the pruned regional phylogeny). The type II category comprises all leading measures mentioned above (Faith's PD, MPD and VPD). In community ecology, the type II approach is dominant, on the other hand, the type I approach is more frequent in conservation studies.

For the purposes of study questions in community ecology, it is usually recommended to perform a null model analysis in order to standardise a PD metric and account for its dependence on species richness (Webb et al. 2002, Swenson 2014, Miller et al. 2017). The idea behind this treatment is to get additional (or maybe even orthogonal) information on top of simple species richness. For example, higher PD than expected in a 20-species quadrat should be considered as more impressive than the same value in a 5-species quadrat. There are various ways how to do these standardisations, e.g. by null model randomisations (Swenson 2014, Miller et al. 2017), direct analytical expressions (Tsirogiannis & Sandel 2016) or rarefaction (Sandel 2018). The most popular methods for PD estimation are implemented in several R packages, such as *picante* (Kembel et al. 2010) or *PhyloMeasures* (Tsirogiannis & Sandel 2016).

2.3 Challenges

The whole process of PD estimation is affected by important methodological decisions and biases. As discussed in Li et. al (2019), researchers have to decide what markers (Which genes to select?) and methods to use (Alignment method? Model of evolution? Maximum likelihood or Bayesian inference framework? What fossil constraints for molecular dating?). Besides that, PD metrics depend on the attributes of phylogenies (that might be an inherent feature of the studied species pool, but also might be the result of errors or unsuitable methods used), such as the degree of tree balance, diversification rate, resolution, taxon sampling or tree reconstruction methods (Swenson

2009, Vellend et al. 2011, Park et al. 2018, Jantzen et al. 2019). Fortunately, some of these issues can be avoided as there are synthesis-based super-trees available for vascular, angiosperm or European plants (e.g. Durka & Michalski 2012, Zanne et al. 2014 or an updated version by Qian and Jin 2016, Hinchliff et al. 2015, Smith & Brown 2018, Janssens et al. 2020). Super-trees save a considerable amount of time and effort, on the other hand, they could have more polytomies (lower resolution). However, as long as these polytomies are at more recent nodes, the impact on PD estimation should be low (Swenson 2009), and Li et al. (2019) showed that published super-trees can be used for commonly employed community phylogenetic analyses. Additionally, Cadotte (2015) also demonstrated that changing branch lengths did not strongly affect relationships between phylogenetic diversity and ecosystem function, suggesting that phylogenetic diversity measures are not so sensitive to the branch lengths of the phylogeny as long as the topology is right.

This robustness led me to an idea to develop a method of PD estimation, which does not rely on phylogenetic distances but is based on categorical taxonomic coding at the family level (Bitomský et al. 2020, Appendix C). These indices only include information about how families (or any defined clade) are represented in a community (presence and relative abundance). A simple simulation (Fig. 1 in Appendix C) indicated that phylogenetic richness (Faith’s PD) tends to increase with increasing proportion of the most distantly related species in comparison with the rest of the species in the community. Phylogenetic divergence (MPD) was relatively high when all defined clades (i.e. monocots, Ranunculales, super-rosids and super-asterids) had equal proportions. Finally, phylogenetic regularity was relatively high (i.e. VPD was low) when the defined clades had proportions proportional to their relative species richness in the species pool. This information was behind the equations of proposed measures (the clade indices) in our study (Table 2 in Appendix C). It appears our method provides estimates that correlate with all leading metrics of PD (richness, divergence and regularity) quite well, except for phylogenetic divergence and regularity at very large phylogenetic scales.

Our approach is a promising alternative for all projects with a phylogenetic framework, which can trade-off a little precision for a significant speed-up and simplification. However, there is still a plenty of recently localised issues (Table 2.1) affecting commonly applied approaches for PD estimation (phylogeny-based and taxonomy-based measures). For the type II approach, the common practice is to infer a full tree of the complete species list (i.e. all species occurring in all sampled units, such as quadrats) and then prune it to small trees for each sampled unit. However, Park et al. (2018) provided convincing evidence that inferring phylogenies with only taxa within the area of interest can lead to spurious estimates of alpha and beta PD, because these inferred phylogenies are biased in their taxon sampling (Table 2.1a). Undersampling taxa leads to the underestimation of divergence times (Linder et al. 2005, Schulte 2013), and to general underestimation of PD and incorrect ranking of PD among sites (when one is just interested in relative differences, Park et al. 2018). Broader taxon sampling also can increase statistical power as more taxa in a study increases the likelihood of observing significantly non-random phylogenetic patterns (Jantzen et al. 2019). These perspectives suggest that super-trees with broader taxon sampling should be preferred.

Null-model based standardisations (Swenson 2014, Miller et al. 2017) are commonly used to remove the effect of species richness on PD estimates. However, Sandel (2018) recently showed that this assumption is incorrect, especially in situations when the environmental filter favours a moderate-sized clade strongly over others and when using the net relatedness index (NRI, a standardised version of MPD). This suggests that samples with varying species richness are not comparable, which might affect type I error rates (Table 2.1b). Sandel (2018) proposed rarefaction to fix this issue, but at the expense of increased error. Some studies use exclusively rarefied versions of PD measures (e.g. Sandel et al. 2020), but it appears so far that both approaches provide similar conclusions about some phylogenetic patterns, e.g. in tropical tree communities along the precipitation gradient (Neves et al. 2020).

Table 2.1 Summary of the most important considerations and issues that affect the estimation of PD and interpretation of phylogenetic patterns in communities.

Issue	Description	Source
a) Taxon sampling	Inferring regional phylogenies using only species from the species pool is biased in its taxon sampling, which leads to underestimation of PD.	Park et al. (2018)
b) Species richness dependence	Null models are not enough to treat for the species richness dependence of PD measures under some conditions.	Sandel (2018)
c) Phylogenetic identity	PD effects can be confounded by phylogenetic identity effects when assemblages with low or high PD tend to be dominated by a single clade.	Hipp et al. (2018)
d) Lineage specificity	When opposing patterns in two clades cancel out, it might result in seemingly random or even opposite phylogenetic processes compared to when considering both clades together.	Ndiribe et al. (2013) Elliott et al. (2016) Graham et al. (2018)
e) Molecular dating	Chronograms (dated phylogenetic trees with branch lengths in units of evolutionary time) can provide different estimates of PD than phylograms (branch lengths in units of substitutions per site).	Elliott et al. (2018) Jantzen et al. (2019)

Taxon sampling is related to another issue: the confounding effect of phylogenetic identity. When the species pool exhibits a phylogenetic bias, low or high PD could be disproportionately associated with certain clades (Hipp et al. 2018). For example, without a sufficient sampling in grasslands, low PD plots could be frequently dominated by a single clade (e.g. Poaceae or Asteraceae), while high PD plots could be frequently dominated by phylogenetically distinct species from clades that are underrepresented in the species pool. As a consequence, it would be difficult to tease apart the effects on a plot being dominated by Poaceae from the effects of having a low PD species assemblage (Table 2.1c). Hipp et al. (2018) suggest a visualisation tool and a taxonomically constrained experimental design to deal with the confounding effect of phylogenetic identity. Some authors have also highlighted that various phylogenetic processes are lineage-specific and dependent on phylogenetic scale (Table 2.1d; Ndiribe et al. 2013, Elliott et al. 2016, Graham et al. 2018). The effect of environmental factors on the distribution of species from different clades can differ (Duarte et al. 2012, Descombes et al. 2017, Graham et al. 2018), for example, when opposing patterns in two clades cancel out, it might result in seemingly random or even opposite phylogenetic processes when considering both clades together (Elliott et al. 2016, Graham et al. 2018). To acknowledge this issue, approaches, which explore lineage-specific patterns that allow for opposing mechanisms across the tree topology or specifically evaluate lineage-specific co-occurrence patterns, might be a good idea (Ndiribe et al. 2013, Elliott et al. 2016).

Finally, researchers have to decide whether to use a chronogram (a dated phylogenetic tree with branch lengths in units of evolutionary time) or a phylogram (branch lengths in units of substitutions per site). Both approaches are quite common, some popular maximum likelihood programs for phylogeny inference (such as RaxML; Stamatakis 2014) produce phylograms that can be dated later using another software (e.g. PATHD8; Britton et al. 2007). On the other hand, BEAST (Suchard et al. 2018) directly infers chronograms using the Bayesian framework. Molecular dating could be influenced by time priors (external information about some node ages), clock or speciation model, therefore, possible differences between phylograms and chronograms should be expected. Jantzen et al. (2019) found that in some cases both approaches provide different results, and Elliott et al. (2018) argue that the choice between phylogram and chronogram can have a dramatic effect on biodiversity hotspot distributions, which could affect conservation decisioning (Table 2.1e).

CHAPTER 3

PHYLOGENY-TRAIT LINKS

3.1 State of the field

Some traits are used as proxies of performance (growth, survival or reproduction) under specific environmental conditions, which indicates that traits should also provide information about species' niches (McGill et al. 2006). Traits evolve steadily through time, thus, branching patterns of phylogenetic trees, such as topology, branch lengths or tree balance, should also produce comparable estimates of niche space to those obtained by traits. This logic advocates the use of phylogenetic measures in the absence of relevant trait data when traits are sufficiently phylogenetically conserved (Kraft et al. 2007, de Bello et al. 2015). Phylogenetic information is easier to obtain as there is no need for time- and money-consuming measuring of various plant traits, therefore, many studies discuss whether PD serves as a reliable proxy for FD. It is clear now that methodological, evolutionary or ecological processes (e.g. the size of plot used, number of species, number of traits measured or the mode and rate of trait evolution) can bias the inferred PD-FD relationship (Pavoine et al. 2013, Gerhold et al. 2015, Cadotte et al. 2017, Burns et al. 2018, Cadotte et al. 2018, Tucker et al. 2018). Nevertheless, good understanding of the PD-FD relationship could have implications for ecosystem functioning, nature conservation or mitigating the impact of biotic invasions (Srivastava et al. 2012, Cadotte et al. 2018, Burns et al. 2018, Mazel et al. 2017, Zheng et al. 2018).

3.2 Analytical approaches

Following Cadotte et al. (2019), I summarise two common types of questions regarding the phylogeny-trait links asked in ecological studies, which are tested by different analytical approaches. Question 1: Are there phylogenetic signals in the measured traits in communities? Phylogenetic signal (PS) is the tendency of related species to resemble each other more than species drawn at random from the same tree (Münkemüller et al. 2012). For continuous traits, measures, such as Pagel's lambda (λ, Pagel 1999) or Blomberg's K (Blomberg et al. 2003), are very common. For categorical traits, there are also some alternatives, e.g. the alpha parameter of the phylogenetic logistic regression (Ives & Garland 2010) and the D statistic (Fritz & Purvis 2010) for binary traits (0/1), or the δ statistic (Borges et al. 2018) for traits with more than two states. Measures of PS depend on various statistical frameworks: permutations (Blomberg's K and D), maximum likelihood (Pagel's λ) or Bayesian (δ), and their performance under various conditions has been thoroughly examined (Münkemüller et al. 2012, Borges et al. 2018).

Generally, Pagel's λ is recommended because it performed well in simulations under the Brownian motion (BM) model of trait evolution (Münkemüller et al. 2012) and seems strongly robust to incompletely resolved phylogenies (Molina-Venegas & Rodríguez 2017). For categorical traits, the δ statistic is probably more robust than D because the D statistic cannot be used in studies where categorical traits do not evolve according to the BM model (Felsenstein 2005, Borges et al. 2018). On the other hand, the δ statistic is developed for gene trees, i.e. a phylogenetic approach that is not common in community phylogenetics (it also restricts the use of published super-trees). Any measure of PS can be easily estimated using several R packages, such as *geiger* (Pennell et al. 2014), *phytools* (Revell 2012) or *phylosignal* (Keck et al. 2016). About 76% of studies in the ecological literature report phylogenetic signal in the measured traits (Cadotte et al. 2019).

Question 2: Are PD and FD correlated across multiple sites, i.e. is there the PD-FD relationship? Intuitively, multiple forms of diversity should covary spatially, which is especially the case for the richness dimension where all indices mathematically depend on species richness (Pavoine et al. 2013). Overall, the PD-FD relationship tends to be tight and positive (Cadotte et al. 2019), which is the expected and intuitive result. However, there is an increasing evidence that interpretation based on parametric correlations can be misleading. Letten & Cornwell (2015) showed that the assumption that trait dissimilarity increases linearly with divergence time is incorrect (under BM trait evolution), and suggested a square root transformation of the phylogenetic distance matrix to fix this non-linearity. Probably, the most important driver of the PD-FD relationship is the character of trait evolution. Tucker et al. (2018) simulated traits according to the BM (a random walk) and Ornstein-Uhlenbeck (OU; where trait is attracted to a selection optimum with some degree of selection strength) models and found that the former suggests an exponential growth model ($FD \sim a\, e^{1 - bPD}$), while the latter suggests a logistic model ($FD \sim a + b \ln PD$) to fit the PD-FD relationship. The degree of covariance between PD and FD tends to non-linearly increase with increasing number of traits used to compute functional distances (there seems to be a threshold around eight traits when the explained variance of functional distances is at its peak; Tucker et al. 2018). To my knowledge, there are other calls for the need to test alternative models of evolution to obtain better understanding of complex relationships between phylogenetic and functional distance (see, for example, Burns et al. 2018, Cadotte et al. 2018).

We considered the above mentioned issues when studying the nature of the relationship between MPD (phylogenetic divergence) and mean functional distance (MFD, functional divergence) of mycorrhizal status in vegetation quadrats sampled in Czech semi-natural grasslands and Scottish coastal habitats (Bitomský et al. in review, Appendix D). Mycorrhizal status denotes whether the roots of a plant species are consistently (obligatory mycorrhiza, OM), sometimes, but not always (facultative mycorrhiza, FM) or never (non-mycorrhiza, NM) colonised by mycorrhizal fungi in nature (Smith & Read 2008, Hempel et al. 2013, Moora 2014). We followed the usual numerical coding of mycorrhizal status: NM plants = 0, FM plants = 0.5 and OM plants = 1 (Moora 2014, Gerz et al. 2016, 2019). This ability to form mycorrhiza is usually phylogenetically conserved (Brundrett 2009, Brundrett & Tedersoo 2019) and has been under stabilising selection that favours mycorrhizal symbiosis over mutualism abandonment in seed plants (Maherali et al. 2016). We found that a non-linear saturating curve explains the PD-FD relationship between MPD and MFD of mycorrhizal status (Appendix D, Fig. 1b), which is expected for a trait under stabilising selection (Tucker et al. 2018). Our results support the idea that simple correlations are not enough to infer the nature of the PD-FD relationship, while the consideration of alternative evolutionary models and non-linear statistical approaches is necessary.

Many models of trait evolution have been developed so far to address various evolutionary processes (Cooper et al. 2016, Harmon 2019). They can now account for stabilising selection (Hansen and Martins 1996, Hansen 1997), multiple optima (Butler and King 2004), and differing rates of evolution across taxa (O’Meara et al. 2006) or through time (Pagel 1999, Blomberg et al. 2003). For discrete traits, a general approach is to fit the Mk model (Lewis 2001, Harmon 2019). To analyse macroevolutionary models and complex PD-FD relationships, packages, such as geiger (Pennell et al. 2014), nlme (Pinheiro et al. 2019) or mgcv (Wood 2017), are useful.

3.3 Challenges

Cadotte et al. (2019) performed a literature survey and reported that the PD-FD relationship is usually positive. Obviously, a publication bias plays a role as studies with significant results are more likely to be published. However, other studies reporting no or even negative PD-FD relationship cannot be simply ignored (e.g. Prinzing et al. 2008, Kelly et al. 2014, Mazel et al. 2017, 2018). These inconsistencies question some of the most prominent arguments why to use PD in nature conservation and community ecology (e.g. Gerhold et al 2015). For example, Mazel et al. (2018) analysed more than 15 000 vertebrate species and concluded that maximising PD to help to protect FD is a risky conservation strategy. But what is the reason of these inconsistencies? Ubiquitous stochastic evolution of some traits? Statistical properties (e.g. error rates, inconsistent parameter estimates) of widely used methods? Wrong assumptions regarding underlying evolutionary processes? I find these challenges worth to study in detail in order to understand the PD-FD relationship, thus, I briefly discuss them here (Table 3.1).

Table 3.1 Summary of the most important considerations and issues that affect the nature of the PD-FD relationship and interpretation of patterns of trait evolution.

Issue	Description	Source
a) The concept of the phylogenetic signal can be misleading	Some properties of the tree (e.g. taxon sampling, size and shape) affect phylogenetic signal estimates. Functional distances could correlated with phylogeny only for a short distance along the tree.	Boettiger et al. (2012) Kelly et al. (2014) Mazel et al.(2017)
b) Phylogeny balance	The shape of the tree affects phylogenetic distances among species. Trait and phylogenetic distances might have different distributions, which affects statistical properties.	Heard (1992) Cadotte et al. (2017)
c) Phylogenies might be uniformative	Some phylogenies can lack reasonable power to detect relevant patterns of trait evolution. Robustness should be tested.	Boettiger et al. (2012) Paterno et al. (2018)
d) Trees with only present-day taxa are limiting	Fossil taxa and their traits might be necessary to obtain correct estimates when fitting trait evolutionary models.	Ho & Ané (2013, 2014)
e) Comparison based on information criteria might be biased	AIC or BIC might favour overly complex (or overly simplistic) evolutionary scenarios.	Ané (2008) Boettiger et al. (2012) Ho & Ané (2014)

Communities assembled from few lineages can have higher variation in trait-states than communities assembled from many lineages, despite the phylogenetic signal present in the measured traits (Prinzing et al. 2008). Mazel et al. (2017) concluded that measuring phylogenetic signal in traits is uninformative for evaluating the effectiveness of using PD as a proxy of FD under some models of trait evolution and tree shapes (Table 3.1a). Probably, the reason is that traits are not informative for deeper splits in the tree (Kelly et al. 2014). In other words, functional distance can be correlated with phylogenetic distance only for a short relative distance along the tree. That is probably the reason why measures of phylogenetic signal heavily depend on whether or not all sister species are present in the data (Boettiger et al. 2012). Adding sister species to the phylogeny increases the likelihood of finding phylogenetic signal, which is one of several reasons why the concept of phylogenetic signal could be misleading. This, again, stresses the effect of taxon sampling, which seems to also affect the PD-FD relationship.

The size (number of tips) and shape of the phylogeny is key to understanding how much information about evolutionary processes it is possible to extract from traits of species at the leaves of the tree (Boettiger et al. 2012). Surrogacy of PD for FD appears to weaken as the species pool richness increases (Mazel et al. 2018). The shape is described by the tree balance, i.e. the degree to which branch points define subgroups of equal size (Heard 1992), that has direct effects on PD estimation and inferences of trait evolution (Table 3.1b; Vellend et al. 2011, Mazel et al. 2017). The tree shape, therefore, influences the statistical distribution of phylogenetic distances, which are often left skewed, while traits are much more likely to be normally distributed (Cadotte et al. 2017, 2019). Of course there some exceptions, for example, herbaceous climbers in Central European herbaceous habitats have extremely large seeds compared to co-occurring plants (Bitomský et al. 2019, Appendix B), which resulted in the right skewed distribution of seed mass. Anyway, it is not clear how distributional differences between phylogenetic and traits distances will influence comparisons across different numbers of species and communities (Cadotte et al. 2019).

So far, I have reviewed how considering evolutionary models can inform our understanding of the PD-FD relationship; however, this set of phylogenetic comparative methods (PCM) also has several known statistical limitations (Freckleton 2009, Ho & Ané 2014, Cooper et al. 2016a,b). Phylogenies could be uninformative and some parameters cannot be inferred precisely, especially under the OU model when the selection strength is small (Table 3.1c; Boettiger et al. 2012). For the OU model, even using large phylogenies does not ensure that the selection optimum can be consistently estimated (Ho & Ané 2013). To increase precision in the estimated selection optimum, the presence of fossil taxa is crucial (Ho & Ané 2014), but the inclusion of fossil taxa in phylogenies for PCM is a challenging task that is usually avoided (Table 3.1d). Finally, robustness of PCM should be quantified in order to be confident about the conclusions (Paterno et al. 2018).

Other considerations include whether PCMs are influenced by polytomies that can inflate rates of evolution and bias models of evolution (Cooper & Purvis 2010). Further, some authors question the traditional methods for model selection, such as Akaike's or Bayesian information criterion (AIC and BIC, respectively; Table 3.1e). For example, standard BIC tends to select overly simple models under BM evolution (Ané 2008), while both AIC and BIC favour overparametrised models when applied to OU models (Boettiger et al. 2012, Cooper et al. 2016b). Limitations of PCM are often inadequately assessed in empirical studies leading to misinterpreted results and poor model fits, but some less technical guidelines can be found in the literature (Cooper et al. 2016b) to minimalise the effect of wrongly used methods. In summary, research efficiency in community phylogenetics would gain from routine correct using of PCM to understand how traits and phylogenetic relationships are linked together.

CHAPTER FOUR

PHYLOGENY-TRAIT-PRODUCTIVITY LINKS

4.1 State of the field

In the face of the current environmental crisis, phylogenetic relatedness and relevant traits might be very useful to predict changes in the vegetation threatened by eutrophication (nutrient enrichment). Nutrient addition in ecosystems usually increases productivity, i.e. a gradient indicating light limitation and the intensity of competition for light. There are several variables that are usually used to proxy the productivity gradient: aboveground biomass (Cadotte et al. 2008, 2009; Roscher et al. 2012, Liu et al. 2015), soil resource availability (Adler et al. 2013), plant cover (Lhotsky et al. 2016), or combinations of these (Manning et al. 2009, Guerrero et al. 2014, Bergholz et al. 2015). There are other possibilities, such as water availability (Harel et al. 2011, May et al. 2013) in arid systems or Ellenberg indicator values for nutrients or light (Santini et al. 2017). Higher productivity inevitably results in increased light limitation and the intensity of competition for light, which is one of the most important drivers of plant biodiversity loss (Hautier et al. 2009, Borer et al. 2014; but see Laliberté et al. 2014 for an alternative hypothesis). It is likely that species loss will be accompanied with the loss of phylogenetic and functional diversity. To evaluate this hypothesis, I focus on some links between plant evolutionary history, traits and variables related to productivity (aboveground biomass, plant cover or nutrient availability).

Researchers have tried to understand how biomass correlates with a number of biodiversity metrics and found that combinations of PD and FD measures predict changes of plant community biomass very well (e.g. Cadotte et al. 2008, 2009; Roscher et al. 2012, Liu et al. 2015). Several explanations why is PD a good indicator of productivity have been proposed. Most of them are based on an assumption that phylogenetic distances reflect multidimensional trait and niche differences (discussed in the section 3.3). For example, plant productivity increases when plant species in symbiotic relationships with N-fixing bacteria (a strong phylogenetic signal to N-fixing symbioses) co-occur with plant species without such relationships (Tilman et al. 1997, Cadotte et al. 2009). Other facilitative mechanisms are expected to be more likely between distantly related species (Valiente-Banuet & Verdú 2013), e.g. nutrient mobilisation between Fe/Zn mobilising graminoids and P-mobilising dicots (Li et al. 2014). These perspectives suggest why we often observe a positive relationship between PD and aboveground biomass.

Functional differences among plants species occurring along the gradient of productivity can be explained by several key axes of specialisation that include the leaf economy spectrum, plant size and dispersal ability (Westoby 1998, Westoby et al. 2002, Wright et al. 2004). A very influential approach to proxy these axes is the use of the leaf-height-seed (LHS) plant ecology strategy scheme (Westoby 1998), which is based on three plant traits: specific leaf area (SLA), canopy height and seed mass (see a trait measurement protocol for more details on these traits, Pérez-Harguindeguy et al. 2016). The theory predicts that species favouring productive conditions are taller, have higher SLA and large seeds. Canopy height is a trait representing shade avoidance and the ability to compete for light, while both high SLA and large seeds are trait-states linked to shade tolerance as they increase survival under shady conditions thanks to enhanced carbon gain and well-provisioned seedlings (Westoby 1998, Evans & Poorter 2001, Valladares & Niinemets 2008, Manning et al. 2009, Muller-Landau 2010, Gommers et al. 2013, Bitomský et al. 2018, 2019; DeMalach et al.

2019). On the other hand, recent findings suggest that clonal and bud bank traits are ecologically more relevant for herbs and herbaceous habitats (Klimešová et al. 2016), which might have implications for the assembly processes along the gradient of productivity. Indeed, clonal and bud bank traits can explain species positions along the productivity gradient (Herben et al. 2018). Apparently, many traits are involved in species responses to productivity, therefore, phylogenetic information might be needed to proxy these multidimensional trait and ecological differences.

Nutrient addition is globally responsible for species loss (DeMalach et al. 2017) due to increased competition for light (Hautier et al. 2009), but we still need to understand how FD and PD are affected. Intuitively, FD and PD should decline with decreasing species richness. For example, Zemunik et al. (2015) studied species richness (SR) and FD of resource-acquisition strategies along a soil chronosequence (representing soil development and fertility) and found a decrease of both SR and FD with increasing soil fertility. In addition, phosphorus-impoverished soils in Australia, South Africa (fynbos vegetation) or Brazilian rupestrian grasslands (*campos rupestres*) are typically hyperdiverse in terms of SR and FD (Lambers et al. 2008, Zemunik et al. 2015, 2018; Verboom et al. 2017). On the other hand, Yang et al. (2018) observed an increase of PD after N addition in a temperate semi-arid steppe as this treatment promoted the colonisation of species distantly related to the residents. Many grassland experiments do not focus on PD measures but report changes in phylogenetic structure (community clade composition). For example, graminoids, especially sedges, benefit from N addition and are more frequent in N-rich sites (Bassin et al. 2012, Zhao et al. 2019), legumes are more P- and K-demanding (Divito & Sadras 2014, Ren et al. 2017), while other super-rosids and super-asterids prefer sites with higher plant-available Ca and Mg in the soil (personal observation).

4.2 Analytical approaches

Measures of functional diversity can be summarised via an analogous framework as in PD (Mason et al. 2005), i.e. functional richness, divergence and regularity (evenness). Additionally, one of the most popular approaches is the concept of community-weighted mean (CWM), i.e. community mean of a trait weighted by species abundances, which is a measure of functional composition. There is a plethora of univariate and multivariate methods testing the trait-environment relationships (Lavorel et al. 2008, Kleyer et al. 2012). Values of CWM are then regressed (or correlated) against an environmental gradient. A similar procedure is done with various PD metrics (raw or null-model corrected; see Tucker et al. 2017 for a summary), which are also regressed against a gradient to infer a relationship between PD and an environment variable. In general, the PD-FD-productivity links are usually studied using quite conventional statistical methods, such as linear models, multivariate methods, RLQ (Dolédec et al. 1996) or fourth-corner (Legendre et al. 1997). Much more rare are studies that directly integrate phylogenetic and functional diversity metrics and/or partitioning variation in phylogenetic or functional diversity into the components that can be explained by external factor (e.g. spatial and environmental gradients) and their covariation (but see Liu et al. 2013, Swenson 2014).

Several R packages have been develop to analyse trait data and changes of FD and PD along environmental gradients. The package *FD* (Laliberté & Legendre 2010) has some basic tools for functional ecology, *ade4* (Dray & Dufour 2007) can be used to test trait-environment links, while *phyr* (Ives et al. 2019) provides some advanced tools to fit phylogenetic generalised linear mixed models or to estimate correlations among functional traits.

4.3 Challenges

Several recent studies have pointed out several critical issues with some routinely used methods to test the trait-environment and phylogeny-environment relationships (Zelený & Schaffers 2012, Hawkins et al. 2017, Peres-Neto et al. 2017, Duarte et al. 2018, Laughlin et al. 2018, Zelený 2018). Zelený and Schaffers (2012) showed that analyses relating mean Ellenberg indicator values (EIV) with other variables derived from the species composition produce highly biased results due to the similarity issue (EIVs inherit information about compositional similarity as they are weighted by species abundances). In my opinion, this is an expected issue as both community and environmental features are inferred from vegetation data in the EIV approach; however, Hawkins et al. (2017) showed that the whole "community-focus" approach suffers from a potentially severe structural problem driven by repeated species co-occurrences. They argue any metric (e.g. functional or phylogenetic), which is generated at the assemblage/community level by assigning values to species and averaging them for the species present within a cell/plot, can have internal statistical relationships of no biological significance across communities (Hawkins et al. 2017). Finally, Peres-Neto et al. (2017) found that the CWM approach can produce inconsistent parameter estimation and inflated Type I error rates (Table 4.1a). In other words, many studies might have provided too optimistic statistical support for the relationship between FD, PD measures and environmental gradients when, in fact, there is no relationship. This is a very disturbing fact given the commonness of all criticised methods (e.g. 2420 studies in *Google Scholar* on 18th March when submitted "community-weighted mean").

For the trait-environment relationship, there are some solutions to control for the type I error, such as the fourth-corner correlation (based on permutations, Peres-Neto et al. 2017), the parametric max test (ter Braak et al. 2018) or a method to control for phylogenetic autocorrelation (Duarte et al. 2018). Besides statistical reasons, the trait-environment relationship must be interpreted with caution. For example, Laughlin et al. (2018) observed inconsistency between the CWM approach and the analysis of species survival regarding the relationship between SLA and soil fertility. While the CWM of SLA was not correlated with soil fertility, survival was highest for species with low SLA in infertile soil, suggesting that the CWM approach is unreliable and cannot be taken as the evidence of the adaptive value of traits. Therefore, analyses of the trait-environment relationship motivated by biological reasoning might be efficient (Table 4.1b). For example, we applied this approach in our study of the CWM of seed mass along the gradient of light limitation (crop cover) in arable fields (Bitomský et al. 2018, Appendix A). Large seeds are expected to endure stressful shading (see the section 4.1). We showed that the community seed mass was driven by the light limitation gradient only in the case of annual plants (positive relationship). Perennials often prioritise many other reproduction strategies (e.g., clonal propagation) or survive under unsuitable conditions in structures other than seeds; thus, seed mass is likely to be ecologically relevant only for annuals. Some future studies need to consider that the trait-environment relationship can be different for different groups (e.g. functional groups or clades).

When studying changes of average trait values along environmental gradients, we ignore within-site trait variability that could be potentially more relevant. Thus, even if there is a significant relationship between a trait and an environmental gradient, in fact, it can only describe a portion of the total trait variability as it ignores the within-site component (Table 4.1c). Trait variation among species within community has gained much interest and has been shown to be relatively higher in various traits than the among-site component (e.g. Gallagher & Leishman 2012, Pescador et al. 2015, Peres-Neto et al. 2017). To acknowledge this, Ackerly & Cornwell (2007) developed a simple trait partitioning framework of species' trait values into alpha and beta components: beta values refer to a species' position along a gradient defined by community-level mean trait values; alpha values are the difference between a species' trait values and the mean of co-occurring taxa. We used this approach to test the trait differences between herbaceous climbers and co-occurring herbs (Bitomský et al. 2019, Appendix B) and found that vines tended to have higher canopy height, seed mass and SLA. In my opinion, it is better to compare trait syndromes between functional groups or clades in terms of co-occurring species because assessments across habitats introduce another source of variation that may simply reflect among-habitat differences and bias conclusions.

Table 4.1 Summary of the most important considerations and issues that affect the PD-FD-productivity links.

Issue	Description	Source
a) Type I error rates	Intrinsic sample attributes (derived from a species composition matrix) or phylogenetic autocorrelation affect type I error rates in some popular approaches for testing the PD-FD changes along environmental gradients.	Hawkins et al. (2017) Peres-Neto et al. (2017) Duarte et al. (2018) ter Braak et al. (2018) Zelený (2018)
b) Biological reasoning	Ecological relevancy of a trait can differ among functional groups or clades. The trait-environment relationship can provide unreliable estimates of the adaptive value of a trait.	Bitomský et al. (2018) Laughlin et al. (2018)
c) Ecological relevancy of the within- and among-site components	One should check whether the trait of interest has higher within- or among-site variation in order to focus on ecologically more relevant trait component.	Ackerly & Cornwell (2007) Pescador et al. (2015) Peres-Neto et al. (2017) Bitomský et al. (2019, in review)
d) Conclusions based on methods fitting the centers of the distributions of response variables may be limiting.	Assumptions of conventional correlation and regression analyses fundamentally conflict with the ecological concept of limiting factors. Estimating a range of regression quantiles of FD and PD metrics provides a comprehensive description of biological response patterns.	Cade et al. (1999) Cade & Noon (2003)

Various measures of functional divergence are used to assess the within-site component but both FD and CWM approaches are not additive (but see Ackerly & Cornwell 2007). Thus, Peres-Neto et al. (2017) developed a framework decomposing variation into the percentage within- and among-site components. We used this method in our study of variation of plant mycorrhizal status in herbaceous habitats (Bitomský et al. in review, Appendix D). We found that within-site variation in mycorrhizal status was always at least 2.2 times (but mostly nine times) larger than among-site variation. This is a bit surprising result considering the commonly accepted ecological theory that relates the ability to form mycorrhizal symbiosis to habitat and environmental factors (e.g. Gerz et al. 2016, 2019; Brundrett & Tedersoo 2018). Intuitively, one should, therefore, expect larger variation among sites differing in environmental conditions. This suggests that we should check whether the trait variation of interest (either within- or among-site) is ecologically relevant.

Finally, I believe that the consideration of alternative statistical tools might improve the research efficiency in phylogenetic and functional ecology. For example, quantile regression is a promising, but a bit overlooked, tool in ecology (Cade et al. 1999, Cade & Noon 2003) that has several advantages compared to classic linear models (and other derived methods) used to test the PD-FD-environment links (Table 4.1d). Conventional linear models focus on changes through the center of the distribution of the response variable. When there are other unmeasured variables that potentially limit FD or PD regardless of productivity (e.g. climate or land use), than the estimates based on linear models might be inaccurate (Cade et al. 1999). On the other hand, upper quantiles are more consistent with the ecological theory of limiting factors (such as nutrient availability or light limitation) than estimates through the center of data distributions (Cade et al. 1999).

In our study (Bitomský et al., unpublished), we suggest a rigorous testing procedure of the trait-environment relationship using an example with mycorrhizal status and nutrient availability (P and N:P). We combined the max test (ter Braak et al. 2018) and quantile regression (Fig. 4.1). The max test is a powerful parametric approach assessing the significances of two linear models, one community-level test as in the CWM approach and one species-level test that fits the environmental means of species, such as species niche centroids (SNC; Ackerly et al. 2002, Peres-Neto et al. 2017), against the trait (ter Braak et al. 2018). The conclusion about the trait-environment relationship is simply based on higher P-value, therefore, both tests must be significant in order to safely conclude that mycorrhizal status is linked to nutrient availability. The max test approach is then extended to quantile regression as the slope estimates of each quantile in both CWM and SNC regressions are compared. Possibly, this approach can be even extended to account for spatial correlation (for the CWM plot level) or phylogenetic correlation (for the SNC species level), but statistical performance of this method still needs to be thoroughly assessed.

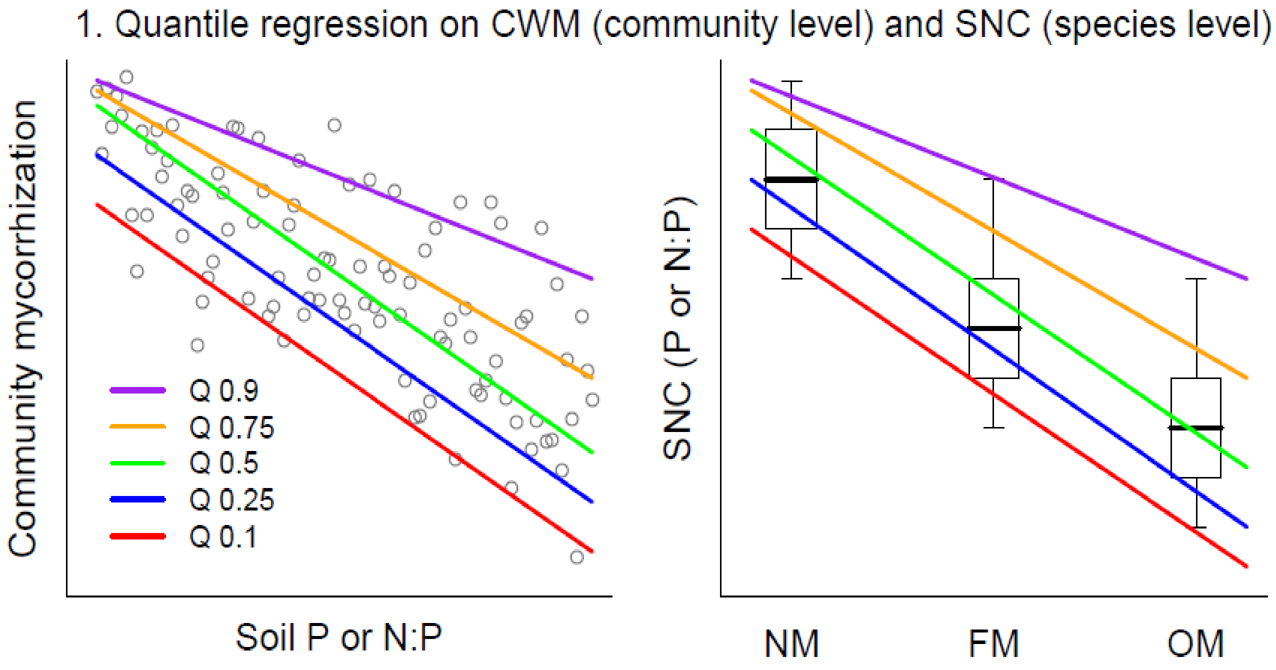


Fig. 4.1 Combination of the parametric max test (ter Braak et al. 2018) and quantile regression to rigorously test the trait-environment relationship (an example with mycorrhizal status and nutrient availability). In our study (Bitomský et al. unpublished) we tested the hypothesis that plants with the ability to form mycorrhiza will be more frequent in nutrient-poor sites, while non-mycorrhizal plants will be more frequent in nutrient-rich sites in various European herbaceous habitats. First figure describes quantile regression at the community level, i.e. the CWM of mycorrhizal status (community mycorrhization sensu Moora 2014) on soil P or N:P. Second figure refers to quantile regression at the species level, i.e. SNC values fitted against the trait. This testing controls for the type I error thanks to the max test approach as both regressions must be significant (ter Braak et al. 2018), while quantile regression enables an assessment of effects across suitable quantiles. Especially, upper quantiles should be of interest (Cade et al. 1999). Mycorrhizal status is coded as 0 (non-mycorrhizal plants: NM), 0.5 (facultative mycorrhiza: FM) and 1 (obligatory mycorrhiza: OM).

CHAPTER FIVE

TOWARDS THE ECOLOGICAL THEORY OF EVERYTHING VIA A PHYLOGENETIC FRAMEWORK: A SYNTHESIS

5.1 Overview

Incorporating phylogenies into ecology is usually due to two reasons: either one wants to use it to study phylogenetic patterns in communities, or one wants to filter its confounding effect. The former is thoroughly discussed in this thesis. The latter, so-called "phylogenetic corrections", is more like a statistical problem that undermines the independence of the data (a crucial assumption of most conventional statistical methods) because the degree of phylogenetic relatedness might indicate the degree of correlation in species measurements (a trait, an environmental response etc.) among species pairs. How to deal with this phylogenetic dependence has been intensively discussed from many angles in the ecological and evolutionary literature (Felsenstein 1985, Westoby et al. 1995, de Bello et al. 2015, Li & Ives 2017). Thus, it is clear that an ecological unifying framework will not work without phylogenetic information because it is needed due to both biological and statistical reasons.

A good ecological unifying framework needs to be based on a strong methodology in order to produce robust predictions and useful information for conservation decisioning and community ecology. However, biased methods produce biased results. To deal with this issue, it is necessary to understand how developed methods perform in different situations. Here, I primarily reviewed analytical approaches and discussed their performance using data and simulations from recent methodological and case studies. First, PD estimates depend on tree properties (taxon sampling, balance or branch lengths units) and phylogenetic patterns in vegetation can be lineage-specific (Chapter Two). In the Chapter Three, I reviewed what is known about the relationship between traits and phylogenetic relatedness, and advocated the use of phylogenetic comparative methods in ecology (and their common pitfalls). Finally, I briefly described phylogenetic and functional patterns in vegetation along the gradient of productivity, and highlighted analytical issues with some widely used approaches in functional and phylogenetic ecology.

An ideal scheme of an analysis in phylogenetic and functional ecology could look like the following. First of all, we need to rely on rigorous phylogenetic analyses (see Anisimova et al. 2013) to generate best phylogenies possible with correct phylogenetic distances among studied species. Second, generated phylogenies need to be examined to check whether the PD-FD patterns truly reflect evolutionary processes but not just some inherent properties of the tree (sections 2.3 and 3.3). Third, statistical and phylogenetic comparative methods can be used to infer the PD-FD relationship or trait evolutionary trajectories (section 3.3). Finally, if one is interested in the role of environmental gradients, I would suggest to use a method that has a better performance and controls for the type I error because some popular approaches are a bit outdated and can produce too optimistic results (section 4.3). This proposed scheme can be a powerful (but certainly time-consuming) tool for predicting alterations in species composition, functional and phylogenetic structure or ecosystem functioning in vegetation threatened by eutrophication and following changes in productivity.

5.2 Major questions

To avoid redundancy, I refer the reader to recent reviews proposing general questions and directions for future research related to the topics discussed here, e.g. Winter et al (2013) regarding the concept of PD in nature conservation and the PD-FD relationship, Srivastava et al. (2012) regarding PD and functioning of ecosystems, Weber and Agrawal (2012) regarding the testing of links between traits and phylogenies, Segar et al. (2020) regarding the role of evolution in shaping ecological interactions among species or Estrada et al. (2016) regarding the usefulness of traits in predicting the impact of environmental changes. In this respect, I think the research agenda should also focus on the following questions:

(i) Is the concept of PD even useful in nature conservation and community ecology? Despite so many studies stressing the usefulness of PD in nature conservation, Winter et al. (2013) claimed that the current knowledge will not convince conservationists and policymakers to apply the concept of phylogenetic diversity in nature conservation because it still depends on many assumptions, uncertainties, and varying messages. Later studies suggest that the most prominent hypothesis that maximising PD maximises FD is a risky conservation tool (Mazel et al. 2017, 2018). A lot of assumptions about phylogenetic patterns in communities has been questioned too (Gerhold et al. 2015). In their reviews, Gerhold et al. (2015, 2018) argue that patterns in PD are not the result of present assembly processes, but they rather reflect how the PD concept explains some macroevolutionary patterns in species pools and local coexistence. This suggests that maybe we have been just asking wrong questions so far.

(ii) Can the PD-FD relationship be summarised using just several types of lines/curves? Can we describe all complex patterns of trait evolution using just few models? Various traits have different evolutionary trajectories that can be fitted using macroevolutionary models. Since the number of evolutionary models is finite, we should expect the finite number of lines and curves fitting the PD-FD relationship. This would be a huge step forward because we would have a summarising and valid framework of the PD-FD links that would work across clades and functional groups. Among other things, this set of predictions can be then used to address the first question. For example, stabilising selection (OU model) suggests a logarithmic increase of FD with increasing PD (Tucker et al. 2018). In this situation, maximising PD is useless because FD is quickly saturated and an increase of FD is negligible after some threshold PD value. Such mechanistic explanations are desperately needed to help us how to use the concept of PD appropriately.

(iii) Can we get robust predictions of shifts of phylogenetic structure (or diversity) in vegetation due to eutrophication? Are phylogenetically distinct groups more threatened (possibly due to some phylogenetically conserved nutrient-acquisition strategies)? Can we expect omnipresent phylogenetic homogenisation due to eutrophication? To the best of my knowledge, there are still no global predictions regarding the loss of evolutionary history (in comparison with species and functional losses) due to nutrient enrichment.

REFERENCES

Ackerly D & Cornwell WK (2007). A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecology Letter*s 10: 135–145.

Ackerly D et al. (2002). Leaf size, specific leaf area and microhabitat distribution of chaparral woody plants: contrasting patterns in species level and community level analyses. *Oecologia* 130: 449–457.

Adler PB et al. (2013). Trait-based tests of coexistence mechanisms. *Ecology Letters* 16: 1294–1306.

Ané C (2008). Analysis of comparative data with hierarchical autocorrelation. *Annals of Applied Statistics* 2: 1078–1102.

Anisimova M et al. (2013). State-of the art methodologies dictate new standards for phylogenetic analysis. *BMC Evolutionary Biology* 13: 161.

Bassin S et al. (2012). Different types of sub-alpine grassland respond similarly to elevated nitrogen deposition in terms of productivity and sedge abundance. *Journal of Vegetation Science* 23: 1024–1034.

Bergholz K et al. (2015). Fertilization affects the establishment ability of species differing in seed mass via direct nutrient addition and indirect competition effects. *Oikos* 124: 1547–1554.

Bitomský M et al. (2018). Light limitation shapes the community seed mass of annual but not of perennial weeds. *Perspectives in Plant Ecology, Evolution and Systematics* 31: 1–6. **[APPENDIX A]**

Bitomský M et al. (2019). Herbaceous climbers in herbaceous systems are shade-tolerant and magnesium-demanding. *Journal of Vegetation Science* 30: 799–808. **[APPENDIX B]**

Bitomský M et al. (2020). Clade composition of a plant community indicates its phylogenetic diversity. *Ecology and Evolution* (early view). doi: 10.1002/ece3.6170 **[APPENDIX C]**

Bitomský M et al. (in review). Variation in mycorrhizal status is much larger within than among sites: a phylogenetically-informed assessment across grassland and coastal floras. **[APPENDIX D]**

Blomberg SP et al. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57: 717–745.

Boettiger C et al. (2012). Is your phylogeny informative? Measuring the power of comparative methods. *Evolution* 66: 2240–2251.

Borer ET et al. (2014). Herbivores and nutrients control grassland plant diversity via light limitation. *Nature* 508: 517–520.

Borges R et al. (2018). Measuring phylogenetic signal between categorical traits and phylogenies. *Bioinformatics* 35: 1862–1869.

Britton T et al. (2007). Estimating divergence times in large phylogenetic trees. *Systematic Biology* 56: 741–752.

Brundrett MC (2009). Mycorrhizal associations and other means of nutrition of vascular plants: understanding the global diversity of host plants by resolving conflicting information and developing reliable means of diagnosis. *Plant and Soil* 320: 37–77.

Brundrett MC & Tedersoo L (2018). Evolutionary history of mycorrhizal symbioses and global host plant diversity. *New Phytologist* 220: 1108–1115.

Brundrett MC & Tedersoo L (2019). Misdiagnosis of mycorrhizas and inappropriate recycling of data can lead to false conclusions. *New Phytologist* 221: 18–24.

Burns JH & Strauss SY (2011). More closely related species are more ecologically similar in an experimental test. *Proceedings of the National Academy of Sciences of the United States of America* 108: 5302–5307.

Burns JH et al. (2018). Tests of alternative evolutionary models are needed to enhance our understanding of biological invasions. *New Phytologist* 222: 701–707.

Butler MA & King AA (2004). Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist* 164: 683–695.

Cade BS et al. (1999). Estimating effects of limiting factors with regression quantiles. *Ecology* 80: 311–323.

Cade BS & Noon BR (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment* 1: 412–420.

Cadotte MW et al. (2008). Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences of the United States of America* 105: 17012–17017.

Cadotte MW et al. (2009). Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *PloS One* 4: e5695.

Cadotte MW et al. (2011). Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology* 48: 1079–1087.

Cadotte MW et al. (2013). The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology Letters* 16: 1234–1244.

Cadotte MW (2015). Phylogenetic diversity–ecosystem function relationships are insensitive to phylogenetic edge lengths. *Functional Ecology* 29: 718–723.

Cadotte MW et al. (2017). Why phylogenies do not always predict ecological differences. *Ecological Monographs* 87: 535–551.

Cadotte MW et al. (2018). Preadaptation and naturalization of nonnative species: Darwin’s two fundamental insights into species invasion. *Annual Review of Plant Biology* 69: 661–684.

Cadotte MW et al. (2019). Do traits and phylogeny support congruent community diversity patterns and assembly inferences? *Journal of Ecology* 107: 2065–2077.

Campbell N et al. (2010). Species richness, taxonomic diversity, and taxonomic distinctness of the deep-water demersal fish community on the Northeast Atlantic continental slope (ICES Subdivision Via). *ICES Journal of Marine Science* 68: 365–376.

Cavender-Bares J et al. (2004). Phylogenetic overdispersion in Floridian oak communities. *The American Naturalist* 163: 823–843.

Cavender-Bares et al. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters* 12: 693–715.

Chao A et al. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society, B: Biological Sciences* 365: 3599–3609.

Chao A et al. (2014). Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through Hill numbers. *Annual Review of Ecology, Evolution, and Systematics* 45: 297–324.

Clarke KR & Warwick RM (2001). A further biodiversity index applicable to species lists: Variation in taxonomic distinctness. *Marine Ecology Progress Series* 216: 265–278.

Cooper N & Purvis A (2010). Body size evolution in mammals: complexity in tempo and mode. *The American Naturalist* 175: 727–738.

Cooper N et al. (2016a). A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. *Biological Journal of the Linnean Society* 118: 64–77.

Cooper N et al. (2016b). Shedding light on the 'dark side' of phylogenetic comparative methods. *Methods in Ecology and Evolution* 7: 693–699.

de Bello F et al. (2015). On the need for phylogenetic 'corrections' in functional trait-based approaches. *Folia Geobotanica* 50: 349–357.

de Bello F et al. (2017). Decoupling phylogenetic and functional diversity to reveal hidden signals in community assembly. *Methods in Ecology and Evolution* 8: 1200–1211.

DeMalach N et al. (2017). Contrasting effects of water and nutrient additions on grassland communities: A global meta-analysis. *Global Ecology and Biogeography* 26: 983–992.

DeMalach N et al. (2019). Mechanisms of seed mass variation along resource gradients. *Ecology Letters* 22: 181–189.

Descombes P et al. (2017). Uneven rate of plant turnover along elevation in grasslands. *Alpine Botany* 127: 53–63.

Divito GA & Sadras VO (2014). How do phosphorus, potassium and sulphur affect plant growth and biological nitrogen fixation in crop and pasture legumes? A meta-analysis. *Field Crops Research* 156: 161–171.

Dolédec S et al. (1996). Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics* 3: 143–166.

Dray S & Dufour A (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22: 1–20.

Duarte LDS et al. (2012). Assessing spatial and environmental drivers of phylogenetic structure in Brazilian *Araucaria* forests. *Ecography* 35: 952–960.

Duarte LDS et al. (2018). Analyzing community-weighted trait means across environmental gradients: should phylogeny stay or should it go? *Ecology* 99: 385–398.

Durka W & Michalski SG (2012). Daphne: a dated phylogeny of a large European flora for phylogenetically informed ecological analyses. *Ecology* 93: 2297.

Elliott MJ et al. (2016). Choice between phylogram and chronogram can have a dramatic impact on the location of phylogenetic diversity hotspots. *Journal of Biogeography* 45: 2190–2201.

Elliott TL et al. (2016). Contrasting lineage-specific patterns conceal community phylogenetic structure in larger clades. *Journal of Vegetation Science* 27: 69–79.

Elton C (1946). Competition and the structure of ecological communities. *Journal of Animal Ecology* 15: 54–68.

Estrada A et al. (2016). Usefulness of species traits in predicting range shifts. *Trends in Ecology & Evolution* 31: 190–203

Evans JR & Poorter H (2001). Photosynthetic acclimation of plants to growth irradiance: the relative importance of specific leaf area and nitrogen partitioning in maximizing carbon gain. *Plant, Cell & Environment* 24: 755–767.

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

Faith DP et al. (2010). Evosystem services: an evolutionary perspective on the links between biodiversity and human well-being. *Current Opinion in Environmental Sustainability* 2: 66–74.

Faith DP (2018). Phylogenetic diversity and conservation evaluation: perspectives on multiple values, indices, and scales of application. In: Scherson RA & Faith DP (eds.). *Phylogenetic diversity: applications and challenges in biodiversity science*. pp. 1–26, Springer.

Felsenstein J (1985). Phylogenies and the comparative method. *The American Naturalis*t 125: 1–15.

Felsenstein J (2005). Using the quantitative genetic threshold model for inferences between and within species. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 1427–1434.

Forest F et al. (2007). Preserving the evolutionary potential of floras in biodiversity hotspots. *Nature* 445: 757–760.

Freckleton RP (2009). The seven deadly sins of comparative analysis. *Journal of Evolutionary Biology* 22: 1367–1375.

Fritz SA & Purvis A (2010). Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conservation Biology* 24: 1042–1051.

Gallagher RV & Leishman MR (2012). Contrasting patterns of trait-based community assembly in lianas and trees from temperate Australia. *Oikos* 121: 2026–2035.

Gerhold P et al. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology* 29: 600–614.

Gerhold P et al. (2018). The deep past controls the phylogenetic structure of present, local communities. *Annual Review of Ecology, Evolution, and Systematics* 49: 477–497.

Gerz M et al. (2016). Plant community mycorrhization in temperate forests and grasslands: relations with edaphic properties and plant diversity. *Journal of Vegetation Science* 27: 89–99.

Gerz M et al. (2019). Responses of plant community mycorrhization to anthropogenic influence depend on the habitat and mycorrhizal type. *Oikos* 128: 1565–1575.

Gommers CMM et al. (2013). Shade tolerance: When growing tall is not an option. *Trends in Plant Science* 18: 65–71.

Graham CH et al. (2018). Phylogenetic scale in ecology and evolution. *Global Ecology and Biogeography* 27: 175–187.

Guerrero I et al. (2014). Non-linear responses of functional diversity and redundancy to agricultural intensification at the field scale in Mediterranean arable plant communities. *Agriculture, Ecosystems & Environment* 195: 36–43.

Hall SJ & Greenstreet SP (1998). Taxonomic distinctness and diversity measures: responses in marine fish communities. *Marine Ecology Progress Series* 166: 227–229.

Hansen TF & Martins EP (1996). Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* 50: 1404–1417.

Hansen TF (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51: 1341–1351.

Harel D et al. (2011). Seed mass and dormancy of annual plant populations and communities decreases with aridity and rainfall predictability. *Basic and Applied Ecology* 12: 674–684.

Harmon LJ (2019). *Phylogenetic comparative methods: learning from trees*. Self published under a CC-BY-4.0 license. https://lukejharmon.github.io/pcm/

Hautier Y et al. (2009). Competition for light causes plant biodiversity loss after eutrophication. *Science* 324: 636–638.

Hawkins BA et al. (2017). Structural bias in aggregated species-level variables driven by repeated species co-occurrences: a pervasive problem in community and assemblage data. *Journal of Biogeography* 44: 1199–1211.

Heard SB (1992). Patterns in tree balance among cladistic, phenetic, and randomly generated phylogenetic trees. *Evolution* 46: 1818–1826.

Hempel S et al. (2013). Mycorrhizas in the Central European flora: relationships with plant life history traits and ecology. *Ecology* 94: 1389–1399.

Herben T et al. (2018). Effects of disturbance frequency and severity on plant traits: An assessment across a temperate flora. *Functional Ecology* 32: 799–808.

Hill MO (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology* 54: 427–431.

Hinchliff CE et al. (2015). Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences of the United States of America* 13: 12764–12769.

Hipp AL et al. (2018). Effects of phylogenetic diversity and phylogenetic identity in a restoration ecology experiment. In: Scherson RA & Faith DP (eds.). *Phylogenetic diversity: applications and challenges in biodiversity science*. pp. 189–210, Springer.

Ho LST & Ané C (2013). Asymptotic theory with hierarchical autocorrelation: Ornstein-Uhlenbeck tree models. *Annals of Statistics* 41: 957–981.

Ho LST & Ané C (2014). Intrinsic inference difficulties for trait evolution with Ornstein-Uhlenbeck models. *Methods in Ecology and Evolution* 5: 1133–1146.

Isaac NJB & Pearse WD (2018). The use of EDGE (evolutionary distinct globally endangered) and EDGE-like metrics to evaluate taxa for conservation. In: Scherson RA & Faith DP (eds.). *Phylogenetic diversity: applications and challenges in biodiversity science*. pp. 27–39, Springer.

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

Ives AR & Garland T (2010). Phylogenetic logistic regression for binary dependent variables. *Systematic Biology* 59: 9–26.

Ives AR et al. (2019). phyr: model based phylogenetic analysis. https://CRAN.R-project.org/package=phyr

Janssens SB et al. (2020). A large-scale species level dated angiosperm phylogeny for evolutionary and ecological analyses. *Biodiversity Data Journal* 8: e39677.

Jantzen JR et al. (2019). Effects of taxon sampling and tree reconstruction methods on phylodiversity metrics. *Ecology and Evolution* 9: 9479–9499.

Keck F et al. (2016). phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution* 6: 2774–2780.

Kelly S et al. (2014). Phylogenetic trees do not reliably predict feature diversity. *Diversity and Distributions* 20: 600–612.

Kembel SW et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26: 1463–1464.

Kleyer M et al. (2012). Assessing species and community functional responses to environmental gradients: which multivariate methods? *Journal of Vegetation Science* 23: 805–821.

Klimešová J et al. (2016). Herbs are different: clonal and bud bank traits can matter more than leaf–height–seed traits. *New Phytologist* 210: 13–17.

Knapp S et al. (2008). Challenging urban species diversity: contrasting phylogenetic patterns across plant functional groups in Germany. *Ecology Letters* 11: 1054–1064.

Kraft NJB et al. (2007). Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist* 170: 271–283.

Kraft NJB et al. (2015). Plant functional traits and the multidimensional nature of species coexistence. *Proceedings of the National Academy of Sciences of the United States of America* 112: 797–802.

Laliberté E & Legendre P (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology* 91: 299–305.

Laliberté E et al. (2014). Environmental filtering explains variation in plant diversity along resource gradients. *Science* 345: 1602–1605.

Lambers H et al. (2008). Plant nutrient-acquisition strategies change with soil age. *Trends in Ecology & Evolution* 23: 95–103.

Lankau R et al. (2011). Incorporating evolutionary principles into environmental management and policy. *Evolutionary Applications* 4: 315–325.

Laughlin DC et al. (2018). Survival rates indicate that correlations between community-weighted mean traits and environments can be unreliable estimates of the adaptive value of traits. *Ecology Letters* 21: 411–421.

Lavergne S et al. (2010). Biodiversity and climate change: integrating evolutionary and ecological responses of species and communities. *Annual Review of Ecology and Systematics* 41: 321–350.

Lavorel S et al. (2008). Assessing functional diversity in the field – methodology matters! *Functional Ecology* 22: 134–147.

Legendre P et al. (1997). Relating behavior to habitat: solutions to the fourth-corner problem. *Ecology* 78: 547–562.

Letten AD & Cornwell WK (2015). Trees, branches and (square) roots: why evolutionary relatedness is not linearly related to functional distance. *Methods in Ecology and Evolution* 6: 439–444.

Lewis PO (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* 50: 913–925.

Lhotsky B et al. (2016). Changes in assembly rules along a stress gradient from open dry grasslands to wetlands. *Journal of Ecology* 104: 507–517.

Li D & Ives AR (2017). The statistical need to include phylogeny in trait-based analyses of community composition. *Methods in Ecology and Evolution* 8: 1192–1199.

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

Li D et al. (2019). For common community phylogenetic analyses, go ahead and use synthesis phylogenies. *Ecology* 100: e02788.

Li L et al. (2014). Plant diversity and overyielding: insights from belowground facilitation of intercropping in agriculture. *New Phytologist* 203: 63–69.

Linder HP et al. (2005). Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. *Molecular Phylogenetics and Evolution* 35: 569–582.

Liu X et al. (2013). The environment and space, not phylogeny, determine trait dispersion in a subtropical forest. *Functional Ecology* 27: 264–272.

Liu J et al. (2015). Explaining maximum variation in productivity requires phylogenetic diversity and single functional traits. *Ecology* 96: 176–183.

Maherali H et al. (2016). Mutualism persistence and abandonment during the evolution of the mycorrhizal symbiosis. *The American Naturalist* 188: E113-E125.

Manning P et al. (2009). Shifts in seed size across experimental nitrogen enrichment and plant density gradients. *Basic and Applied Ecology* 10: 300–308.

Mason NWH et al. (2005). Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos* 111: 112–118.

May F et al. (2013). Plant functional traits and community assembly along interacting gradients of productivity and fragmentation. *Perspectives in Plant Ecology, Evolution and Systematics* 15: 304–318.

Mayfield MM & Levine JM (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters* 13: 1085–1093.

Mazel F et al. (2017). Conserving phylogenetic diversity can be a poor strategy for conserving functional diversity. *Systematic Biology* 66: 1019–1027.

Mazel F et al. (2018). Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nature Communications* 9: 28888.

McGill BJ et al. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution* 21: 178–186.

Miller ET et al. (2017). Phylogenetic community structure metrics and null models: a review with new methods and software. *Ecography* 40: 461–477.

Molina-Venegas R & Rodríguez MÁ (2017). Revisiting phylogenetic signal; strong or negligible impacts of polytomies and branch length information? *BMC Evolutionary Biology* 17: 53.

Moora M (2014). Mycorrhizal traits and plant communities: perspectives for integration. *Journal of Vegetation Science* 25: 1126–1132.

Muller-Landau HC (2010). The tolerance-fecundity trade-off and the maintenance of diversity in seed size. *Proceedings of the National Academy of Sciences of the United States of America* 107: 4242–4247.

Münkemüller T et al. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution* 3: 743–756.

Ndiribe C et al. (2013). Phylogenetic plant community structure along elevation is lineage specific. *Ecology and Evolution* 3: 4925–4939.

Neves DM et al. (2020). Evolutionary diversity in tropical tree communities peaks at intermediate precipitation. *Scientific reports* 10: 1188.

O'Meara BC et al. (2006). Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60: 922–933.

Ovaskainen O et al. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters* 20: 561–576.

Park DS et al. (2018). Taxon sampling effects on the quantification and comparison of community phylogenetic diversity. *Molecular Ecology* 27: 1296–1308.

Paterno GB et al. (2018). sensiPhy: An r-package for sensitivity analysis in phylogenetic comparative methods. *Methods in Ecology and Evolution* 9: 1461–1467.

Pavoine S et al. (2013). Correlations between phylogenetic and functional diversity: mathematical artefacts or true ecological and evolutionary processes? *Journal of Vegetation Science* 24: 781–793.

Pennell MW et al. (2014). geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* 30: 2216–2218.

Peres-Neto PR et al. (2017). Linking trait variation to the environment: critical issues with community-weighted mean correlation resolved by the fourth-corner approach. *Ecography* 40: 806–816.

Pérez-Harguindeguy N et al. (2016). Corrigendum to: New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany* 64: 715–716.

Pescador DS et al. (2015). Plant trait variation along an altitudinal gradient in Mediterranean high mountain grasslands: controlling the species turnover effect. *PloS One* 10: e0118876.

Pinheiro J et al. (2019). nlme: Linear and nonlinear mixed effects models. https://CRAN.R-project.org/package=nlme>

Prinzing A et al. (2008). Less lineages – more trait variation: phylogenetically clustered plant communities are functionally more diverse. *Ecology Letters* 11: 809–819.

Purvis A et al. (2000). Nonrandom extinction and the loss of evolutionary history. *Science* 288: 328–330.

Redding DW & Mooers AO (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology* 20: 1670–1678.

Ren F et al. (2017). Phosphorus does not alleviate the negative effect of nitrogen enrichment on legume performance in an alpine grassland. *Journal of Plant Ecology* 10: 822–830.

Revell LJ (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.

Roscher C et al. (2012). Using plant functional traits to explain diversity-productivity relationships. *PloS One* 7: e36760.

Qian H & Jin Y (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology* 9: 233–239.

Sandel B (2018). Richness-dependence of phylogenetic diversity indices. *Ecography* 41: 837–844.

Sandel B et al. (2020). Current climate, isolation and history drive global patterns of tree phylogenetic endemism. *Global Ecology and Biogeography* 29: 4–15.

Santini BA et al. (2017). The triangular seed mass–leaf area relationship holds for annual plants and is determined by habitat productivity. *Functional Ecology* 31: 1770–1779.

Segar ST et al. (2020). The role of evolution in shaping ecological networks. *Trends in Ecology and Evolution* (in press). doi: 10.1016/j.tree.2020.01.004

Scheiner SM et al. (2017). The components of biodiversity, with a particular focus on phylogenetic information. *Ecology and Evolution* 7: 6444–6454.

Schulte JA (2013). Undersampling taxa will underestimate molecular divergence dates: An example from the South American lizard clade Liolaemini. *International Journal of Evolutionary Biology*: 628467.

Smith SA & Brown JW (2018). Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany* 105: 302–314.

Smith SE & Read DJ (2008). *Mycorrhizal symbiosis*. Academic Press.

Stamatakis A (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.

Srivastava DS et al. (2012). Phylogenetic diversity and the functioning of ecosystems. *Ecology Letters* 15: 637–648.

Suchard MA et al. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* 4: vey016.

Swenson NG (2009). Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PLoS One* 4: e4390.

Swenson NG (2014). *Functional and phylogenetic ecology in R*. Springer.

ter Braak CJF et al. (2018). Simple parametric tests for trait–environment association. *Journal of Vegetation Science* 29: 801–811.

Tilman D et al. (1997). The influence of functional diversity and composition on ecosystem processes. *Science* 277: 1300–1302.

Tsirogiannis C & Sandel B (2016). PhyloMeasures: a package for computing phylogenetic biodiversity measures and their statistical moments. *Ecography* 39: 709–714.

Tucker CM et al. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews* 92: 698–715.

Tucker CM et al. (2018). On the relationship between phylogenetic diversity and trait diversity. *Ecology* 99: 1473–1479.

Valiente-Banuet A & Verdú M (2013). Plant facilitation and phylogenetics. *Annual Review of Ecology and Systematics* 44: 347–366.

Valladares F & Niinemets Ü (2008). Shade tolerance, a key plant feature of complex nature and consequences. *Annual Review of Ecology, Evolution, and Systematics* 39: 237–257.

Vane-Wright RI et al. (1991). What to protect? Systematics and the agony of choice. *Biological Conservation* 55: 235–254.

Vellend M (2011). Measuring phylogenetic diversity. In: Magurran AE & McGill (eds.). *Biological diversity: frontiers in measurement and assessment.* pp. 193–206, Oxford University Press.

Verboom GA et al. (2017). Specialization to extremely low-nutrient soils limits the nutritional adaptability of plant lineages. *The American Naturalist* 189: 1–17.

Warwick RM & Clarke KR (1995). New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine Ecology Progress Series* 129: 301–305.

Webb CO et al. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics* 33: 475–505.

Weber MJ & Agrawal AA (2012). Phylogeny, ecology, and the coupling of comparative and experimental approches. *Trends in Ecology & Evolution* 27: 394–403.

Westoby M et al. (1995). On misinterpreting the phylogenetic correction. *Journal of Ecology* 83: 531–534.

Westoby M (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil* 199: 213–227

Westoby M et al. (2002). Plant ecological strategies: some leading dimensions of variation between species. *Annual Review of Ecology and Systematics* 33: 125–159.

Winter M et al. (2009). Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. *Proceedings of the National Academy of Sciences of the United States of America* 106: 21721–21725.

Winter M et al. (2013). Phylogenetic diversity and nature conservation: where are we? *Trends in Ecology and Evolution* 28: 199–204.

Wood SN (2017). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Wright IJ et al. (2004). The worldwide leaf economics spectrum. *Nature* 428: 821–827.

Yang X et al. (2018). Nitrogen fertilization, not water addition, alters plant phylogenetic community structure in a semi-arid steppe. *Journal of Ecology* 106: 991–1000.

Zanne AE et al. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature* 506: 89–92.

Zelený D & Schaffers AP (2012). Too good to be true: pitfalls of using mean Ellenberg indicator values in vegetation analyses. *Journal of Vegetation Science* 23: 419–431.

Zelený D (2018). Which results of the standard test for community-weighted mean approach are too optimistic? *Journal of Vegetation Science* 29: 953–966.

Zemunik G et al. (2015). Diversity of plant nutrient-acquisition strategies increases during long-term ecosystem development. *Nature Plants* 1: 15050.

Zemunik G et al. (2018). High abundance of non-mycorrhizal plant species in severely phosphorus-impoverished Brazilian *campos rupestres*. *Plant and Soil* 424: 255–271.

Zhao Y et al. (2019). Community composition, structure and productivity in response to nitrogen and phosphorus additions in a temperate meadow. *Science of the Total Environment* 654: 863–871.

Zheng YI et al. (2018). Species composition, functional and phylogenetic distances correlate with success of invasive *Chromolaena odorata* in an experimental test. *Ecology Letters* 21: 1211–1220.

Perspectives in Plant Ecology, Evolution and Systematics 31 (2018) 1–6

Contents lists available at ScienceDirect

Perspectives in Plant Ecology, Evolution and Systematics

journal homepage: www.elsevier.com/locate/ppees

Research article

Light limitation shapes the community seed mass of annual but not of perennial weeds

Martin Bitomský[a], Jan Mládek[a], Šárka Cimalová[b,*]

[a] Department of Ecology and Environmental Sciences, Palacký University, Šlechtitelů 241/27, CZ-783 71 Olomouc, Czech Republic
[b] Department of Biology and Ecology, University of Ostrava, Chittusiho 10, CZ-710 00 Ostrava, Czech Republic

ARTICLE INFO

Keywords:
Arable vegetation
Crop cover
Life span
Light competition
Productivity

ABSTRACT

How seed mass determines plant performance is dependent on the type of herbaceous community. Here, we hypothesized that life span is a decisive driver of the seed mass response to an environmental gradient. We examined whether separating community data into annuals and perennials sheds new light on seed mass importance in herbaceous weed vegetation of arable fields. We studied the seed mass response to a gradient of light limitation and tested the prediction that seed mass will increase with light limitation but that the trend will differ in annuals compared to perennials. In summary, only the seed mass of annuals reacted to the light limitation gradient. The seed mass community-weighted mean (CWM), controlled for crop type and seasonality, was positively linked with crop coverage, i.e., a proxy for light limitation. The seed mass CWM of perennials exhibited a random distribution. In annuals which are strongly dependent on seeds, large seeds are advantageous under dense crop canopies. We showed that considering the relevance of a trait for a particular strategy can improve our understanding of community assembly. This approach can help to explain some differences among published studies regarding the effect of an environmental gradient on the community seed mass.

1. Introduction

Studying the patterns of trait changes along environmental gradients is one of the major topics of trait-based ecology. In particular, regeneration traits such as seed mass are very often considered (Lhotsky et al., 2016; Metz et al., 2010; Santini et al., 2017). Seed mass is one of the three components of the LHS (leaf-height-seed) plant ecology strategy scheme (Westoby, 1998), a key trait in the dynamics of community structure (Leishman, 2001) and a part of the leading dimension of variation between species (Díaz et al., 2016; Westoby et al., 2002). In herbaceous vegetation, the role of seed mass in structuring communities has been thoroughly studied but the evidence is mixed (Coomes and Grubb, 2003; Fenner, 1985; Klimešová et al., 2016; Levine and Rees, 2002). The relevance of seed mass varies according to the type of vegetation and type of plants (summarized in Coomes and Grubb, 2003). This uncertainty of the importance of seed mass in herbs can be obscured by the significant proportion of perennial herbs relying on clonal propagation instead of seed production (Klimešová et al., 2016). In annuals, however, seed mass is a trait involved in all important aspects of plant ecology, such as dispersal, colonization, recruitment, tolerance of environmental hazards and competition.

Large-seeded plants are thought to have an advantage over small-seeded plants beyond seedling establishment (Metz et al., 2010); however, this advantage is dependent on site conditions (e.g., the level of shade, soil nutrients or disturbance). Light limitation (sensu Borer et al., 2014) is a key factor in plant communities (Hautier et al., 2009), and it is closely associated with site productivity. In line with the tolerance-fecundity trade-off (Muller-Landau, 2010), it has been hypothesized that seedlings from large-seeded plants better endure stressful shading (Bergholz et al., 2015; Leishman et al., 2000; Manning et al., 2009) because of at least two mechanisms. First, higher resource reserves provide a competitive advantage as plants can spend additional resources on shoot biomass. Large-seeded species, therefore, should be better competitors for light (Tilman, 1988). Second, seed reserves boost survival below denser canopies, i.e., plants with larger seeds tolerate lower levels of light in closed vegetation (Manning et al., 2009) or under litter (Lönnberg and Eriksson, 2013; Thompson, 1987). Nevertheless, Bergholz et al. (2015) recently suggested that the hypothesized positive relationship between community seed mass and plant canopy density is not linear but rather U shaped. If soil nutrients are considered (positively correlated with canopy density), these shifts appear to be the consequence of site productivity (competition for light) and the varying stress tolerance of seeds. On the one hand, large seeds can be advantageous under oligotrophic conditions (Adler et al., 2013;

---

* Corresponding author.
E-mail addresses: bitomsky.martin@gmail.com (M. Bitomský), jan.mladek@upol.cz (J. Mládek), sarka.cimalova@osu.cz (Š. Cimalová).

https://doi.org/10.1016/j.ppees.2017.11.005
Received 3 April 2017; Received in revised form 15 November 2017; Accepted 15 November 2017
Available online 17 November 2017
1433-8319/ © 2017 Elsevier GmbH. All rights reserved.

Dainese and Sitzia, 2013), because these better provisioned seeds have a higher chance of reaching maturity. On the other hand, productive sites also select for large-seeded species because non-limiting nutrients enhance light competition and stressful shading, which are necessary for seedlings to overcome (Bergholz et al., 2015; Hautier et al., 2009).

Arable weed communities are an excellent example of annual-dominated and frequently disturbed communities (Lososová et al., 2006) where seed mass is a vital trait (Manning et al., 2009). Weeds grow in the crop understorey and shade impacts them during their whole life cycle. Recently, trait-based studies that dealt with seed mass as an essential part of the LHS scheme (Westoby, 1998) questioned its relevance in herbs (Klimešová et al., 2016) or arable communities (Perronne et al., 2015). Currently, the relationship between community seed mass and light limitation gradients is under debate. While some studies found a positive linear relationship (Leishman et al., 2000; Manning et al., 2009; May et al., 2013), others reported non-linear trends (Bergholz et al., 2015; Guerrero et al., 2014) or found weak or no evidence (Lhotsky et al., 2016; Lönnberg and Eriksson, 2013; Santini et al., 2017). Moreover, it should be noted that the majority of studies reporting an increase in seed mass along plant density or productivity gradients were conducted on annuals (Harel et al., 2011; Manning et al., 2009; Santini et al., 2017) or annual-dominated communities (Guerrero et al., 2014; May et al., 2013), but see Bergholz et al. (2015). This finding suggests that life span might play an important role in the association between seed mass and environmental conditions.

In this study, we investigated whether inconsistent and contradicting evidence in the literature stems from the disparity of plant strategies: annual and perennial life span. As far as we know, no study has challenged the established theory of seed mass changes in response to environmental variation from the separated perspective of annuals and perennials. It is, therefore, essential to provide a study that could reveal clear patterns in seed mass changes in annuals and perennials due to variation in one of the most often studied ecological filters (light limitation). Hence, we tested the following hypothesis: average seed mass (community-weighted mean CWM) increases along a crop coverage gradient (proxy of light limitation or productivity), i.e., species with a higher seed mass are favoured under high crop coverage because these species exhibit better light limitation (stressful shading) tolerance. We specifically tested whether the changes in the seed mass CWM along a crop coverage gradient are consistent or differ according to life span strategy (Fig. 1).

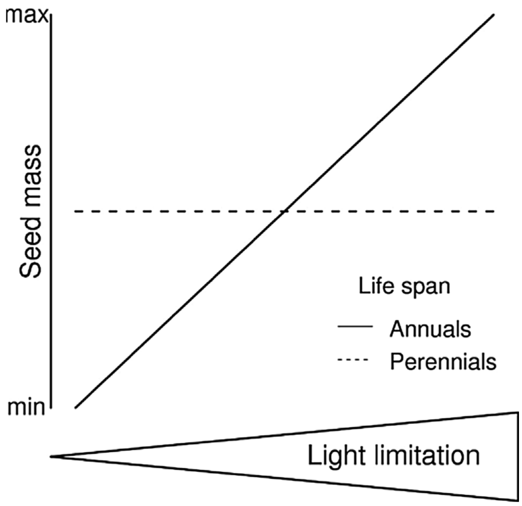


Fig. 1. The hypothesis tested in this study. The strength of the relationship between community seed mass and increasing light limitation is dependent on seed mass relevance. Here, we hypothesized that this relevance differs in annual and perennial weeds. When plant performance is largely dependent on seed mass (annuals), a response is detectable. On the other hand, one can observe no response in cases where recruitment is weakly dependent on seed mass (perennials with clonal reproduction). Note that the slope and intercept of the regression line change according to relevance.

2. Material and methods

2.1. Study sites and sampling

The study area (Fig. 2) covers 11,000 km^2 in northern and central Moravia, which is located in the northeastern part of the Czech Republic (49° 08′–50° 27′ N, 16° 43′–18° 52′ E). The mean annual temperature is 4.5–8.5 °C, and the annual precipitation ranges from 575 to 1 300 mm. The altitude of the study sites spans from 195 to 815 m a. s. l. We recorded vegetation data on arable land for the period 2001–2003. We used a stratified sampling scheme for obtaining a stratified dataset, which produced land categories (polygons) with unique combinations of soil, climate and potential natural vegetation. For more details about the stratifying procedure see Cimalová and Lososová (2009). We sampled in both cereal (autumn seeding: wheat and rye; spring seeding: oat and barley) and root crops (sugar beet, potatoes). Plots of a standard size of 20 m^2 were randomly situated along edges where the effects of herbicide use are relatively low in conventionally managed fields. The total number of sampled plots was 233.

2.2. Traits and variables

We retrieved average seed mass data from the databases BiolFlor (Klotz et al., 2002) and LEDA (Kleyer et al., 2008). A total of 208 species were recorded (103 annuals and 105 perennials; listed in Appendix A). For the species *Armoracia rusticana* we found no data; hence, its seed mass was estimated from the regression model of the inter-trait relationship between seed mass and plant height (data from LEDA database). To assess the impact on weed community seed mass, we used three variables: (1) crop coverage (estimated cover of the planted crop), (2) crop type (cereals: barley, oat, rye and wheat; root crops: beet and potato) and (3) sampling period (part of the season coded as a number of days since the start of the year). We used crop coverage as a suitable proxy for the light availability gradient. Weeds grow in the understorey; hence, the density of the crop canopy has to be strongly correlated with the light availability. We determined the crop coverage gradient by visual estimation of percentage cover on the van der Maarel scale (van der Maarel, 1979).

Weed communities are specifically shaped by cultivated plants and their associated management, and exhibit significant seasonal variation (Cimalová and Lososová, 2009; Gross et al., 2015; Nowak et al., 2015). Crop type is often used as an approximate measure of the differences in agricultural practices (like disturbance, sowing date, herbicide use or inherent traits of the crop; Lososová et al., 2004; Nowak et al., 2015; Šilc et al., 2009) that have been shown to be influential drivers of the seed mass response (Gaba et al., 2014; Gunton et al., 2011). We treated the crop type (an indicator of management practices) and season (an indicator of phenological changes in weed vegetation) as confounding variables. Crop coverage was the variable of direct interest.

2.3. Data processing and software

Prior to analysis we excluded crops, juvenile trees and non-seeded plants from the dataset. Prior to CWM calculation, we divided recorded plants into two groups based on the life span (annuals or perennials) and log-transformed their seed mass values. The seed mass CWM (weighted by percentage cover of each species) was calculated separately for each group in each plot. To explore the effect of light limitation on the seed mass CWM, we first tested the significance of the interaction term: life span x crop coverage. To do so, we combined annual and perennial datasets and fitted a linear model using the *lm* function in R (version 3.3.1; R Core Team, 2016) with all other predictors. Life span was treated as a binary variable. Given the non-constant proportions of annuals or perennials in sampled plots, we also

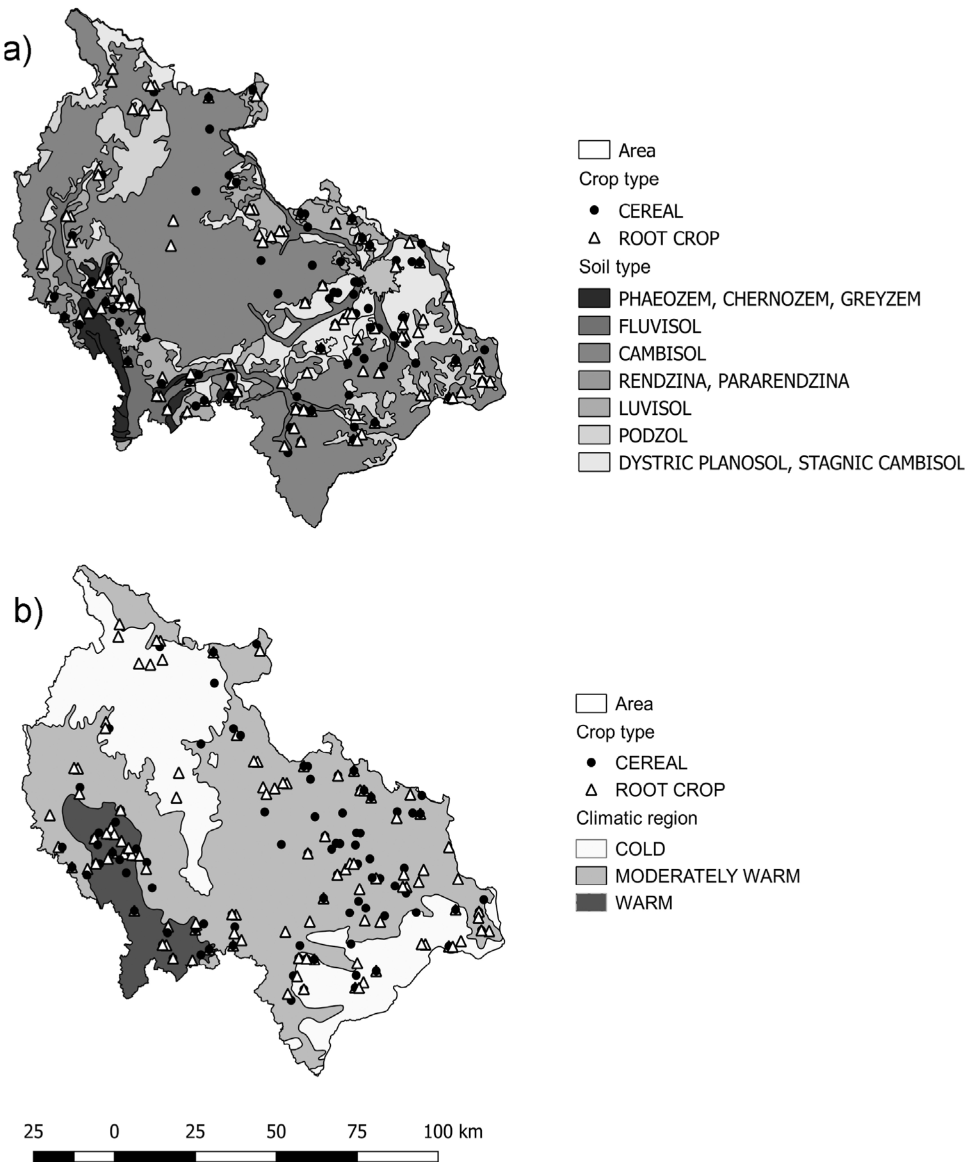


Fig. 2. Study area. We sampled in the northeastern part of the Czech Republic. Plots are distinguished according to crop type (circles = cereals, triangles = root crops). a) Soil map and b) Climate map.

tested the interaction term with weighted CWM values (based on the number of species in a plot used for calculation).

Then, we performed separate models for only annual or perennials plants. We used generalized least-squares models using the *nlme* library (Pinheiro et al., 2016). We employed this method in order to account for potential spatial autocorrelation (nearby plots could have similar species composition and therefore similar CWM) in the response variable. We evaluated all correlation structures available in *nlme* library and chose the most suitable according to Akaike's information criterion (AIC). We included three explanatory variables (see the Traits and variables section) in the model comparison process. The seed mass CWM variation in response to seasonality or crop type was not of direct interest hence we used them as confounding covariates. We performed forward selection as a model building approach. In particular, we compared three models related to our hypotheses (Table 1). Model 1 was a null model, model 2 had season and crop type variables and model 3 contained all variables (season, crop type and crop coverage). This approach allowed us to test the effect of light availability against a null model and a model that already accounted for seasonal and management bias. Model comparison was based on AIC and Akaike's weights (Burnham and Anderson, 2002) with the maximum likelihood (ML) estimation method. We then refitted the best model with the

Table 1
The comparison of several generalized least squares (GLS) models testing the effect of light limitation on the seed mass community-weighted mean (CWM) of a) annual and b) perennial weeds. Model comparison is based on AIC (lower value suggests better fit) and Akaike's weights (the probability that a particular model is the best from the set of given models). Model 1 was without predictors, model 2 contained only confounding variables (crop type and season), and model 3 included confounding variables and crop coverage (an indicator of light limitation). Spatial correlation was accounted for, and seed mass values were log-transformed. c) Coefficients (standard error and P-value) of the crop coverage for both life span groups.

Model	Predictors	AIC	Aw
a) Annuals			
1	none	35.34	< 0.01
2	season + crop type	26.92	0.18
3	season + crop type + crop coverage	23.93	0.82
b) Perennials			
1	none	132.98	< 0.01
2	season + crop type	119.42	0.72
3	season + crop type + crop coverage	121.27	0.18
c) Life span	Crop coverage coefficient	SE	P-value
Annuals	0.02	0.008	0.025
Perennials	− 0.004	0.01	n.s.

Table 2
Summary information about the relative proportions of annuals and perennials in sampled plots.

Life span	Annuals (%)	Perennials (%)
Average	65.2	34.8
sd	18.4	18.4
min	2.2	4.7
max	95.3	97.8
n	103 species	105 species

restricted maximum likelihood (REML) method to obtain more precise and unbiased parameters' estimates. We graphically evaluated the assumptions of residual normal distribution and uncorrelatedness. All data are accessible in Mendeley Data repository (Bitomský et al., 2017).

3. Results

The seed mass of the considered weed species ranged between 0.01 and 39.4 mg. Both life span groups had similar seed mass distribution (annuals average ± SD = 2.6 ± 5.48 mg; perennials average ± SD = 2.7 ± 5.33 mg). On average, annuals were more frequent in the sampled plots (Table 2).

A linear model using the combined data of annuals and perennials revealed a significant interaction between crop coverage and life span (F = 6.52, P = 0.01). A weighted linear model (weights represented by the number of species) provided the same result (F = 7.02, P = 0.008). Separate models supported the prediction that the seed mass CWM of annuals is more sensitive to the light limitation gradient than that of perennials. Based on AIC (the lowest value indicates a better fit to the data) and Akaike's weights (the relative likelihood of the model, i.e., the probability that a particular model has the best fit from a set of given models), adding crop coverage improved the model fit in only annuals (Table 1). The seed mass CWM of annuals was positively dependent on crop coverage (β = 0.02, SE = 0.008, P = 0.025), while perennials showed no relationship (β = −0.004, SE = 0.01, P = 0.7; Fig. 3). The quadratic term was not significant (F = 0.01, P = 0.91). In both models, incorporating exponential spatial correlation (corExp function in R, Pinheiro et al., 2016) slightly improved the model fit.

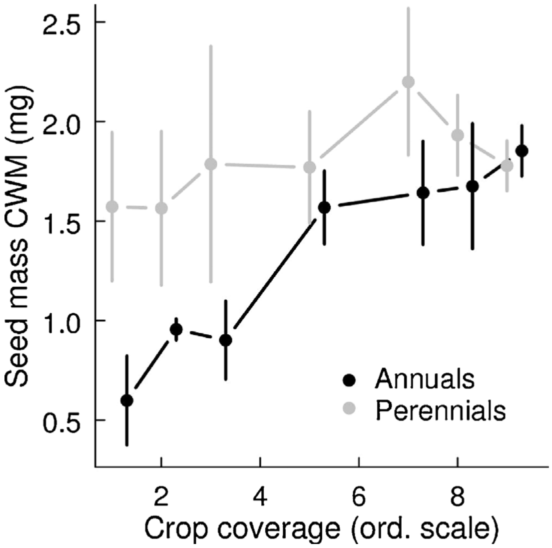


Fig. 3. Trends in seed mass community-weighted mean (CWM) along the light limitation gradient separated for annual (black) and perennial (grey) plants. Average seed mass CWM +/− SE of weeds in response to crop coverage. Estimated crop coverage is on the ordinal van der Maarel scale (van der Maarel, 1979); values from 1 to 9 account for 1, 2, 3, 5, 8, 13, 18, 38, 63 and 88% coverage, respectively. Raw untransformed data are shown.

4. Discussion

The community-weighted mean (CWM) of the seed mass of annuals tended to be higher under dense crop canopies (Fig. 3). This finding represents quite robust evidence since it notes a clear response of seed mass after accounting for the noise of season and crop type. Light availability, therefore, seems to be an important driver of the seed mass response of annuals in arable weed communities. This finding is in concordance with the idea that annual plants with large seeds are better able to cope with light shortage than annuals with small seeds. This result can be explained via two complementary hypotheses. Either large-seeded annual weeds have higher survival in shady crop understorey (Manning et al., 2009), or they can more effectively compete for light with crops (DeMalach et al., 2016; Tilman, 1988). More reserves in larger seeds could mitigate the impact of light limitation (stressful shading) during seedling development. Seedling survival is dependent on seed mass, and seedlings from larger seeds perform better under shaded conditions (Leishman et al., 2000), especially during early stages of seedling growth (Walters and Reich, 2000). Small-seeded species do not occur in shady conditions not only because of smaller reserves but also because of higher light requirements for germination (Jensen and Gutekunst, 2003). It should also be noted that the importance of seed mass in mediating the effects of shade is largely dependent on ontogeny (Moles and Westoby, 2004; Niinemets, 2006). Although weeds can grow to the same (or higher) height as crops, during germination and seedling establishment in the crop understorey (early ontogenetic phase), they face strong or weak light limitation depending on crop canopy density. During this phase, seed mass is relatively more vital in comparison with other traits related to shade tolerance.

Our findings provide new insight into the mixed evidence for seed mass response to light limitation or closely associated productivity gradients. Positive response, non-linear response or no association can be found in the ecological literature. As we found a positive, linear relationship between the seed mass CWM of annuals and crop cover, our results differ from the findings of Guerrero et al. (2014) and Santini et al. (2017). Guerrero et al. (2014) showed an increase in the seed mass of arable weeds along an intensification gradient (proxied by yield, sowing density and nitrogen fertilizer), but this relationship was non-linear. In contrast, Santini et al. (2017) did not find any significant relationship between the seed mass and shade (Ellenberg indicator values for light) using a large dataset of annuals from the United Kingdom. Some mixed evidence could stem from the selection of the variable that represents a productivity gradient (indicating light limitation and the intensity of competition for light). Soil resource availability (Adler et al., 2013), plant cover (this study, Lhotsky et al., 2016) or both (Bergholz et al., 2015; Guerrero et al., 2014; Manning et al., 2009) are often used. There are other possibilities, such as water availability (Harel et al., 2011; May et al., 2013) in arid systems or Ellenberg indicator values for nutrients or light (Santini et al., 2017). Another possible reason for the inconsistency may be due to combining data across habitats. The nature of the relationship between seed mass and light limitation is likely habitat-dependent. For example, a greater seed mass of herbs in a forest understorey provides a negligible advantage, as it is not possible for herbs to reach the canopy or overgrow trees. Mixing annuals occurring in herbaceous habitats with annuals occurring in forests in one analysis might remove the detectable association between seed mass and shade in the study by Santini et al. (2017). Conditions in a forest understorey, compared to herbaceous systems, might select for different strategies (not dependent on seed mass) for how to cope with low light during the whole life cycle.

Another question is how important is light as a driver of community seed mass (effect size). For instance, Lönnberg and Eriksson (2012,2013) argued that shade is a weak filter of seed size. To understand often contradicting results, it is necessary to take into account that shade under canopies is correlated with other environmental conditions (Valladares et al., 2016). Shade moderates temperature fluctuations in the understorey or retains humidity. Plants in unshaded plots may even experience harsher conditions (drought or extreme temperatures) than plants under dense canopies. This complexity of shade effects could create some noise in our data. Still, our model suggests that light limitation is a stronger seed mass filter than other counteracting environmental factors related to shade. Finally, noise in the data could be caused by weeds that avoided competition for light, such as weeds that emerged immediately at the start of the growing season and grew faster than the planted crop. These relatively small-seeded, early species (e.g., *Capsella bursa-pastoris*, *Stellaria media*, *Veronica persica* and *Viola arvensis*) were not affected by crop canopy in the fields of spring cereals and root crops.

In observational studies focusing on the CWM, it is crucial to discuss the limitations and other possible causal links. First, identified patterns may be the result of the correlation with another trait, such as plant height. Taller plants are dominant when competing for light and typically occur in productive habitats (Moles et al., 2009). However, this correlation is not necessarily a problem but rather a feature. Plant height is a function of seed mass, and both are part of the same strategy (Díaz et al., 2016). A seedling from a large seed performs better in the shady understorey, and a taller adult plant is a better competitor for light. Second, some issues with CWM calculation have to be considered (for the criticism of the CWM approach, see Peres-Neto et al., 2017). Given the non-constant proportions of annuals or perennials in sampled plots, there is a risk that CWM estimate will be unreliable when only a few species occur in a plot. To overcome this issue, we also performed the same analyses using weights (based on the number of species used for CWM calculation) with the same outcome, i.e., the life span x crop coverage interaction was significant (see results). Finally, the potential effect of phylogeny should be considered. We found 208 species belonging to 31 families (Appendix A). Our data seemed to show no systematic bias in phylogenetic information because annual and perennial species belonged to the same families and because most of the common families were equally distributed according to life span (e.g., Asteraceae, Fabaceae and Polygonaceae) except Brassicaceae (almost all species were annuals).

5. Conclusions

We showed that the community seed mass was driven by the light limitation gradient only in the case of annual plants. One, therefore, needs to consider if seed mass is a relevant trait in studied species, not only in a given system. We want to emphasize that studies with seed mass in perennial herbs should be conducted with caution, as perennials often prioritize many other reproduction strategies (e.g., clonal propagation) or survive under unsuitable conditions in structures other than seeds. This phenomenon could result in stochastic or even false relationships among seed mass and the studied ecological filters. The strength of the relationship between an ecological filter and seed mass is determined by the degree to which seed mass is a vital trait for the considered species. Given low importance, plants should exhibit stochastic average seed mass along the light limitation gradient. If seed mass is a crucial trait, its response to light should be strong, i.e., a steeper slope (as predicted in Fig. 1). Here, we showed that seed mass importance is determined by life span, i.e., deterministic patterns occur in assemblages of annuals (Bergholz et al., 2017; Leishman, 2001; Manning et al., 2009). The same applies for other non-clonal plants such as trees where the vital role of seed mass has been thoroughly documented (e.g., Walters and Reich, 2000). In contrast, in habitats where plants rely on clonal propagation (e.g., many perennial herbs), random patterns of community seed mass may appear. For example, when studying trait-based community assembly patterns, bud bank or clonal traits could be better predictors of distribution in perennials (Klimešová et al., 2016). Here, we suggest that considering life span regarding seed mass changes along an environmental gradient could effectively improve the interpretation of results, as whole community data can be quite noisy in some systems.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

MB was supported by grants from the Internal Grant Agency of Palacký University (IGA_PrF_2016_019 and IGA_PrF_2017_015). Data collection was supported by the project MSM 0021622416 of the Ministry of Education of the Czech Republic. We are grateful to Milan Chytrý for valuable suggestions regarding the sampling design. We thank Stanislav Hejduk and Jan Winkler for information about agricultural practices. The clarity of the manuscript was substantially improved thanks to the suggestions of editor Florian Jeltsch and two anonymous reviewers.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ppees.2017.11.005.

References

Adler, P.B., Fajardo, A., Kleinhesselink, A.R., Kraft, N.J.B., 2013. Trait-based tests of coexistence mechanisms. Ecol. Lett. 16, 1294–1306. http://dx.doi.org/10.1111/ele.12157.

Bergholz, K., Jeltsch, F., Weiss, L., Pottek, J., Geißler, K., Ristow, M., 2015. Fertilization affects the establishment ability of species differing in seed mass via direct nutrient addition and indirect competition effects. Oikos 124, 1547–1554. http://dx.doi.org/10.1111/oik.02193.

Bergholz, K., May, F., Giladi, I., Ristow, M., Ziv, Y., Jeltsch, F., 2017. Environmental heterogeneity drives fine-scale species assembly and functional diversity of annual plants in a semi-arid environment. Perspect. Plant Ecol. Evol. Syst. 24, 138–146. http://dx.doi.org/10.1016/j.ppees.2017.01.001.

Bitomský, M., Mládek, J., Cimalová, Š., 2017. Data From: Light Limitation Shapes the Community Seed Mass of Annual but Not of Perennial Weeds. Mendeley Data, v1. http://dx.doi.org/10.17632/fm38vrhjyp.1.

Borer, E.T., Seabloom, E.W., Gruner, D.S., Harpole, W.S., Hillebrand, H., Lind, E.M., Adler, P.B., Alberti, J., Anderson, T.M., Bakker, J.D., Biederman, L., Blumenthal, D., Brown, C.S., Brudvig, L.A., Buckley, Y.M., Cadotte, M., Chu, C., Cleland, E.E., Crawley, M.J., Daleo, P., Damschen, E.I., Davies, K.F., DeCrappeo, N.M., Du, G., Firn, J., Hautier, Y., Heckman, R.W., Hector, A., HilleRisLambers, J., Iribarne, O., Klein, J.A., Knops, J.M.H., La Pierre, K.J., Leakey, A.D.B., Li, W., MacDougall, A.S., McCulley, R.L., Melbourne, B.A., Mitchell, C.E., Moore, J.L., Mortensen, B., O'Halloran, L.R., Orrock, J.L., Pascual, J., Prober, S.M., Pyke, D.A., Risch, A.C., Schuetz, M., Smith, M.D., Stevens, C.J., Sullivan, L.L., Williams, R.J., Wragg, P.D., Wright, J.P., Yang, L.H., 2014. Herbivores and nutrients control grassland plant diversity via light limitation. Nature 508, 517–520. http://dx.doi.org/10.1038/nature13144.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference, first ed. Springer, New York.

Cimalová, Š., Lososová, Z., 2009. Arable weed vegetation of the northeastern part of the Czech Republic: effects of environmental factors on species composition. Plant Ecol. 203, 45–57. http://dx.doi.org/10.1007/s11258-008-9503-1.

Coomes, D.A., Grubb, P.J., 2003. Colonization, tolerance, competition and seed-size variation within functional groups. Trends Ecol. Evol. 18, 283–291. http://dx.doi.org/10.1016/S0169-5347(03)00072-7.

Díaz, S., Kattge, J., Cornelissen, J.H., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Prentice, I.C., Garnier, E., Bonisch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Wright, S.J., Sheremet'ev, S.N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S., Gunther, A., Falczuk, V., Ruger, N., Mahecha, M.D., Gorne, L.D., 2016. The global spectrum of plant form and function. Nature 529, 167–171. http://dx.doi.org/10.1038/nature16489.

Dainese, M., Sitzia, T., 2013. Assessing the influence of environmental gradients on seed mass variation in mountain grasslands using a spatial phylogenetic filtering approach. Perspect. Plant Ecol. Evol. Syst. 15, 12–19. http://dx.doi.org/10.1016/j.ppees.2012.10.005.

DeMalach, N., Zaady, E., Weiner, J., Kadmon, R., 2016. Size asymmetry of resource competition and the structure of plant communities. J. Ecol. 104, 899–910. http://dx.doi.org/10.1111/1365-2745.12557.

Fenner, M.W., 1985. Seed Ecology, first ed. Chapman & Hall, New York.

Gaba, S., Fried, G., Kazakou, E., Chauvel, B., Navas, M.L., 2014. Agroecological weed control using a functional approach: a review of cropping systems diversity. Agron. Sustain. Dev. 34, 103–119. http://dx.doi.org/10.1007/s13593-013-0166-5.

Gross, K.L., Emery, S., Davis, A.S., Smith, R.G., Robinson, T.M.P., 2015. Plant community dynamics in agricultural and successional fields. In: Hamilton, S.K., Doll, J.E., Robertson, G.P. (Eds.), The Ecology of Agricultural Landscapes. Oxford University Press, New York, pp. 158–187.

Guerrero, I., Carmona, C.P., Morales, M.B., Oñate, J.J., Peco, B., 2014. Non-linear responses of functional diversity and redundancy to agricultural intensification at the field scale in Mediterranean arable plant communities. Agric. Ecosyst. Environ. 195, 36–43. http://dx.doi.org/10.1016/j.agee.2014.05.021.

Gunton, R.M., Petit, S., Gaba, S., 2011. Functional traits relating arable weed communities to crop characteristics. J. Veg. Sci. 22, 541–550. http://dx.doi.org/10.1111/j.1654-1103.2011.01273.x.

Harel, D., Holzapfel, C., Sternberg, M., 2011. Seed mass and dormancy of annual plant populations and communities decreases with aridity and rainfall predictability. Basic Appl. Ecol. 12, 674–684. http://dx.doi.org/10.1016/j.baae.2011.09.003.

Hautier, Y., Niklaus, P.A., Hector, A., 2009. Competition for light causes plant biodiversity loss after eutrophication. Science 324, 636–638. http://dx.doi.org/10.1126/science.1169640.

Jensen, K., Gutekunst, K., 2003. Effects of litter on establishment of grassland plant species: the role of seed size and successional status. Basic Appl. Ecol. 4, 579–587. http://dx.doi.org/10.1078/1439-1791-00179.

Kleyer, M., Bekker, R.M., Knevel, I.C., Bakker, J.P., Thompson, K., Sonnenschein, M., Poschlod, P., van Groenendael, J.M., Klimeš, L., Klimešová, J., Klotz, S., Rusch, G.M., Hermy, M., Adriaens, D., Boedeltje, G., Bossuyt, B., Dannemann, A., Endels, P., Gotzenberger, L., Hodgson, J.G., Jackel, A.K., Kuhn, I., Kunzmann, D., Ozinga, W.A., Romermann, C., Stadler, M., Schlegelmilch, J., Steendam, H.J., Tackenberg, O., Wilmann, B., Cornelissen, J.H.C., Eriksson, O., Garnier, E., Peco, B., 2008. The LEDA Traitbase: a database of life-history traits of the Northwest European flora. J. Ecol. 96, 1266–1274. http://dx.doi.org/10.1111/j.1365-2745.2008.01430.x.

Klimešová, J., Tackenberg, O., Herben, T., 2016. Herbs are different: clonal and bud bank traits can matter more than leaf-height-seed traits. New Phytol. 210, 13–17. http://dx.doi.org/10.1111/nph.13788.

Klotz, S., Kühn, I., Durka, W., 2002. BIOLFLOR–Eine Datenbank mit biologisch-ökologischen Merkmalen zur Flora von Deutschland. Schriftenr für Veg. 38, 1–334.

Lönnberg, K., Eriksson, O., 2012. Seed size and recruitment patterns in a gradient from grassland to forest. Ecoscience 19, 140–147. http://dx.doi.org/10.2980/19-2-3521.

Lönnberg, K., Eriksson, O., 2013. Rules of the seed size game: contests between large-seeded and small-seeded species. Oikos 122, 1080–1084. http://dx.doi.org/10.1111/j.1600-0706.2012.00249.x.

Leishman, M.R., Wright, I.J., Moles, A.T., Westoby, M., 2000. The evolutionary ecology of seed size. In: Fenner, M. (Ed.), Seeds: The Ecology of Regeneration in Plant Communities. CAB International, Wallingford, pp. 31–57.

Leishman, M.R., 2001. Does the seed size/number trade-off model determine plant community structure? An assessment of the model mechanisms and their generality. Oikos 93, 294–302. http://dx.doi.org/10.1034/j.1600-0706.2001.930212.x.

Levine, J.M., Rees, M., 2002. Coexistence and relative abundance in annual plant assemblages: the roles of competition and colonization. Am. Nat. 160, 452–467.

Lhotsky, B., Kovács, B., Ónodi, G., Csecserits, A., Rédei, T., Lengyel, A., Kertész, M., Botta-Dukát, Z., 2016. Changes in assembly rules along a stress gradient from open dry grasslands to wetlands. J. Ecol. 104, 507–517. http://dx.doi.org/10.1111/1365-2745.12532.

Lososová, Z., Chytrý, M., Cimalová, S., Kropáč, Z., Otýpková, Z., Pyšek, P., Tichý, L., 2004. Weed vegetation of arable land in Central Europe: gradients of diversity and species composition. J. Veg. Sci. 15, 415–422. http://dx.doi.org/10.1111/j.1654-1103.2004. tb02279.x.

Lososová, Z., Chytrý, M., Kühn, I., Hájek, O., Horáková, V., Pyšek, P., Tichý, L., 2006. Patterns of plant traits in annual vegetation of man-made habitats in central Europe. Perspect. Plant Ecol. Evol. Syst. 8, 69–81.

Manning, P., Houston, K., Evans, T., 2009. Shifts in seed size across experimental nitrogen enrichment and plant density gradients. Basic Appl. Ecol. 10, 300–308. http://dx.doi.org/10.1016/j.baae.2008.08.004.

May, F., Giladi, I., Ristow, M., Ziv, Y., Jeltsch, F., 2013. Plant functional traits and community assembly along interacting gradients of productivity and fragmentation. Perspect. Plant Ecol. Evol. Syst. 15, 304–318. http://dx.doi.org/10.1016/j.ppees.2013.08.002.

Metz, J., Liancourt, P., Kigel, J., Harel, D., Sternberg, M., Tielbörger, K., 2010. Plant survival in relation to seed size along environmental gradients: a long-term study from semi-arid and Mediterranean annual plant communities. J. Ecol. 98, 697–704. http://dx.doi.org/10.1111/j.1365-2745.2010.01652.x.

Moles, A.T., Westoby, M., 2004. Seedling survival and seed size: a synthesis of the literature. J. Ecol. 92, 372–383. http://dx.doi.org/10.1111/j. 0022-0477.2004.00884.x.

Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W., Zanne, A.E., Pitman, A., Hemmings, F.A., Leishman, M.R., 2009. Global patterns in plant height. J. Ecol. 97, 923–932. http://dx.doi.org/10.1111/j. 1365-2745.2009.01526.x.

Muller-Landau, H.C., 2010. The tolerance-fecundity trade-off and the maintenance of diversity in seed size. Proc. Natl. Acad. Sci. U. S. A. 107, 4242–4247. http://dx.doi.org/10.1073/pnas.0911637107.

Niinemets, Ü., 2006. The controversy over traits conferring shade-tolerance in trees: ontogenetic changes revisited. J. Ecol. 94, 464–470. http://dx.doi.org/10.1111/j.1365-2745.2006.01093.x.

Nowak, A., Nowak, S., Nobis, M., Nobis, A., 2015. Crop type and altitude are the main drivers of species composition of arable weed vegetation in Tajikistan. Weed Res. 55, 525–536. http://dx.doi.org/10.1111/wre.12165.

Peres-Neto, P.R., Dray, S., ter Braak, C.J.F., 2017. Linking trait variation to the environment: critical issues with community-weighted mean correlation resolved by the fourth-corner approach. Ecography 40, 806–816. http://dx.doi.org/10.1111/ecog.02302.

Perronne, R., Le Corre, V., Bretagnolle, V., Gaba, S., 2015. Stochastic processes and crop types shape weed community assembly in arable fields. J. Veg. Sci. 26, 348–359. http://dx.doi.org/10.1111/jvs.12238.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., 2016. Nlme: Linear and Nonlinear Mixed Effects Models. https://cran.r-project.org/web/packages/nlme/index.html.

R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Santini, B.A., Hodgson, J.G., Thompson, K., Wilson, P.J., Band, S.R., Jones, G., Charles, M., Bogaard, A., Palmer, C., Rees, M., 2017. The triangular seed mass–leaf area relationship holds for annual plants and is determined by habitat productivity. Funct. Ecol. http://dx.doi.org/10.1111/1365-2435.12870.

Šilc, U., Vrbničanin, S., Božić, D., Čarni, A., Stevanović, Z.D., 2009. Weed vegetation in the North-Western Balkans: diversity and species composition. Weed Res. 49, 602–612. http://dx.doi.org/10.1111/j.1365-3180.2009.00726.x.

Thompson, K., 1987. Seeds and seed banks. New Phytol. 106, 23–34. http://dx.doi.org/10.1111/j. 1469-8137.1987.tb04680.x.

Tilman, D., 1988. Plant Strategies and the Dynamics and Structure of Plant Communities. Princeton University Press, Princeton, New Jersey.

Valladares, F., Laanisto, L., Niinemets, Ü., Zavala, M.A., 2016. Shedding light on shade: ecological perspectives of understorey plant life. Plant Ecol. Divers. 9, 237–251. http://dx.doi.org/10.1080/17550874.2016.1210262.

van der Maarel, E., 1979. Transformation of cover-abundance values in phytosociology and its effects on community similarity. Vegetatio 39, 97–114. http://dx.doi.org/10.1007/BF00052021.

Walters, M.B., Reich, P.B., 2000. Seed size, nitrogen supply, and growth rate affect tree seedling survival in deep shade. Ecology 81, 1887–1901. http://dx.doi.org/10.1890/0012-9658(2000)081[1887:SSNSAG]2.0.CO;2.

Westoby, M., Falster, D.S., Moles, A.T., Vesk, P.A., Wright, I.J., 2002. Plant ecological strategies: some leading dimensions of variation between species. Annu. Rev. Ecol. Syst. 33, 125–159. http://dx.doi.org/10.1146/annurev.ecolsys.33.010802.150452.

Westoby, M., 1998. A leaf-height-seed (LHS) plant ecology strategy scheme. Plant Soil 199, 213–227. http://dx.doi.org/10.1023/A:1004327224729.

Received: 1 August 2018 | Revised: 22 March 2019 | Accepted: 16 April 2019

DOI: 10.1111/jvs.12768

RESEARCH ARTICLE

Journal of Vegetation Science IAVS

Herbaceous climbers in herbaceous systems are shade-tolerant and magnesium-demanding

Martin Bitomský[1] | Pavla Mládková[1,2] | Šárka Cimalová[3] | Jan Mládek[1]

[1]Department of Ecology and Environmental Sciences, Palacký University, Olomouc, Czech Republic

[2]Department of Ecology, Czech University of Life Sciences, Prague, Czech Republic

[3]Department of Biology and Ecology, University of Ostrava, Ostrava, Czech Republic

Correspondence
Martin Bitomský, Department of Ecology and Environmental Sciences, Palacký University, Olomouc, Czech Republic.
Email: bitomsky.martin@gmail.com

Funding information
Data collection in species-rich grasslands was supported by the Ministry of the Environment of the Czech Republic (RPV SP/2D3/179/07), the Grant Agency of the Czech University of Life Sciences (CIGA 42110/1313/3114) and the Ministry of Agriculture of the Czech Republic (5400/2013-MZE-13212). Data collection in arable fields was supported by the Ministry of Education of the Czech Republic (MSM 0021622416). MB was supported by a grant from the Internal Grant Agency of Palacký University (IGA_PrF_2018_020) and by a Scholarship of the city of Ostrava.

Co-ordinating Editor: Jodi Price

Abstract

Questions: Climbers do not invest in self-supporting architecture, allowing them to grow quickly and search for light and nutrients in canopy gaps. The climbing strategy requires both disturbance and external support; however, in herbaceous systems, disturbances often homogeneously remove external support. As a result, there should be a negative relationship between light availability and the presence of external support. Here, we asked if the distribution of herbaceous climbers is positively or negatively associated with disturbance, light availability and plant-available nutrients. Further, we tested if climbers differ in traits compared to co-occurring herbs.

Location: Czech Republic.

Methods: We used observations from a phytosociological database, species-rich grasslands and arable fields. First, we examined the presence of climbers in response to disturbance variables (frequency, severity and regime) and canopy cover (external support and light availability). Second, we examined the distribution of climbers along gradients of plant-available nutrients in species-rich grasslands. Finally, we compared seven functional traits between 31 climbers and 1,138 co-occurring herbs growing in 18 Central European herbaceous habitats.

Results: We found no relationship between any disturbance variables and the occurrence of climbers; indeed, climbers tended to occur more frequently in sites with denser canopies. In species-rich grasslands, the presence of climbers was positively related to plant-available magnesium. Climbers tended to be taller, with greater seed mass and specific leaf area (SLA) than co-occurring species.

Conclusions: Herbaceous climbers grow in plots with low light availability but plenty of external support and require magnesium for the formation of flexible stems. Compared to co-occurring herbs, climbers allocate resources saved from their lack of self-support to both shade-avoidance (canopy height) and shade-tolerance traits (seed mass, SLA). Compared to lianas, the distribution of herbaceous climbers is not affected by disturbance but is dependent on the availability of external support and magnesium.

KEYWORDS
canopy height, disturbance, light limitation, magnesium, seed mass, soil nutrients, specific leaf area, vines

© 2019 International Association for Vegetation Science

1 | INTRODUCTION

Climbers use the architecture of other plants to ascend to the canopy, and are typically classified according to growth form, either woody (often called lianas) or herbaceous (often called herbaceous climbers or vines). Most studies have focused on woody climbers, while studies on vines are limited. However, the available evidence suggests that the ecological strategy of vines is based on different mechanisms than those of woody climbers. Pierce et al. (2017) examined the ecological strategies of all climber growth habits applying the competitor, stress-tolerator, ruderal theory (Grime, 1977), and found that woody climbers tend to follow a CS/CSR strategy, while vines a CR/CSR strategy. This suggests that the strategy of vines is slightly more ruderal (a tendency to invest a large proportion of resources in propagules) in comparison with woody climbers. Furthermore, Hu, Li, and Li (2010) found different responses of vines and lianas to geographical and environmental gradients (rainfall and temperature), suggesting different strategies for coping with various environmental stresses. To understand the general principles of structural parasitism in all plant communities, it is therefore essential to examine the ecological strategy of vines in herbaceous systems.

Lianas are more frequent in canopy gaps (Dalling et al., 2012; Ledo & Schnitzer, 2014; Schnitzer, 2015), as they are fast-growing plants that exploit the high resource availability (light and nutrients) in disturbed sites (Kazda, 2015; Schnitzer & Bongers, 2002; Wyka, Oleksyn, Karolewski, & Schnitzer, 2013). However, typical disturbances in herbaceous systems are quite different from spatially heterogeneous gap dynamics in forests. For instance, mowing and tilling remove the above-ground biomass uniformly. Grazing may be spatially heterogeneous but still removes a substantial proportion of the biomass. In addition to changing light conditions after a disturbance, vines also have to cope with the removal of external support. Given the absence of gap disturbance dynamics in herbaceous systems, vines may either be more frequent in disturbed sites (without external support) with high light conditions or may be more frequent in undisturbed (but shady) sites with external support present. Assessments across herbaceous habitats testing these two competing hypotheses could provide new insights into the ongoing discussion of whether climbers are light-demanding plants or not and how climbers respond to disturbances in general (e.g., Carrasco-Urra & Gianoli, 2009; Schnitzer & Bongers, 2002; Valladares, Gianoli, & Saldaña, 2011).

In tropical ecosystems, climbers are fast-growing and nutrition-demanding, and their abundance is positively associated with soil N and P (Addo-Fordjour, Rahmad, & Shahrul, 2014; Homeier, Englert, Leuschner, Weigelt, & Unger, 2010; Kazda, 2015). In addition, climbers have higher mass-based concentrations of nutrients, such as N, P, K, Ca or Mg, in their tissues (Asner & Martin, 2012; Kazda, 2015; Wyka et al., 2013). These macronutrients are vital for the physiological processes of most plant species, but some are also directly involved in plant stem elongation and extensibility, important features for vines. For instance, Mg is responsible for cell wall extensibility (Nakajima, Morikawa, Igarashi, & Senda, 1981), and vines might require higher Mg concentrations in order to ascend to the canopy. However, there is a lack of general consensus on the distribution of climbers along gradients of soil nutrients because the availability of host plants and their characteristics may be more important than the direct effects of soil variables (Macía, Ruokolainen, Tuomisto, Quisbert, & Cala, 2007; van der Heijden & Phillips, 2008).

Disturbance, light and nutrient availability all affect species composition through their vegetative propagation, dispersal capacity, competitive ability, growth rate, and life strategy (Grime, 1977; Herben, Klimešová, & Chytrý, 2018). Several traits are critical to these plant functions, such as bud bank size and depth (the number of dormant meristems and their vertical distribution), specific leaf area (SLA), canopy height, seed mass (leaf-height-seed; LHS traits sensu Westoby, 1998), leaf dry matter content (LDMC) and the beginning of flowering. In the face of repeated disturbances, herbs typically invest in more buds stored at greater depths, and are shorter, with smaller seeds and a higher SLA (Grime, 1977; Klimešová & Klimeš, 2007; Pierce et al., 2017; Westoby, 1998). In nutrient-rich, less disturbed and shady environments, herbs produce fewer buds, and have higher LHS traits (Herben et al., 2018; Westoby, 1998). However, climbers can be environmentally filtered into a particular habitat based on different combinations of traits than those of other growth forms. For example, lianas possess smaller seeds, a higher SLA and higher leaf N than co-occurring trees as they tend to have wind-dispersed seeds, and higher leaf turnover and photosynthetic rates (Gallagher & Leishman, 2012). Moreover, it is becoming increasingly apparent that trait variability is relatively most important within sites as demonstrated in lianas, trees and herbs (Gallagher & Leishman, 2012; Pescador, de Bello, Valladares, & Escudero, 2015). Therefore, to properly compare trait ranges of vines and herbs, the focus should be on the within-site level (alpha trait component, sensu Ackerly & Cornwell, 2007).

Here, using three comprehensive datasets we examined the response of vines to important ecological gradients and tested potential differences in trait-based environmental filtering to the same habitat between vines and co-occurring herbs. We addressed three questions: (a) Is the distribution of vines positively or negatively associated with disturbance frequency, severity and light availability (which is altered by disturbance)? We examined whether vines tend to occur in undisturbed sites with developed vegetation and dense canopies or in disturbed sites with higher light levels. (b) Is the distribution of vines associated with plant-available nutrients? We expected the distribution to be positively associated with soil N, P and Mg. (c) Do vines differ from co-occurring herbs in any of seven functional traits? We compared trait values of vines with co-occurring herbs that are subject to the same environmental filters. Vines were expected to produce more buds, and have smaller seeds and higher SLA than co-occurring herbs (Table 1), assuming they follow a ruderal strategy (Pierce et al., 2017). Vines may also be taller than neighbouring plants, and have lower LDMC because climbers are characterised by low leaf construction costs and rapid leaf turnover

TABLE 1 Summary of hypothesized trait differences between vines and co-occurring herbs. Predicted trait values are based on an assumption that vines are ruderal and fast-growing plants filtered to the environment on the basis of different combinations of traits than other growth forms

Trait	Expected value	Possible explanation
Bud bank depth	Higher	Higher number of buds stored in a high depth enhance vegetative propagation
Bud bank size	Higher	
Beginning of flowering	Later	Vines have to wait for external support to grow. Vines may delay flowering due to this waiting strategy
Canopy height	Higher	To grow taller than neighbouring plants, vines should be taller
Leaf dry matter content	Lower	Fast-growing plants in resource-rich sites tend to have lower LDMC
Seed mass	Smaller	Small-seeded plants are often more frequent in disturbed and resource-rich sites
Specific leaf area	Higher	Fast-growing plants in resource-rich sites tend to have higher SLA

(Gallagher & Leishman, 2012; Wyka et al., 2013). Finally, climbers display different flowering phenology compared to trees, shrubs or herbs (Cortés-Flores, Hernández-Esquivel, González-Rodríguez, & Ibarra-Manríquez, 2017; Putz & Windsor, 1987), and we hypothesized that vines would exhibit delayed flowering because they have to wait for external support plants to grow.

2 | METHODS

2.1 | Data collection and preparation

For the purposes of this study, we utilised three data sets (Table 2): vegetation plots from the Czech National Phytosociological Database (Dataset 1; Chytrý & Rafajová, 2003; Chytrý, Pyšek, Tichý, Knollová, & Danihelka, 2005); experimental plots in two landscape-protected areas situated in the borderland between the Czech Republic and Slovakia (Beskydy and White Carpathian Mountains, Dataset 2); and vegetation plots sampled in arable fields in the northeastern part of the Czech Republic (Dataset 3). From Dataset 1, we only used plots where at least one vine occurred (number of plots = 15,224; Appendix S1). Vines (31 in total; Appendix S2) belonged to the following families: Cucurbitaceae (1), Convolvulaceae (2), Caryophyllaceae (1), Cuscutaceae (2), Polygonaceae (2), Rubiaceae (3) and Fabaceae (20). Dataset 2 included 240 plots from 12 long-term experimental sites in species-rich grasslands. Experimental sites were monitored after nine years (or seven in Beskydy) of management manipulation including

TABLE 2 Summary of all datasets used in this study. The species pool in Dataset 1 refers to the number of all herbs occurring in plots occupied by at least one vine

	Dataset 1	Dataset 2	Dataset 3
Description	18 herbaceous habitats	Species-rich grasslands	Arable fields
Location	Czech Republic	Borderland between the Czech Republic and Slovakia: Beskydy and White Carpathian Mountains	Northeastern part of the Czech Republic
Species pool	1,169	179	208
Number of climbers	31	9	12
Number of plots	15,224	240	231
Number of plots occupied at least by one climber	6,254	134	197
Variables	(1) herbaceous climbers (presence/absence)		
	(2) three disturbance parameters	(2) four disturbance regimes	(2) crop canopy cover
	(3) herb canopy cover	(3) herb canopy cover	
	(4) seven functional traits	(4) soil nutrients (C, Ca, C:N, K, Mg, N, P)	
Study questions	Climbers ~ disturbance + canopy cover (Question 1)	Climbers ~ disturbance regimes + canopy cover (Question 1)	Climbers ~ crop canopy cover (Question 1)
	Trait differences climbers versus co-occurring herbs (Question 3)	Climbers ~ soil nutrients (Question 2)	

four different disturbance regimes (summer mowing, grazing, grazing + spring burning of litter once every three years, left undisturbed). A detailed description of the sites and experimental sampling design is given in Appendix S3 and in Mládková et al. (2015), respectively. Dataset 3 contained 231 vegetation plots sampled in arable fields. The study area covers 11,000 km^2 in northern and central Moravia, which is located in the northeastern part of the Czech Republic (49°08′–50°27′ N, 16°43′–18°52′ E). Detailed information about data sampling is provided in Cimalová and Lososová (2009). Briefly, 10 m × 2 m belts were randomly situated along edges where the effects of herbicide use are relatively low in conventionally managed fields.

2.1.1 | The effect of disturbances and light availability on vine occurrence

To address this question, we focused on variables indicating disturbance frequency, severity and regime (Datasets 1 & 2), and used canopy cover as a surrogate of light availability. For Dataset 1, we assigned three disturbance parameters (Appendix S1) based on expert judgement (Herben, Chytrý, & Klimešová, 2016): (a) herb layer disturbance frequency (an estimate of the mean number of years between two consecutive disturbances common in that vegetation type); (b) disturbance severity (the proportion of aboveground herb biomass removed in a single disturbance event); and (c) degree of soil disturbance (the proportional change in bare ground cover after a single disturbance event). We used the proportion of plots with vine occurrence (plots with vine/all plots) in every habitat as a dependent variable for the three disturbance measures and performed generalized linear models (GLM) assuming a binomial error structure. Before fitting the models, we accounted for sampling effort and rarefied observed percentages of plots with a vine in order to ensure comparability among habitats with different numbers of plots sampled, using the *rarefy* function from the *vegan* package (R Core Team, R Foundation for Statistical Computing, Vienna, Austria). For Dataset 2, we used vegetation records in permanent plots, each 1 m × 1 m in size, in June 2013. Plots were subject to different disturbance regimes, including summer mowing, grazing, grazing + spring burning of litter once every three years and no disturbance. For detailed information about the statistical analysis, see the next section. All analyses were conducted in R software (R Core Team, R Foundation for Statistical Computing, Vienna, Austria).

We tested the effect of canopy cover on vine presence/absence in all datasets. Percent canopy cover was estimated visually. For Dataset 3 (arable fields), we used crop cover measured on the van der Maarel scale (van der Maarel, 1979). We treated vine occurrence as a binary variable (presence or absence), assumed a binomial error structure and used a logit link function in all analyses. For Datasets 1 and 3, we used all plots with known herb canopy cover and geographical coordinates ($n = 12,956$ and $n = 231$, respectively) and performed a GLM. We also checked for spatial autocorrelation by estimating Moran's *I* using the *correlog* function (*pgirmess* package; R Core Team, R Foundation for Statistical Computing, Vienna, Austria) on model residuals. In both datasets, residuals showed no spatial correlation. Moran's *I* was very close to zero (max = 0.084, min = −0.080) in all distance classes (data not shown); hence, we did not account for spatial autocorrelation in these models. For Dataset 2, we performed a generalized linear mixed-effect model (GLMM) using the restricted maximum-likelihood method with Laplace approximation (*glmer* function) in the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) considering experimental site as a random term (random intercept).

2.1.2 | The effect of soil nutrients on vine occurrence

We tested for the effects of soil nutrients and disturbance regimes on the distribution of vines in long-term experiments in species-rich grasslands (Dataset 2). Three soil samples collected by a 5-cm-diameter probe from the 0–20 cm layer in each plot were pooled for laboratory analysis. After removing plant residues, the soil samples were air-dried, ground in a mortar and sieved to 2 mm. We determined pH (in 0.01 M $CaCl_2$), plant-available P, Ca, K and Mg with the Mehlich 3 reagent (Mehlich, 1984), the percent content of organic carbon (C) with oxidation with potassium dichromate (Walkley–Black method; Black, 1965), and the percent content of total nitrogen (N) with the Kjeldahl method (AOAC, 1984).

We computed the relative importance of soil nutrients and disturbance regimes in explaining vine occurrence across experimental sites using hierarchical partitioning (Chevan & Sutherland, 1991). This method expresses the relative importance of each variable as an independent effect indicated by the explanatory power not influenced by joint actions with other considered variables. We computed independent effects using the *hier.part* function in the *hier.part* package (R Core Team, R Foundation for Statistical Computing, Vienna, Austria) with the root-mean-square prediction error as a goodness of fit measure. Further, we fitted a GLMM and selected the model with the best fit on the basis of AIC with a forward selection approach. Due to correlations among soil nutrients, we also checked for multicollinearity using variance inflation factors when adding new terms to a model. We treated experimental site as a variable with random effect (random intercept model).

2.1.3 | Trait differences between climbers and non-climbers

To create a trait database of vines and co-occurring herbs, we extracted plots occupied by vines from Dataset 1. All woody plants (several trees, shrubs and lianas) and non-native vines were removed. We excluded woody plants due to the focus on herbaceous communities and due to the sparse occurrence of seedlings of trees and shrubs in the studied systems. In addition, woody plants would have biased the trait analysis (e.g., due to their higher canopy height or seed mass). Non-native vines (*Echinocystis lobata* and *Tropaeolum majus*) were removed because of their different biogeographic origins that may reflect different processes. *Echinocystis lobata* is originally a forest species and *Tropaeolum majus* is usually grown in gardens as a decorative plant. In addition, both plants tend to be invasive in some regions and their invasive potential might be based on traits that are not directly linked to structural parasitism. For analyses, we selected seven traits: SLA, canopy height, seed mass, LDMC, beginning of flowering, bud bank size (BBsum) and bud bank depth (BBdepth). Canopy height, LDMC, SLA and seed mass were retrieved from the LEDA database (Kleyer et al., 2008) using the TR8 package (Bocci, 2015), and the beginning of flowering was obtained from BiolFlor (Klotz, Kühn, Durka, & Briemle, 2002). Finally, the bud bank size and bud bank depth were obtained from the CLO-PLA database (Klimešová, Danihelka, Chrtek, de Bello, & Herben, 2017).

After retrieving trait values from all sources, the trait dataset still contained some missing values. If a species had no trait information available we deleted it (61 species out of 1231). For species with incomplete trait data, we input the missing values using a regularized iterative principal components analysis (Josse & Husson, 2012) implemented in the *missMDA* package (Josse & Husson, 2016). This is a useful method for handling missing values because it takes into account overall trait similarity between species and links between traits. We used two dimensions to predict missing values as suggested by the *estim_ncpPCA* function (Josse & Husson, 2016). Using this approach, we inputted 169 missing values in bud bank depth, 131 in bud bank size, 75 in flowering, 183 in canopy height, 358 in LDMC, 260 in seed mass and 307 in SLA.

To examine trait differences between vines and co-occurring herbs, we used the approach of Ackerly and Cornwell (2007) for decomposing trait variation into alpha (within-site) and beta (among-site) components. An alpha trait component describes how the traits of species vary in relation to the species with which they co-occur (a positive number indicates that a species has a higher trait value than co-occurring species, while a negative value means a lower trait value compared to co-occurring species). We calculated alpha trait values for each vine and tested whether vines differ from co-occurring plants for each trait separately. We conducted a series of univariate tests on alpha trait values of vines if they significantly differed from zero (which indicates a scaled average trait value of co-occurring herbs) and either used a two-tailed *t* test or a non-parametric test (when the trait distribution visually deviated from the normal distribution). Canopy height, seed mass and SLA were log-transformed prior to the computation of alpha values.

3 | RESULTS

3.1 | The effect of disturbances and light availability on vine occurrence

Disturbance parameters (frequency, severity and degree of soil disturbance) in herbaceous habitats had no effect on vine presence/absence (GLM; frequency: deviance = 0.038, $p = 0.975$; severity: deviance = 0.003, $p = 0.953$; degree of soil disturbance: deviance = 0.002, $p = 0.968$; $n = 18$ in all cases). This was also true in species-rich grasslands, where we analysed four different disturbance regimes. The model (GLMM) including disturbance regimes had a higher AIC (three units) than the null model (no predictors with experimental site as a random term) and both models explained vine occurrence equally ($\chi^2 = 3.15$, $df = 3$, $p = 0.369$). Therefore, the disturbance regimes (summer mowing, grazing, grazing + spring burning of litter and non-disturbed plots) did not affect vine presence. However, vines showed a positive association with canopy cover in all three datasets (Figure 1), suggesting that the probability of vine occurrence increased with a denser canopy. Canopy cover positively affected vine occurrence but explained only a small proportion of the variance in all datasets (Dataset 1: $b = 0.02$, SE = 0.001, $p < 0.001$, $R^2 = 3\%$; Dataset 2: $b = 0.03$, SE = 0.012, $p = 0.010$, $R^2 = 4\%$; Dataset 3: $b = 0.29$, SE = 0.082, $p < 0.001$, $R^2 = 6\%$; all coefficients on the logit scale, i.e., the logarithm of the odds).

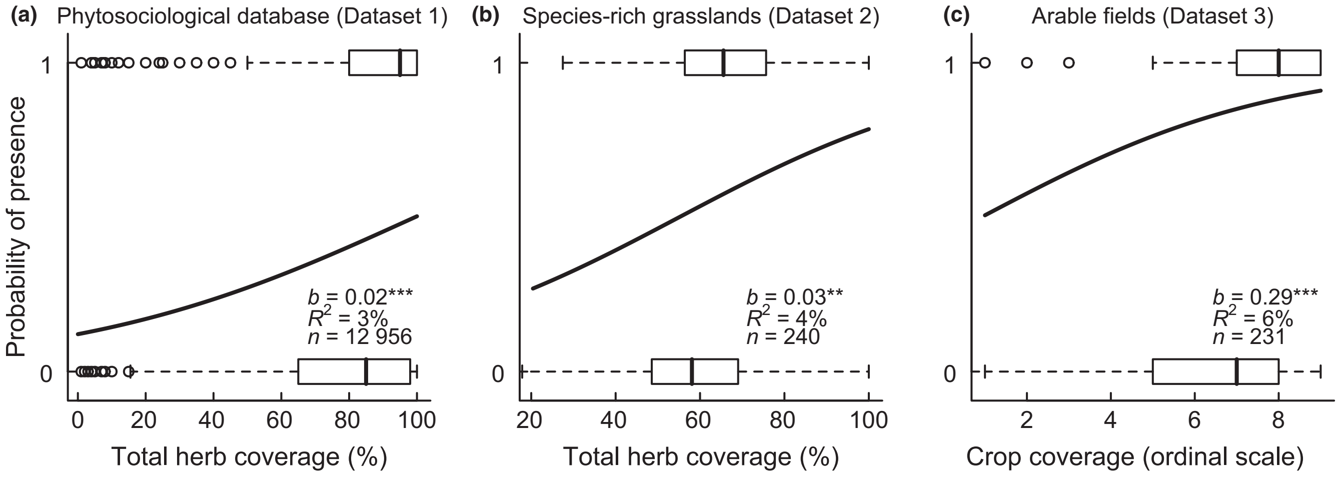


FIGURE 1 Relationships between the probability of the presence of a vine and the canopy cover (indicator of light availability) in all three datasets: (a) phytosociological database, (b) species-rich grasslands and (c) arable fields (see Table 2 and Methods for description). Estimated crop cover in (c) is on the ordinal van der Maarel scale (van der Maarel, 1979); values from 1 to 9 account for 1, 2, 3, 5, 8, 13, 18, 38, 63 and 88% cover, respectively. The coefficients are on a logit scale. The distribution of data points is presented using boxplots

3.2 | The effect of soil nutrients on vine occurrence

In species-rich grasslands, hierarchical partitioning revealed that Mg had the highest explanatory power (independent effect = 40.7%) and outperformed the other variables (Figure 2) in explaining vine occurrence. The model (GLMM) including Mg, K and their interaction best fit vine presence/absence (Table 3), thus demonstrating the importance of these two macronutrients at the experimental site level. The best model also suggested a positive effect of K; however, this nutrient showed low explanatory power (Figure 2) and was not significant when testing each nutrient separately. The probability of vine occurrence was highest in plots with high Mg but with relatively low K concentrations (Figure 3). However, modelling vine presence/absence did not account for species information, which might have led to biased results. To acknowledge this issue, we also performed multivariate analyses at the species level, which generally revealed similar patterns (Appendix S4). Except for *Galium aparine* and *Lathyrus pratensis*, all vines had a positive response to plant-available Mg (Appendix S5). Since the majority of vines belonged to the family Fabaceae, we also tested whether the response to Mg and K was a phylogenetic confounding effect. Analysing these relationships within Fabaceae showed that vines had a significantly higher probability of occurrence in plots with higher plant-available Mg compared to other Fabaceae plants (Appendix S6). In other words, the probability that a Fabaceae plant in a particular plot was a vine positively scaled with plant-available Mg. Finally, the vine–Mg relationship was also apparent within experimental sites (Appendix S7). In nine out of 12 sites, mean Mg concentration was higher in plots occupied at least by one vine compared to plots without vines.

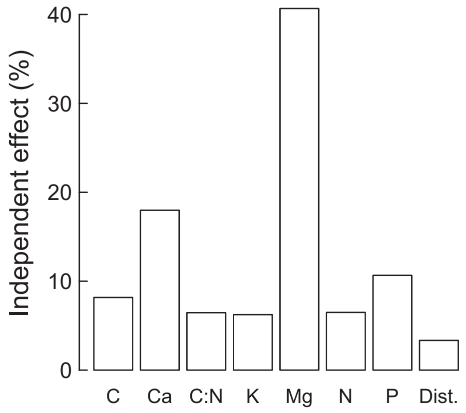


FIGURE 2 Independent effects (relative explanatory power of each variable without the cross-correlation with other predictors) of measured plant-available nutrients and management treatments indicating disturbance (Dist.). Independent effects indicate which variables are likely to be most influential in controlling variation in the presence/absence of vines in species-rich grasslands (Dataset 2, see Table 2 and Methods for description)

3.3 | Trait differences between climbers and co-occurring non-climbers

Canopy height, seed mass and SLA alpha values were significantly greater than zero (Table 4), indicating that vines had higher trait values than co-occurring herbs. On average, vines were 0.26 m taller, possessed a 2.2 mg greater seed mass and had a 1 cm^2/g higher SLA compared to co-occurring herbs. After standardization, the largest effect size was for seed mass (Figure 4). We did not observe any detectable differences in the LDMC, beginning of flowering, bud bank size or depth, suggesting that vines do not display unique strategies in these four traits compared to co-occurring non-climbers in the same habitat. We also checked for the necessity of incorporating phylogeny, but phylogenetic linear models did not improve the model fit (based on AIC; Appendix S8).

4 | DISCUSSION

Vines were not associated with any disturbance variable (frequency, severity, or regime), but their probability of occurrence increased with increasing canopy cover (Figure 1). This is in contrast with studies from tropical forests highlighting a higher abundance of woody climbers in disturbed canopy gaps (e.g., Dalling et al., 2012; Kazda, 2015; Ledo & Schnitzer, 2014; Schnitzer, 2015). This disparity is likely due to disturbance differences between herbaceous and forest habitats, as mowing or grazing is more homogeneous disturbance than gap dynamics in forests, suggesting that heterogeneous disturbances increase the incidence of climbers because they ensure patches of host plants providing external support. This is also supported by our finding of the positive effect of canopy cover on vine occurrence in all datasets.

Our results also suggest that vines are more common at sites with dense vegetation, based on their higher SLA, larger seeds and taller canopies compared with co-occurring herbs (Figure 4). A potential explanation for this shift towards higher LHS traits is that vines require traits connected to shade tolerance because their strategy is dependent (and therefore more sensitive in comparison with other herbs) on dense vegetation with plenty of external support. Shade tolerance is based on many traits, including larger seeds and a higher SLA (reviewed in Gommers, Visser, St Onge, Voesenek, & Pierik, 2013; Valladares & Niinemets, 2008). First, large seeds are particularly important for shade tolerance in the understorey (Manning, Houston, & Evans, 2009), mainly for annual herbs (Bitomský, Mládek, & Cimalová, 2018). During their period in the understorey, vines optimize light capture and utilization with a higher SLA. Second, vines were taller on average (Figure 4). Canopy height is a trait representing shade avoidance and the ability to compete for light, which enables vines to escape from the deep shade of the understorey. These results on the shade tolerance of vines thus provide new insight into the debate regarding whether climbers are light-demanding species or not (e.g., Carrasco-Urra & Gianoli, 2009; Dalling et al., 2012; Schnitzer & Bongers, 2002; Valladares et al., 2011).

TABLE 3 Summary of the generalized mixed-effect model with the best fit describing vine presence/absence in species-rich grasslands (Dataset 2, see Table 2 and Methods for description). The fit was compared with the null model (the model without fixed effect terms including only experimental site as a random term). Standardized coefficients (Coef.), standard errors (SE) and variance inflation factors (VIF) are shown. The estimates are on a logit scale (binomial error structure). Coefficients of determination (marginal and conditional R^2) were computed according to Nakagawa and Schielzeth (2013)

Model fit	Fixed effects	Coef.	SE	*p*-Value	VIF
AIC = 245.6	Mg	0.85	0.485	0.08	1.6
ΔAIC = −13 (vs. null model)	K	0.52	0.287	0.07	1.4
Marginal R^2 = 31.3%	Mg × K	−0.87	0.325	0.008	1.3
Conditional R^2 = 59.3%	Random effect				
	Experimental site	Var = 2.26	SD = 1.51		

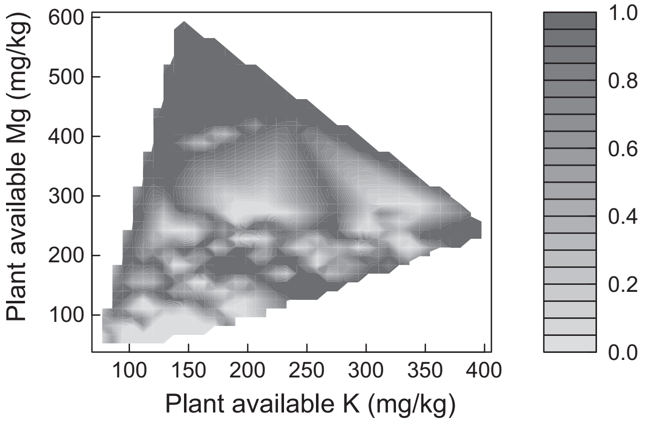


FIGURE 3 Occurrence of vines along plant-available potassium (K) and magnesium (Mg) gradients in soil in species-rich grasslands (Dataset 2, see Table 2 and Methods for description). Darker grey indicates a higher probability of a climber's presence

TABLE 4 Summary of testing whether alpha trait values of vines differ from zero on average (Dataset 1, see Table 2 and Methods2 for description). Significant results indicate differences in a particular trait compared to co-occurring plants. Canopy height, seed mass and specific leaf area were higher (Figure 4). For skewed distributions, we used a non-parametric test (Wilcox). Otherwise, we performed two-tailed *t* tests. Canopy height, seed mass and specific leaf area were log-transformed

Trait	Test (alpha values ≠ 0)	Test statistic (*df* = 30)	*p*-Value
Bud bank depth	*t* test	*t* = −1.03	n.s.
Bud bank size	*t* test	*t* = −1.06	n.s.
Flowering	*t* test	*t* = −1.15	n.s.
Canopy height	Wilcox	*V* = 386	0.006
Leaf dry matter content	*t* test	*t* = −0.37	n.s.
Seed mass	Wilcox	*V* = 461	<0.001
Specific leaf area	*t* test	*t* = 2.27	0.03

In species-rich grasslands, we showed that plant-available Mg is the most important driver of vine occurrence (Table 3, Figure 2). Several studies have described the role of Mg (Addo-Fordjour et al. 2014; Macía et al., 2007), but it has never before been shown to be more influential than more commonly measured nutrients such as N and P. In fact, Homeier et al. (2010) found no effect of Mg on climber density. Magnesium is important for photosynthesis, nucleotide metabolism, and plant growth (Gransee & Führs, 2013; Verbruggen & Hermans, 2013), and has been documented to have higher mass-based concentrations in woody lianas in the tropics (Asner & Martin, 2012). Agricultural research on climbing crops such as peas (*Pisum sativum*) and beans (*Phaseolus vulgaris*) suggests that Mg shortage limits plant growth observable by reduced stem or (mainly) root biomass (Cakmak, Hengeler, & Marschner, 1994; Gransee & Führs, 2013; Verbruggen & Hermans, 2013). Moreover, Nakajima et al. (1981) demonstrated that Mg is responsible for cell wall extensibility in *Pisum sativum*. An extensible stem is crucial to the climbing strategy, especially in the case of vines that must cope with the occasional collapse of their non-woody architectural support in open habitats. It is worth noting that K reduced the positive effect of Mg (Table 3, Figure 3) on vine occurrence, probably through antagonistic cation competition (Gransee & Führs, 2013). When K concentrations are high, Mg uptake by plants is reduced. In summary, we hypothesize that vines require higher concentrations of Mg in order to ensure a fast growth rate and the elongation of stems to successfully reach the top of the canopy.

Large-scale observational studies such as this are often accompanied by limitations, including the possible effects of other non-measured variables and the lack of direct comparisons with other growth forms. Given the high water transport capacity of tropical climbers and their peak of relative abundance in seasonally dry tropical forests (reviewed in Santiago, Pasquini, & De Guzman, 2015) it is possible that our results were confounded by the effect of water availability. Unfortunately, we did not measure any variable related to soil water content and therefore we were not able to address this potentially important feature. We also did not perform comparisons with other growth forms (see Macía et al., 2007) because we were limited by the sampling design (presence/absence data). Some common herbaceous life forms (such as hemicryptophytes, therophytes, or geophytes) are almost ubiquitous in all herbaceous habitats, and we would have needed abundance data for comparisons. In summary, a specifically designed experiment with on-site trait measurements (to

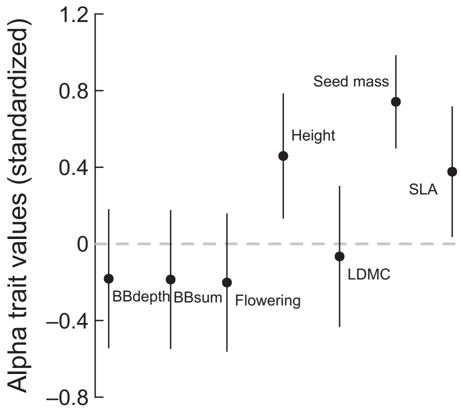


FIGURE 4 Mean alpha trait values (the difference between species' trait values and the mean of co- occurring taxa) of vines. Vertical lines indicate the 95% confidence intervals, which were computed assuming a normal distribution. For skewed distributions (canopy height and seed mass) we used a non-parametric test for significance testing (Table 4). A positive mean with a confidence interval not overlapping zero indicates that vines have a higher trait value than co-occurring species. Traits were standardized. Traits and abbreviations: bud bank depth (BBdepth), bud bank size (BBsum), beginning of flowering (flowering), canopy height (height), leaf dry matter content (LDMC), seed mass, and specific leaf area (SLA). The analysis was based on traits of 31 vines and 1138 co-occurring herbs (Dataset 1, see Table 2 and Methods for description)

acknowledge intraspecific variability) and manipulated environmental conditions could further address the question of whether vines are more or less sensitive to various environmental gradients than other functional groups or life-forms.

5 | CONCLUSIONS

The higher canopy height, seed mass, and SLA of vines along with their more frequent occurrence under denser canopies all suggest that the herbaceous climbing plant strategy is suited for low light conditions. Vines appear to invest more in canopy height (shade avoidance) and seed mass (shade tolerance), and they tend to have a slightly higher SLA (and shade tolerance) in comparison with co-occurring plants. We did not find any direct role of disturbance in the occurrence of vines. In species-rich grasslands, we discovered that plant-available Mg played a prominent role in the prediction of vine occurrence. This higher demand for Mg in vines might be explained by its role in supporting growth rate and stem extensibility. Moreover, this finding suggests that testing the effects of nutrients other than nitrogen and phosphorus (most often used in the ecological literature) can provide novel information about the distribution of growth forms along soil fertility gradients. We believe this study is an important step in understanding how climbers respond to ecological gradients and whether climbers display a different trait-based strategy compared to other growth forms in the same habitat.

ACKNOWLEDGEMENTS

We are grateful to Martin Kočí, Erika Lošáková and David Opálka for help with field measurements in species-rich grasslands, to Milan Chytrý for providing data from the Czech phytosociological database, and to David Hardekopf for English editing. The clarity of the manuscript was substantially improved thanks to the suggestions of the editor Jodi Price and two anonymous reviewers.

AUTHOR CONTRIBUTIONS

M.B. and J.M. conceived the ideas. J.M., P.M. and S.C. designed the methodology and collected data. M.B. analysed data with contributions from P.M. M.B. wrote the manuscript with contributions from all authors. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY

All data supporting the results are archived in the Mendeley Data depository (https://data.mendeley.com). https://doi.org/10.17632/2mj8tmgdzv.2

ORCID

Martin Bitomský https://orcid.org/0000-0002-0022-8473
Jan Mládek https://orcid.org/0000-0002-8301-4824

REFERENCES

Ackerly, D. D., & Cornwell, W. K. (2007). A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecology Letters*, *10*, 135–145. https://doi.org/10.1111/j.1461-0248.2006.01006.x

Addo-Fordjour, P., Rahmad, Z. B., & Shahrul, A. M. S. (2014). Environmental factors influencing liana community diversity, structure and habitat associations in a tropical hill forest, Malaysia. *Plant Ecology & Diversity*, *7*, 485–496. https://doi.org/10.1080/17550874.2013.782369

AOAC (1984). *Official methods of analysis* (14th ed.). Arlington, VA: AOAC.

Asner, G. P., & Martin, R. E. (2012). Contrasting leaf chemical traits in tropical lianas and trees: implications for future forest composition. *Ecology Letters*, *15*, 1001–1007. https://doi.org/10.1111/j.1461-0248.2012.01821.x

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bitomský, M., Mládek, J., & Cimalová, Š. (2018). Light limitation shapes the community seed mass of annual but not of perennial weeds. *Perspectives in Plant Ecology, Evolution and Systematics*, *31*, 1–6. https://doi.org/10.1016/j.ppees.2017.11.005

Black, C. A. (1965). *Methods of soil analysis, part I and II*. Madison, WI: American Society of Agronomy, Inc.

Bocci, G. (2015). TR8: An R package for easily retrieving plant species traits. *Methods in Ecology and Evolution*, *6*, 347–350. https://doi.org/10.1111/2041-210x.12327

Cakmak, I., Hengeler, C., & Marschner, H. (1994). Partitioning of shoot and root dry matter and carbohydrates in bean plants suffering from phosphorus, potassium and magnesium deficiency. *Journal* *of Experimental Botany*, *45*, 1245–1250. https://doi.org/10.1093/jxb/45.9.1245

Carrasco-Urra, F., & Gianoli, E. (2009). Abundance of climbing plants in a southern temperate rain forest: host tree characteristics or light availability? *Journal of Vegetation Science*, *20*, 1155–1162. https://doi.org/10.1111/j.1654-1103.2009.01115.x

Chevan, A., & Sutherland, M. (1991). Hierarchical partitioning. *The American Statistician*, *45*, 90–96.

Chytrý, M., Pyšek, P., Tichý, L., Knollová, I., & Danihelka, J. (2005). Invasions by alien plants in the Czech Republic: A quantitative assessment across habitats. *Preslia*, *77*, 339–354.

Chytrý, M., & Rafajová, M. (2003). Czech National Phytosociological Database: Basic statistics of the available vegetation-plot data. *Preslia*, *75*, 1–15.

Cimalová, Š., & Lososová, Z. (2009). Arable weed vegetation of the northeastern part of the Czech Republic: effects of environmental factors on species composition. *Plant Ecology*, *203*, 45–57. https://doi.org/10.1007/s11258-008-9503-1

Cortés-Flores, J., Hernández-Esquivel, K. B., González-Rodríguez, A., & Ibarra-Manríquez, G. (2017). Flowering phenology, growth forms, and pollination syndromes in tropical dry forest species: Influence of phylogeny and abiotic factors. *American Journal of Botany*, *104*, 39–49. https://doi.org/10.3732/ajb.1600305

Dalling, J. W., Schnitzer, S. A., Baldeck, C., Harms, K. E., John, R., Mangan, S. A., … Hubbel, S. P. (2012). Resource-based habitat associations in a neotropical liana community. *Journal of Ecology*, *100*, 1174–1182. https://doi.org/10.1111/j.1365-2745.2012.01989.x

Gallagher, R. V., & Leishman, M. R. (2012). Contrasting patterns of trait-based community assembly in lianas and trees from temperate Australia. *Oikos*, *121*, 2026–2035. https://doi.org/10.1111/j.1600-0706.2012.20025.x

Gommers, C. M. M., Visser, E. J. W., St Onge, K. R., Voesenek, L. A. C. J., & Pierik, R. (2013). Shade tolerance: When growing tall is not an option. *Trends in Plant Science*, *18*, 65–71. https://doi.org/10.1016/j.tplants.2012.09.008

Gransee, A., & Führs, H. (2013). Magnesium mobility in soils as a challenge for soil and plant analysis, magnesium fertilization and root uptake under adverse growth conditions. *Plant and Soil*, *368*, 5–21. https://doi.org/10.1007/s11104-012-1567-y

Grime, J. P. (1977). Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *American Naturalist*, *111*, 1169–1194. https://doi.org/10.1086/283244

Herben, T., Chytrý, M., & Klimešová, J. (2016). A quest for species-level indicator values for disturbance. *Journal of Vegetation Science*, *27*, 628–636. https://doi.org/10.1111/jvs.12384

Herben, T., Klimešová, J., & Chytrý, M. (2018). Effects of disturbance frequency and severity on plant traits: An assessment across a temperate flora. *Functional Ecology*, *32*, 799–808. https://doi.org/10.1111/1365-2435.13011

Homeier, J., Englert, F., Leuschner, C., Weigelt, P., & Unger, M. (2010). Factors controlling the abundance of lianas along an altitudinal transect of tropical forests in Ecuador. *Forest Ecology and Management*, *259*, 1399–1405. https://doi.org/10.1016/j.foreco.2010.01.008

Hu, L., Li., M., & Li., Z. (2010). Geographical and environmental gradients of lianas and vines in China. *Global Ecology and Biogeography*, *19*, 554–561. https://doi.org/10.1111/j.1466-8238.2010.00527.x

Josse, J., & Husson, F. (2012). Handling missing values in exploratory multivariate data analysis methods. *Journal de la Société Française de Statistique*, *153*, 79–99.

Josse, J., & Husson, F. (2016). missMDA: A package for handling missing values in multivariate data analysis. *Journal of Statistical Software*, *70*, 1–31. https://doi.org/10.18637/jss.v070.i01

Kazda, M. (2015). Liana–nutrient relations. In S. A. Schnitzer, F. Bongers, R. J. Burnham, & F. E. Putz (Eds.), *Ecology of lianas* (pp. 309–322). Oxford, UK: Wiley-Blackwell.

Kleyer, M., Bekker, R. M., Knevel, I. C., Bakker, J. P., Thompson, K., Sonnenschein, M., … Peco, B. (2008). The LEDA Traitbase: A database of life-history traits of the Northwest European flora. *Journal of Ecology*, *96*, 1266–1274. https://doi.org/10.1111/j.1365-2745.2008.01430.x

Klimešová, J., Danihelka, J., Chrtek, J., de Bello, F., & Herben, T. (2017). CLO-PLA: A database of clonal and bud-bank traits of the Central European flora. *Ecology*, *98*, 1179. https://doi.org/10.1002/ecy.1745

Klimešová, J., & Klimeš, L. (2007). Bud banks and their role in vegetative regeneration — A literature review and proposal for simple classification and assessment. *Perspectives in Plant Ecology, Evolution and Systematics*, *8*, 115–129. https://doi.org/10.1016/j.ppees.2006.10.002

Klotz, S., Kühn, I., Durka, W., &Briemle, G. (2002). BIOLFLOR – Eine Datenbank mit biologisch-ökologischen Merkmalen zur Flora von Deutschland. *Schriftenreihe für Vegetationskunde*, Vol. *38*. Hiltrup, Germany: Bundesamt für Naturschutz.

Ledo, A., & Schnitzer, S. A. (2014). Disturbance and clonal reproduction determine liana distribution and maintain liana diversity in a tropical forest. *Ecology*, *95*, 2169–2178. https://doi.org/10.1890/13-1775.1

Macía, M. J., Ruokolainen, K., Tuomisto, H., Quisbert, J., & Cala, V. (2007). Congruence between floristic patterns of trees and lianas in a southwest Amazonian rain forest. *Ecography*, *30*, 561–577. https://doi.org/10.1111/j.2007.0906-7590.05124.x

Manning, P., Houston, K., & Evans, T. (2009). Shifts in seed size across experimental nitrogen enrichment and plant density gradients. *Basic and Applied Ecology*, *10*, 300–308. https://doi.org/10.1016/j.baae.2008.08.004

Mehlich, A. (1984). Mehlich no 3 soil test extractant: A modification of Mehlich no 2 extractant. *Communication in Soil Science and Plant Analysis*, *15*, 1409–1416. https://doi.org/10.1080/00103628409367568

Mládková, P., Mládek, J., Hejduk, S., Hejcman, M., Cruz, P., Jouany, C., & Pakeman, R. J. (2015). High-nature-value grasslands have the capacity to cope with nutrient impoverishment induced by mowing and livestock grazing. *Journal of Applied Ecology*, *52*, 1073–1081. https://doi.org/10.1111/1365-2664.12464

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nakajima, N., Morikawa, H., Igarashi, S., & Senda, M. (1981). Differential effect of calcium and magnesium on mechanical properties of pea stem cell walls. *Plant & Cell Physiology*, *22*, 1305–1315. https://doi.org/10.1093/oxfordjournals.pcp.a076283

Pescador, D. S., de Bello, F., Valladares, F., & Escudero, A. (2015). Plant trait variation along an altitudinal gradient in Mediterranean high mountain grasslands: controlling the species turnover effect. *PLoS One*, *10*, e0118876. https://doi.org/10.1371/journal.pone.0118876

Pierce, S., Negreiros, D., Cerabolini, B. E. L., Kattge, J., Díaz, S., Kleyer, M., … Tampucci, D. (2017). A global method for calculating plant CSR ecological strategies applied across biomes world-wide. *Functional Ecology*, *31*, 444–457. https://doi.org/10.1111/1365-2435.12722

Putz, F. E., & Windsor, D. M. (1987). Liana phenology on Barro Colorado Island, Panama. *Biotropica*, *19*, 334–341. https://doi.org/10.2307/2388631

Santiago, L. S., Pasquini, S. C., & De Guzman, M. E. (2015). Physiological implications of the liana growth form. In S. A. Schnitzer, F. Bongers,

R. J. Burnham, & F. E. Putz (Eds.), *Ecology of lianas* (pp. 288–298). Oxford, UK: Wiley-Blackwell.

Schnitzer, S. A. (2015). Increasing liana abundance in neotropical forests: Causes and consequences. In S. A. Schnitzer, F. Bongers, R. J. Burnham, & F. E. Putz (Eds.), *Ecology of lianas* (pp. 451–464). Oxford, UK: Wiley-Blackwell.

Schnitzer, S. A., & Bongers, F. (2002). The ecology of lianas and their role in forests. *Trends in Ecology and Evolution*, *17*, 223–230. https://doi.org/10.1016/s0169-5347(02)02491-6

Valladares, F., Gianoli, E., & Saldaña, A. (2011). Climbing plants in a temperate rainforest understorey: searching for high light or coping with deep shade? *Annals of Botany*, *108*, 231–239. https://doi.org/10.1093/aob/mcr132

Valladares, F., & Niinemets, Ü. (2008). Shade tolerance, a key plant feature of complex nature and consequences. *Annual Review of Ecology, Evolution, and Systematics*, *39*, 237–257. https://doi.org/10.1146/annurev.ecolsys.39.110707.173506

Van Der Heijden, G. M. F., & Phillips, O. L. (2008). What controls liana success in Neotropical forests? *Global Ecology and Biogeography*, *17*, 372–383. https://doi.org/10.1111/j.1466-8238.2007.00376.x

van der Maarel, E. (1979). Transformation of cover–abundance values in phytosociology and its effects on community similarity. *Vegetatio*, *39*, 97–114. https://doi.org/10.1007/bf00052021

Verbruggen, N., & Hermans, C. (2013). Physiological and molecular responses to magnesium nutritional imbalance in plants. *Plant and Soil*, *368*, 87–99. https://doi.org/10.1007/s11104-013-1589-0

Westoby, M. (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, *199*, 213–227. https://doi.org/10.1023/a:1004327224729

Wyka, T. P., Oleksyn, J., Karolewski, P., & Schnitzer, S. A. (2013). Phenotypic correlates of the lianescent growth form: a review. *Annals of Botany*, *112*, 1667–1681. https://doi.org/10.1093/aob/mct236

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Summary of herbaceous habitats where at least one vine occurred (Dataset 1).

Appendix S2. List of all vines native in the Czech Republic with an indication of their presence in all datasets.

Appendix S3. Location and GPS coordinates of the experimental sites in species-rich grasslands (Dataset 2).

Appendix S4. Vines' response to plant-available nutrients at the species level (Dataset 2).

Appendix S5. Response of vines to plant-available Mg and K concentrations at species level in species-rich grasslands (Dataset 2).

Appendix S6. Results of the fourth-corner analyses within Fabaceae (Dataset 2).

Appendix S7. Magnesium concentration gradients within the experimental sites in species-rich grasslands (Dataset 2).

Appendix S8. Assessment of the model fit of linear versus phylogenetic linear models testing the trait differences of vines compared to co-occurring herbs (Dataset 1).

How to cite this article: Bitomský M, Mládková P, Cimalová Š, Mládek J. Herbaceous climbers in herbaceous systems are shade-tolerant and magnesium-demanding. *J Veg Sci*. 2019;00:1–10. https://doi.org/10.1111/jvs.12768

Graphical Abstract

The contents of this page will be used as part of the graphical abstract of html only. It will not be published as part of main article.

Climbers in forests are more frequent in canopy gaps with high light and nutrient availability; however, in herbaceous systems, disturbances are homogeneous. Vines were not associated with disturbance but were positively associated with the availability of external support and magnesium. Vines cope with shade as they are taller, with greater seed mass and specific leaf area than co-occurring herbs.

Received: 11 July 2019 | Revised: 21 January 2020 | Accepted: 21 February 2020

DOI: 10.1002/ece3.6170

ORIGINAL RESEARCH

Ecology and Evolution Open Access WILEY

Clade composition of a plant community indicates its phylogenetic diversity

Martin Bitomský[1] | **Pavla Mládková**[2] | **Robin J. Pakeman**[3] | **Martin Duchoslav**[4]

[1]Department of Ecology and Environmental Sciences, Palacky University, Olomouc, Czech Republic

[2]Department of Ecology, Czech University of Life Sciences, Prague, Czech Republic

[3]The James Hutton Institute, Aberdeen, UK

[4]Department of Botany, Palacky University, Olomouc, Czech Republic

Correspondence
Martin Bitomský, Department of Ecology and Environmental Sciences, Palacky University, Slechtitelu 27, CZ-783 71 Olomouc, Czech Republic.
Email: bitomsky.martin@gmail.com

Funding information
Palacký University, Grant/Award Number: IGA_PrF_2018_020 and IGA_PrF_2019_021; Ministry of the Environment of the Czech Republic, Grant/Award Number: RPV SP/2D3/179/07; Grant Agency of the Czech University of Life Sciences, Grant/Award Number: 42110/1313/3114; Ministry of Agriculture of the Czech Republic, Grant/Award Number: 5400/2013-MZE-13212

Abstract

Phylogenetic diversity quantification is based on indices computed from phylogenetic distances among species, which are derived from phylogenetic trees. This approach requires phylogenetic expertise and available molecular data, or a fully sampled synthesis-based phylogeny. Here, we propose and evaluate a simpler alternative approach based on taxonomic coding. We developed metrics, the clade indices, based on information about clade proportions in communities and species richness of a community or a clade, which do not require phylogenies. Using vegetation records from herbaceous plots from Central Europe and simulated vegetation plots based on a megaphylogeny of vascular plants, we examined fit accuracy of our proposed indices for all dimensions of phylogenetic diversity (richness, divergence, and regularity). For real vegetation data, the clade indices fitted phylogeny-based metrics very accurately (explanatory power was usually higher than 80% for phylogenetic richness, almost always higher than 90% for phylogenetic divergence, and often higher than 70% for phylogenetic regularity). For phylogenetic regularity, fit accuracy was habitat and species richness dependent. For phylogenetic richness and divergence, the clade indices performed consistently. In simulated datasets, fit accuracy of all clade indices increased with increasing species richness, suggesting better precision in species-rich habitats and at larger spatial scales. Fit accuracy for phylogenetic divergence and regularity was unreliable at large phylogenetic scales, suggesting inadvisability of our method in habitats including many distantly related lineages. The clade indices are promising alternative measures for all projects with a phylogenetic framework, which can trade-off a little precision for a significant speed-up and simplification, such as macroecological analyses or where phylogenetic data is incomplete.

KEYWORDS
biodiversity, clade index, phylogenetic divergence, phylogenetic regularity, phylogenetic richness

1 | INTRODUCTION

The concept of phylogenetic diversity has revolutionized research in nature conservation and community ecology, as it enables the setting of conservation priorities or helps to identify which community assembly processes may have structured a community (Faith, 1992; Webb, Ackerly, McPeek, & Donoghue, 2002). Phylogenetic diversity estimation is based on phylogenetic distances (the

---

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2020 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

amount of time since the most common ancestor of a pair of species), which are derived from dated phylogenies. Researchers have developed more than 70 metrics for quantifying alpha (within-site) and beta (among sites) phylogenetic diversity, which are summarized under several frameworks (Scheiner, Kosman, Presley, & Willig, 2017; Tucker et al., 2017). It is worth noting that there is no agreement on the best or the most suitable metric. Phylogenetic diversity reflects diversification of lineages, geographic movement of lineages, and deep-past and present assembly processes (Gerhold, Carlucci, Proches, & Prinzing, 2018; Webb et al., 2002; Yguel et al., 2016) that can be lineage specific (Elliott, Waterway, & Davies, 2016; Ndiribe et al., 2013). Considering such complexity, it is not possible to address phylogenetic patterns in communities using only one number. Therefore, this plethora of metrics is inevitable because each metric was designed to capture a specific aspect of phylogenetic diversity. Fortunately, various phylogenetic diversity metrics tend to correlate (Swenson, 2014; Vellend, Cornwell, Magnuson-Ford, & Mooers, 2011) suggesting redundancy of some of them, thus, there has been an attempt to select a leading measure for each dimension of phylogenetic diversity (richness, divergence, and regularity; sensu Tucker et al., 2017; Table 1).

To construct dated phylogenies requires considerable effort, and the whole process is affected by methodological biases and subjective decisions (Jantzen et al., 2019; Li et al., 2019). Further, calculated phylogenetic diversity metrics depend on the attributes of phylogenies, such as the degree of balance, diversification rate, resolution, taxon sampling, or tree reconstruction methods (Jantzen et al., 2019; Park, Worthington, & Xi, 2018; Swenson, 2009; Vellend et al., 2011). Here, we propose and evaluate an approach based on the idea of considering species phylogeny as a categorical variable (i.e., affiliation to a phylogenetic clade) rather than continuous (i.e., phylogenetic distances among species). A similar approach based on taxonomic relatedness (derived from a hierarchical Linnaean classification with applied taxonomic weights proportional to the level of the taxonomic rank two species hold in common, i.e., genus, family, or order) has proven to be useful to estimate biodiversity patterns in fish communities (Campbell, Neat, Burns, & Kunzlik, 2010; Hall & Greenstreet, 1998; Warwick & Clarke, 1995). There is also a clear parallel in functional ecology, clades can be considered as analogous to plant functional types (PFT) and their proportions can be utilized to indicate phylogenetic diversity of a community. Such a categorical approach to phylogeny might be a tool for ecologists who are not specialists in phylogenetics and might be useful in communities where some taxa do not have available DNA sequences or in studies where a little precision can be traded-off for significant speed-up and simplification.

This framework certainly causes a loss of information as we basically introduce a polytomy at a node of a defined clade, i.e. the categorical approach still separates species according to their clade affiliation, but it ignores phylogenetic information within clades. On the other hand, there is some indirect support that this loss of phylogenetic information within clades would have a marginal effect. Li et al. (2019) compared purpose-built phylogenies (estimated from sequence data) with published synthesis-based supertrees (which usually have more polytomies than the former) and showed that phylogenetic diversity metrics computed from both types of phylogenies were highly correlated. Cadotte (2015) also demonstrated that changing branch lengths did not strongly affect relationships between phylogenetic diversity and ecosystem function, suggesting that phylogenetic diversity measures are not so sensitive to the branch lengths of the phylogeny as long as the topology is right. One important criterion for choosing among metrics is their conceptual and mathematical simplicity (Vellend et al., 2011). Therefore, if the categorical approach provides sufficiently correlated values with other phylogeny-based measures, than its use can be justified in order to simplify and speed-up phylogenetic diversity estimation.

The phylogenetic categorical approach cannot rely on phylogenetic distances, but we can include information about how clades are represented in a community (presence and relative abundance) to estimate its phylogenetic diversity. Consider a simple example phylogeny of 10 species (Figure 1a), which covers all major clades of the whole species pool of our first case study (Figure S1). We simulated 1,000 communities where these 10 species occurred, but we let their proportions in a community randomly vary. For each community, we estimated phylogenetic richness, divergence, and regularity (sensu Tucker et al., 2017) using a leading metric of each dimension (see Methods for more information). Visual inspection of phylogeny-based measures showed several interesting features. Phylogenetic richness increased with increasing proportion of the most distantly related species (*Ranunculus repens* in this case) in comparison with the rest of the species in the community (Figure 1b). Phylogenetic divergence was relatively high when all defined clades (i.e., monocots, Ranunculales, superrosids, and superasterids) had equal proportions (Figure 1c). Finally, phylogenetic regularity was relatively high (i.e., the variance of phylogenetic distances was low) when the defined clades had proportions proportional to their relative species richness in the species pool (Figure 1d).

Based on the conclusions from the conceptual example described above, we propose here three alternative measures, the clade indices that do not require dated phylogenies for their computation, but instead they utilize information about clade proportions in a community and species richness of a community or defined clades (Table 2). We assessed their fit accuracy for leading phylogeny-based measures of the three dimensions of phylogenetic diversity: richness, divergence, and regularity (sensu Tucker et al., 2017). To do so, we examined the performance of the proposed clade indices in two case studies, firstly with a dataset with a purpose-built phylogeny (sensu Li et al., 2019) consisting of relatively small number of taxa in the species pool and second dataset with a synthesis-based phylogeny (sensu Li et al., 2019) consisting of relatively large number of taxa in the species pool. In this first case study,

TABLE 1 Summary of three dimensions of phylogenetic diversity (defined by Tucker et al., 2017)

Dimension	Richness	Divergence	Regularity
Leading metric	Faith's phylogenetic diversity (Faith's PD)	Mean pairwise distance (MPD)	Variation of pairwise distances (VPD)
Mathematical function	Sum	Mean distance	Variation
Indicator of	Total evolutionary history	Similarity (phylogenetic relatedness)	Distribution of phylogenetic similarity
Main use	Conservation, predictor of future evolutionary potential	Proxy of trait similarity, test of habitat filtering versus limiting similarity	Testing competitive interactions
Example of a community with high value	Species-rich communities	Clade-rich communities	Communities with low asymmetric competition

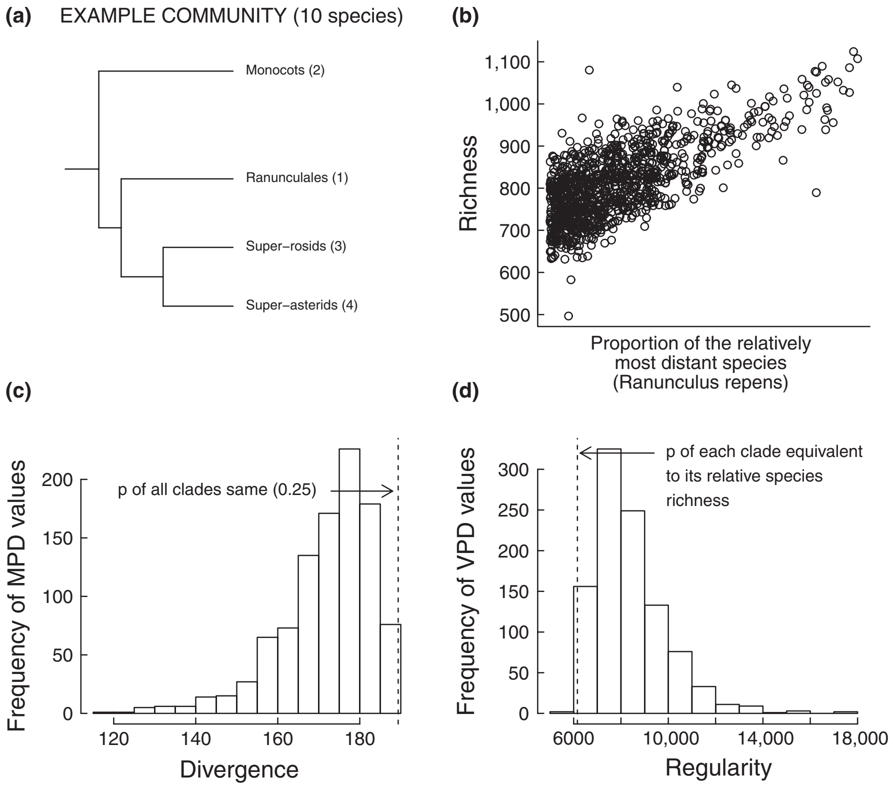


FIGURE 1 A conceptual example demonstrating how clade proportions (relative cover) affect values of leading metrics of all dimensions of phylogenetic diversity (Faith's PD = richness, MPD = divergence, and VPD = regularity). (a) We randomly selected 10 species: two monocots (*Agrostis capillaris* L. and *Bromus erectus* Huds.), one Ranunculales (*Ranunculus repens* L.), three superrosids (*Fragaria viridis* Weston, *Trifolium pratense* L., and *Vicia cracca* L.), and four superasterids (*Aegopodium podagraria* L., *Centaurea jacea* L., *Campanula patula* L., and *Plantago major* L.) in order to cover all major clades of the whole species pool (Figure S1). The number of species in each clade approximately reflects relative species richness of clades of the species pool of the case study in species-rich grasslands. Then, we simulated 1,000 communities using all the 10 species and let their proportions randomly vary. Phylogenetic richness, divergence, and regularity were estimated for each simulated community. (b) Faith's PD particularly increased with increasing proportion of *R. repens* (i.e., the relatively most phylogenetically distant species compared to the rest). Distant branches contribute more to phylogenetic richness as they are longer, suggesting that increase in their weight (reflecting species proportion in a community) also increases phylogenetic richness of a community. (c) Histogram of simulated MPD values. MPD of a community when all four clades are equally abundant ($p_{\text{monocots}} = p_{\text{Ranunculales}} = p_{\text{superrosids}} = p_{\text{superasterids}}$) is indicated. (d) Histogram of simulated VPD values. VPD of a community when each species has same proportion (i.e., proportion of each clade is equivalent to its relative species richness) is indicated

we also examined what clade resolution (at the super-order, order, and family level) for the clade index definition is the most suitable in terms of fit accuracy for phylogeny-based measures. Secondly, we used simulated community matrices based on a megaphylogeny of 31,389 vascular plants (Qian & Jin, 2016) to demonstrate how the clade indices perform at various phylogenetic scales (Graham, Storch, & Machac, 2018), at different species pool sizes and along a species richness gradient.

2 | MATERIALS AND METHODS

2.1 | Data collecting

The focus of the case studies was on herbaceous terrestrial systems. First, we used data from species-rich grasslands located in two Protected Landscape Areas on the border between the Czech Republic and Slovakia: Beskydy Mountains (N 49.45°, E 18.33°) and White Carpathian Mountains (N 48.97°, E 17.82°). We collected vegetation records in 240 permanent plots (1 × 1 m in size) in 12 long-term management experiments (hereafter exclosures) at six localities (Table S1) in 2013. Community data included 171 plant species. Second, we assembled vegetation plots from a stratified dataset (for detailed information, see Chytrý, Pyšek, Tichý, Knollová, & Danihelka, 2005) extracted from the Czech National Phytosociological Database (hereafter CNPD; Chytrý & Rafajová, 2003). This dataset included 16,542 plots and 1,608 species and covered 26 Central European herbaceous habitats (see Table S2 for a habitat classification). We limited our analysis to herbaceous angiosperms that dominate all systems used in this study. In the grassland dataset, tree taxa were omitted in the initial phase of the vegetation recording, but this most likely did not affect estimation of phylogenetic diversity as we found only a few tree seedlings in a few plots. We deleted Pteridophyta from both datasets, whereas gymnosperms did not occur in any dataset.

2.2 | Phylogenetic inference and molecular dating

Prior to the phylogenetic analysis, we checked species lists and edited some species names in order to follow the NCBI nomenclature. For the species-rich grasslands, we constructed a molecular-based phylogeny for our 171 species using 20 orthologous loci downloaded from GenBank (Benson et al., 2017) via an online tool OneTwoTree (Drori et al., 2018). We used *Piper nigrum* L. from the Magnoliids group (a sister clade to clades occurring in our dataset; APG IV, 2016) as an out-group. Due to missing sequence data, we replaced *Potentilla heptaphylla* L. with a relatively close congener *Potentilla crantzii* (Crantz) Beck ex Fritsch (Dobeš, Rossa, Paule, & Hülber, 2013) that had available DNA data. Sequences were aligned using a fast option (FFT-NS-2) in MAFFT (Katoh & Standley, 2013) under default settings available at the OneTwoTree website (6mer pairwise alignment method). The alignment was then cured using the Gblocks online tool (under less stringent selection settings; Castresana, 2000).

We constructed the dated tree using BEAST version 1.10.4 (Suchard et al., 2018) in the CIPRES portal (Miller, Pfeiffer, & Schwartz, 2010). To do so, we manually set constraints according to the APG IV angiosperm phylogeny (APG IV, 2016) and set the uncorrelated relaxed clock as a clock model, Yule process as a speciation model and GTR+G+I (with four gamma categories) as a nucleotide substitution model. To translate genetic distances into absolute times, we exploited the TimeTree database (Kumar, Stecher, Suleski, & Hedges, 2017) and set several time priors with normally distributed errors (median and standard deviation computed from all studies available in the TimeTree database reporting a given divergence time estimate). We performed three independent runs (with different starting seeds) for 100 million generations each. Finally, we checked convergence in Tracer v1.7.1 (Rambaut, Drummond, Xie, Baele, & Suchard, 2018) and combined all runs (10% generations as a burn-in). The dated maximum clade credibility tree (Figure S1) was sampled from 30,000 trees (10% trees as a burn-in).

For the species in the dataset from the CNPD, we extracted species phylogeny from the dated supertree of the European flora (Durka & Michalski, 2012) and followed their nomenclature.

2.3 | Phylogenetic diversity dimensions and metrics

We applied the framework of Tucker et al. (2017) and selected three leading metrics describing three phylogenetic diversity dimensions:

TABLE 2 Summary of the proposed clade indices

Index	Equation	Treatment	Rationale
(a) Clade richness	$\log(S) + 3 \cdot \sum_{i=1}^{k} \frac{p_i}{\mathrm{CR}_i}$	Species-rich clades are penalized as they get lower weight proportional to their clade richness. Higher proportions of species-poor clades increase the clade richness index values	Species from species-poor clades have higher probability to be relatively phylogenetically distant to the rest of a community and their increasing proportion increases phylogenetic richness of a community (Figure 1b)
(b) Clade divergence	$1 - \sum_{i=1}^{k} \left(p_i - \frac{1}{\mathrm{CR}_{\mathrm{SP}}}\right)^2$	Larger deviations from optimal proportions (i.e., 1/number of defined clades in the whole species pool) decrease the value of the clade divergence index. Scales from 0 to 1	Phylogenetic divergence tends to be close to its peak when a community consists of all clades of a species pool and their proportions are equal (Figure 1c)
(c) Clade regularity	$1 - \sum_{i=1}^{k} \left(p_i - \frac{\mathrm{CR}_i}{S_{\mathrm{SP}}}\right)^2$	Larger deviations from the optimal proportions (i.e., clade species richness/total species pool richness) decrease the value of the clade regularity index. Scales from 0 to 1	Phylogenetic regularity tends to be close to its peak (the lowest VPD) when a community consists of all clades of a species pool and their proportions are proportional to their relative clade richness given a species pool (Figure 1d)

Note: S = species richness of a plot; p_i = proportion of the *i*th clade in a plot; CR_i = species richness of the *i*th clade in the whole species pool (all species in the dataset); $\mathrm{CR}_{\mathrm{SP}}$ = the number of all defined clades in the whole species pool; S_{SP} = species richness of the whole species pool.

richness, divergence, and regularity (Table 1). Faith's PD (Faith, 1992) describes the amount of evolutionary history across species (sum of branch lengths) and is a leading measure of phylogenetic richness. Mean phylogenetic distance between each pair of species (MPD; Webb et al., 2002) is a leading measure of phylogenetic divergence. Variation of pairwise phylogenetic distances between each pair of species (VPD; Clarke & Warwick, 2001) is a leading measure of phylogenetic regularity (lower variation indicates higher regularity). We also identified species richness in each plot.

According to Vellend et al. (2011), one can distinguish two qualitatively different types of phylogenetic diversity indices. Faith's PD, MPD, and VPD are type II metrics which are calculated using a subset phylogeny of a focal subset of species (e.g., a vegetation plot). Type I indices are based on the whole species pool phylogeny; each species has its distinctness score calculated. These scores are then used to calculate a phylogenetic diversity measure of a plot (for example, summed evolutionary distinctiveness; Redding & Mooers, 2006). However, type I indices are highly correlated with Faith's PD (Vellend et al., 2011), suggesting they are closely related to the phylogenetic richness dimension, and so we did not consider them. We calculated indices using functions (*pd* and *mpd*) from the *picante* package (Kembel et al., 2010). To compute VPD, we modified the *mpd* function to calculate the variation of pairwise phylogenetic distances (not the mean as in the original function). All metrics were abundance weighted by percentage cover. To calculate abundance-weighted Faith's PD (Barker, 2002), we used the R function of Swenson (2014).

2.4 | Definition of the clade indices

Species affiliation to a clade was based on the recent APG IV classification (APG IV, 2016). The proposed clade indices are summarized in Table 2. They all need information about clade proportions in a community (e.g., relative cover, biomass or abundances). The key idea behind the clade richness index is to penalize proportions of species-rich clades (by reverse clade species richness) because species from species-rich clades are unlikely to be relatively distantly related to the rest of co-occurring species in a community. By chance, more species from a species-rich clade can occur in a community, which would decrease phylogenetic richness as these species are relatively closely related. Species richness can be a very good indicator of phylogenetic richness by its own (Swenson, 2014; Vellend et al., 2011); hence, it is useful to include it in the equation (Table 2a). For phylogenetic divergence, when clades are equally abundant in a community, phylogenetic divergence is close to its peak (Figure 1c). Thus, any deviations from these equal proportions should decrease phylogenetic divergence (Table 2b). For instance, if all clades are present and have equal (i.e., optimal) proportions, the clade divergence index equals one. Finally, the clade regularity index has a similar computation to the clade divergence index, but the optimal proportions are proportional to the relative clade species richness (Table 2c). An R script for computation of the clade indices is stored in the supplemental dataset (https://data.mendeley.com/datasets/gbv472pxsb/1).

2.5 | Performance of the clade indices: case studies

We did all statistical analyses and data simulations in R version 3.6.0. (R Core Team, 2019). Faith's PD was square-root transformed, and VPD was log-transformed prior to the analysis. First, we examined how the different phylogenetic resolutions affect values of the clade indices and their correlations with phylogeny-based indices. To do so, we used the grassland dataset and tested three clade resolutions: (a) super-order level (monocots, Ranunculales, superrosids, and superasterids), (b) order (based on affiliation to 20 orders), and (c) family (based on affiliation to 32 families). We calculated the clade indices and assessed their fit of phylogeny-based indices using linear models (the *lm* function in R) and estimated R^2 values. We also checked for the significance of quadratic terms. In the case of phylogenetic regularity, we used generalized least squares models (the *gls* function, *nlme* package; Pinheiro, Bates, DebRoy, & Sarkar, 2019) to acknowledge heteroscedasticity (using the exponential variance class, *varExp*), which we detected during the model diagnostics.

2.6 | Performance of the clade indices: simulated datasets

Simulation workflow was specifically designed to cover several aspects that can affect phylogenetic diversity estimation, that is, taxon sampling (Park et al., 2018), the number of taxa included in the regional phylogeny (Jantzen et al., 2019) or species richness of a community (Sandel, 2018; Swenson, 2014). Thus, these factors could also affect fit accuracy of the clade indices for all dimensions of phylogenetic diversity. The simulation workflow is summarized in Figure S2. Simulation was based on a megaphylogeny of vascular plants (Zanne et al., 2014, updated by Qian & Jin, 2016). We set three phylogenetic scales: vascular plants, angiosperms, and superasterids. For each phylogenetic scale, we set three species pool sizes: 2,000, 500, and 250 species. These species pools were created by randomly assigning species from a given phylogeny (vascular plants, angiosperms, or superasterids). For each combination of phylogenetic scale and species pool size, we generated community matrices under several species richness ranges: 10–160, 10–80, 10–40, 10–20, 5–10, and 2–5 species per community. For each species richness range, we generated 50 community matrices with 240 sites (same data size as in the grassland case study). Species proportions in communities were random but their sums were always one. In total, we generated 2,700 unique species pools with 2,700 unique corresponding community matrices (900 for each phylogenetic scale).

For each community matrix, we computed both phylogeny-based metrics (Faith's PD, MPD and VPD) and the clade indices. Then, we performed linear models with phylogeny-based metrics as response variables and clade indices as explanatory variables and extracted each models R^2 values. Faith's PD was always square-root transformed; VPD was always log-transformed. To assess the importance of all determinants potentially affecting the relationship between phylogeny-based metrics and the clade indices, we calculated relative variances of R^2 values attributed to either phylogenetic scale, species pool size or species richness range using the *VarCorr* function (*nlme* package, Pinheiro et al., 2019). The determinants were hierarchically structured in the model random-effect formula (phylogenetic scale/species pool size/species richness range).

3 | RESULTS

For all phylogenetic diversity dimensions, fit accuracy of the clade indices increased with fineness of phylogenetic resolution in species-rich grasslands (Table S3); hence, we present here the clade indices based on the resolution at the family level in all case studies and simulated communities. For phylogenetic richness and divergence, the fit was reasonably high and similar in both case studies (Figure 2a–d), and in all herbaceous habitats (the CNPD dataset) when fitted separately (Figures S3 and S4). For phylogenetic regularity, fit accuracy increased with increasing values of the family regularity index (Figure 2e,f) as the relationship was accompanied with decreasing heteroscedasticity. Models are summarized in Table S4.

Heteroscedasticity was mainly apparent in the CNPD dataset (Figure 2f) at the left end of the clade regularity index gradient (a range of values from 0.0 to 0.2, approximately). This was partly the reason of habitat dependency because the clade regularity index showed changeable fit accuracy across habitats (Table S5), and the heteroscedasticity issues at the left end were mainly caused by several habitats (Figure S5), such as C1 (surface standing waters) or C2 (surface running waters). Partly, the broader taxon sampling in the CNPD phylogeny was the reason for a large range of VPD values (approximately, three times higher than in species-rich grasslands). The variance of VPD values was largest at the left end, where the clade regularity index explained VPD less accurately (Figure 2f). Nevertheless, R^2 rapidly increased (72.3%) when we only included plots with the clade regularity index higher than 0.2 (93.8% of all plots). For phylogenetic richness and divergence, fit accuracy across habitats was usually similar (more than 70% for phylogenetic richness and more than 90% for phylogenetic divergence) with several exceptions with lower R^2 values, such as H2 (screes) or E4 (alpine and subalpine grasslands). Fit accuracy in all habitats is given in Table S5.

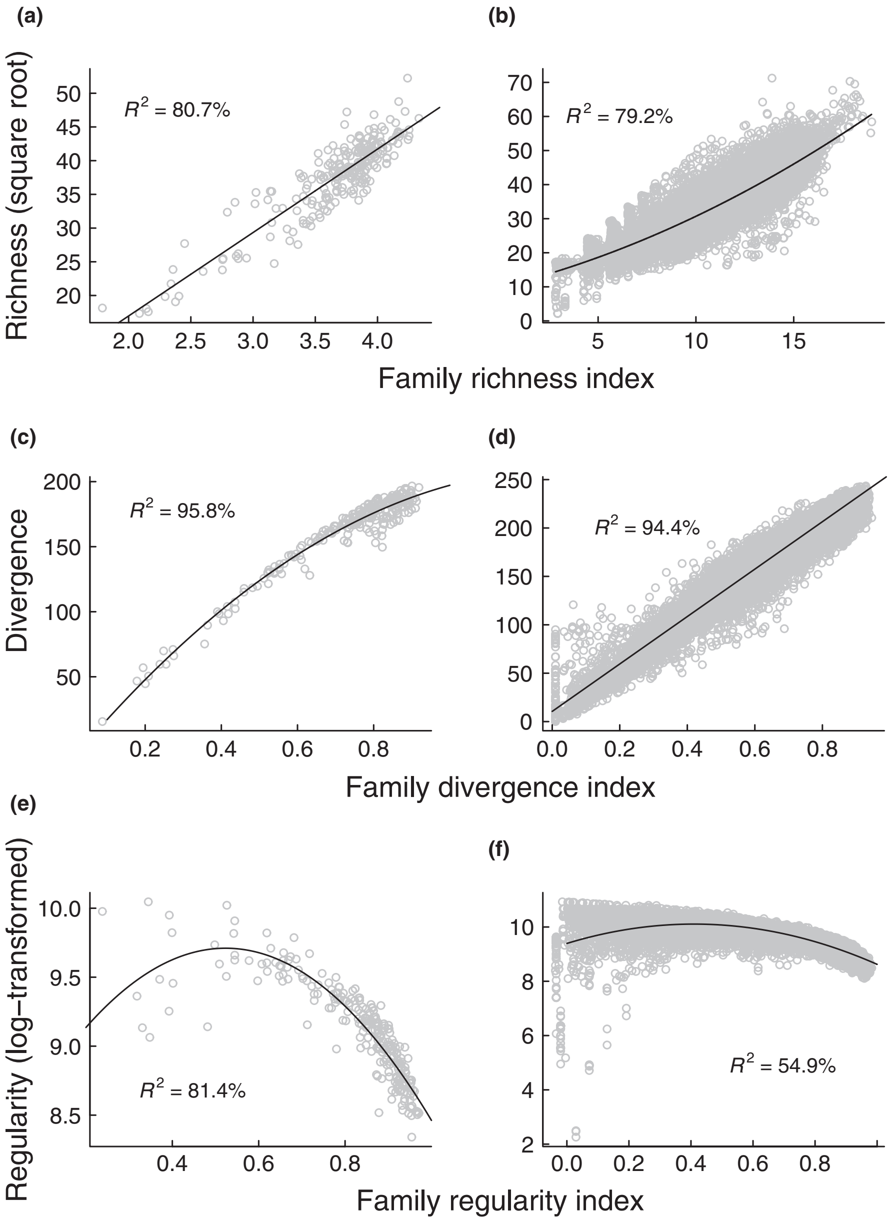
Simulated datasets revealed that species richness range was the most important determinant of fit accuracy of the clade richness index, while phylogenetic scale mainly affected fit accuracy of the clade divergence and regularity indices, followed by species richness (Table 3). Species pool size did not influence fit accuracy for any phylogenetic diversity dimension (Table 3). For phylogenetic richness and regularity, R^2 values increased with increasing species richness range (Figure 3a, Figure S6d,e). For phylogenetic divergence and regularity, fit accuracy increased with decreasing phylogenetic scale, R^2 was highest in community matrices sampled from the phylogeny of superasterids (Figure 3b,c), while the clade indices for these two dimensions were less reliable at the largest phylogenetic scale, that is, vascular plants (Figure 3b,c). At smaller phylogenetic scales (angiosperms and superasterids), fit accuracy for phylogenetic regularity also increased with increasing species richness range (Figure S6d,e), but this was not the case when we sampled community matrices using the whole phylogeny of vascular plants, that is, the largest phylogenetic scale considered (Figure S6f). Interestingly, the R^2 values for phylogenetic divergence were generally lower compared with the case studies where the family divergence index provided exceptional fit accuracy (95.8% and 94.4%), while the R^2 values very rarely exceeded 80% in simulated communities and the average was only 39%. In general, fit accuracy tended to be lower in simulated communities with low species richness, suggesting unreliability of the clade indices as surrogates of phylogenetic diversity in species-poor habitats or at very small spatial scales.

4 | DISCUSSION

We have shown that simple taxonomic coding at the family level can be used to accurately indicate phylogenetic diversity in plant communities. We propose three simple surrogates of phylogenetic diversity, the clade indices, which only require information about species affiliation to a clade and clade proportions in samples, while phylogenetic distances among species are not necessary (Table 2). Our indices provided an accurate fit to leading phylogenetic diversity metrics as shown for our two case studies: 1 × 1 m plots from species-rich grasslands and phytosociological relevés of various sizes from 26 Central European herbaceous habitats (Figure 2, Figures S3–S5). Our simulations indicate that the clade indices are highly correlated with phylogeny-based metrics in samples with 10 or more species (richness and regularity dimensions) and in species pools sampled from relatively smaller phylogenetic scales, such as angiosperms (estimated root age around 181 MYA, Kumar et al., 2017) or hierarchically lower clades (divergence and regularity dimensions; Figure 3, Figure S6). Using phylogenetic trees definitely provides the most detailed information about phylogenetic patterns in communities, however, due to the lack of resources (time or money) our proposed method can be used as a reliable proxy of phylogeny-based measures.

Clade indices can be used to simplify some aspects of the whole workflow behind phylogenetic diversity estimation. First, it enables the speeding-up of vegetation recording for any project involving a phylogenetic framework as species need to be correctly determined only at the higher taxonomic level (but note that the clade richness index requires species richness of a community for its computation). This is a welcome simplification, especially when dealing with hard to determine taxa. Second, phylogenetic diversity estimation using the clade indices requires less effort, expertise, and cost, as there is no need to obtain molecular data, performs a phylogenetic analysis

FIGURE 2 Fit accuracy of the clade indices in the case studies: species-rich grasslands (a,c,e) and the Czech National Phytosociological Database (b,d,f). (a,b) The phylogenetic richness dimension (described by Faith's PD). (c,d) The phylogenetic divergence dimension (described by MPD). (e,f) The phylogenetic regularity dimension (described by VPD). Number of plots: species-rich grasslands = 240, the Czech National Phytosociological Database = 16,542



and molecular dating (the latter is not necessary when phylograms are used, i.e., branch lengths in units substitutions per site; but see Jantzen et al., 2019 for discussion of how phylogenetic diversity measures can be affected by using either phylograms or dated phylogenies). As discussed in Li et al. (2019), researchers have to decide what markers (Which genes to select?) and methods to use (Alignment method? Model of evolution? Maximum likelihood or Bayesian inference framework? What fossil constraints for molecular dating?). All these difficult methodological decisions can be also avoided using phylogenies pruned from supertrees, for example, Daphne (Durka & Michalski, 2012) or the Open Tree of Life (Hinchliff et al., 2015), which have been shown to provide estimates of phylogenetic diversity well correlated to those derived from purpose-built phylogenies (Li et al., 2019) and, additionally, have broader taxon sampling coverage that is important to correctly estimate phylogenetic diversity (Jantzen et al., 2019; Park et al., 2018). On the other hand, for many taxonomic groups, supertrees are poorly sampled and unavailable (e.g., Daphne covers only a part of the European flora) or do not include branch lengths (Open Tree of Life) that need to be additionally calculated (Li et al., 2019).

The larger CNPD phylogeny with a broader taxonomic sampling created an almost three times larger range of VPD values in the CNPD compared to the grassland dataset. Due to this issue, we particularly encountered problems with heteroscedasticity. In species- and clade-poor habitats, the fit was generally poor (Table S5, Figure S5). For example, water habitats (C1 and C2) or carr and fen scrubs (F9.2) usually host specialized species from very few clades (e.g., Alismataceae or Salicaceae, respectively). Phylogenetic regularity of communities in these habitats will be highly dependent on the presence of other arms from the angiosperm radiation, as more distantly related lineages decrease phylogeny balance more than closely related ones, that is, the degree to which branch points define subgroups of equal size

(Heard, 1992). Vellend et al. (2011) provide relevant discussion of the effect of tree imbalance on phylogenetic diversity assessment. Thus, we suggest using the clade regularity index in relatively species-rich communities where its values are higher than 0.2, and recommend the estimation of phylogenetic regularity using phylogeny-based measures in communities where the clade regularity index ranges from 0 to 0.2. For phylogenetic richness and divergence, fit accuracy of the clade indices was consistent across all the studied habitats (Table S5, Figures 2, S3 and S4) and was, therefore, not affected by taxon sampling in the case studies.

Simulated community matrices highlighted the effect of species richness and phylogenetic scale on fit accuracy of the clade indices (Table 3). Species richness affects the values of phylogeny-based measures either directly or indirectly through shaping their range of possible values (Swenson, 2014; Vellend et al., 2011). In species-poor communities, the range of possible values of phylogeny-based measures was relatively high (Figure S7), and the clade indices (richness and regularity) tracked this variance less accurately (Figure 3a, Figure S6d,e). This suggests lower reliability of our method at very small spatial scales where plots consist of few species (<10). In contrast to species richness, increasing phylogenetic scale increases the possible range of phylogenetic distances because more distantly related species can occur in a community. As expected, fit accuracy for phylogenetic divergence and regularity was better at smaller phylogenetic scales (superasterids and angiosperms). For phylogenetic divergence, we observed a disparity in fit accuracy between case studies (substantial R^2 values) and simulated community matrices (moderate R^2 values). This could be probably attributed to the simulation protocol. Simulated community matrices were completely random in terms of species selection and species proportions, which does not reflect nonrandom assembly processes in nature. Sometimes, fit accuracy was greatly improved by log-transforming MPD values, but this mainly depended on the generated community matrix and we did not find consistent improvements after the log-transformation when comparing phylogenetic scales or species richness ranges. On the other hand, our case studies indicate that the phylogenetic divergence index is a very precise surrogate of MPD for real vegetation data (Figure 2c,d). In summary, the results suggest we should expect tight correlations between the clade indices and all dimensions of phylogenetic diversity in angiosperm-dominated habitats where samples have more than 10 species.

TABLE 3 Variance components of the hierarchically structured factors used for generating artificial communities

Factor	Richness	Divergence	Regularity
Phylogenetic scale	<0.1	62.1	51.5
Species pool size	<0.1	<0.1	<0.1
Species richness range	86.8	20.2	33.7
Residual	13.2	17.6	14.8

Note: Values (%) depict relative variance of R^2 values (fit accuracy of the proposed clade indices for all dimensions of phylogenetic diversity) attributed to a factor. Phylogenetic scale reflects a clade used for species pool generating (vascular plants, angiosperms, or superasterids). A megaphylogeny of vascular plants was taken from Qian and Jin (2016). Species pool size indicates the number of species in a regional phylogeny (2,000, 500, or 250). Species richness range indicates a range restricting the number of species in artificial communities (2–5, 5–10, 10–20, 10–40, 10–80, and 10–160). In total, 2,700 unique species pools and corresponding community matrices were generated.

Community and phylogenetic data influence the computation, behavior, or type I and II errors of phylogenetic diversity estimates (Cadotte et al., 2010; Miller, Farine, & Trisos, 2017; Tucker et al., 2017; Vellend et al., 2011). Certain features need to be considered when using clade proportions as an indicator of phylogenetic diversity. First, an outcome is dependent on the type of community data (presence/absence versus abundance weighted). Since the clade indices proposed here require information about relative abundances, they are not useful for presence/absence data. Second, phylogenetic diversity is expected to provide additional information than species

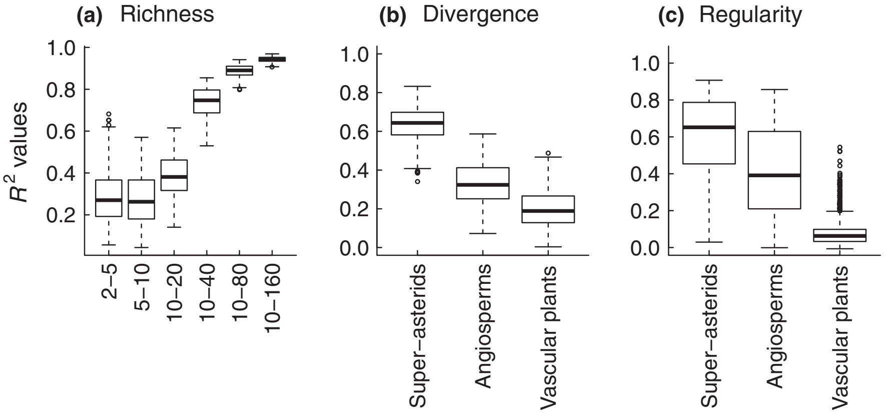


FIGURE 3 Major determinants of fit accuracy of the clade indices in simulated communities (species richness range for phylogenetic richness and phylogenetic scale for divergence and regularity; Table 3). (a) Phylogenetic richness: Faith's PD against family richness index in different species richness ranges, (b) phylogenetic divergence: MPD against family divergence index at different phylogenetic scales, and (c) phylogenetic regularity: VPD against family regularity index at different phylogenetic scales

richness and diversity. Usually, phylogenetic diversity metrics are positively correlated with species richness (Faith's PD) or at least the range of their possible values declines as the number of species increases (MPD; Swenson, 2014). As expected, clade indices showed the same decline of possible values with increasing species richness (Figure S7). To account for possible bias due to species richness variation, null models or rarefaction is recommended (Miller et al., 2017; Sandel, 2018; Swenson, 2014). Both tools can be used to treat species richness-dependence of clade indices. Finally, phylogenetic resolution influences the performance of the clade-based approach. As expected, our results indicate that increasing fineness of phylogenetic resolution increases the tightness of the relationship between phylogeny-based measures and clade indices (Table S3). This agrees with case studies and simulated phylogenies that showed a lower impact of the lack of resolution or poorly estimated branch lengths at more recent nodes on phylogenetic diversity (Allen et al., 2019; Swenson, 2009). Naturally, our method can be prone to taxonomic errors as it assumes proper species assignments to defined taxonomic groups.

Our goal was to show the link between clade composition and phylogenetic diversity. Our results suggest that the clade indices proposed here, which are based on taxonomic resolution at the family level, are a good indicator of all phylogenetic diversity dimensions in angiosperm-dominated habitats with 10 and more species per sampling unit (e.g., 1 m^2 or larger plots in grasslands). Even though this study focused on vascular plants, our results should generalize to any taxonomic group with a well-developed taxonomic classification supported by molecular data. In general, if a taxonomic classification of a group reflects current molecular phylogenies we should expect close correlations between taxonomy-based metrics (e.g., this study, Warwick & Clarke, 1995) and molecular-based phylogenetic metrics. Our approach has a potential in studies working with a lot of taxa when phylogenetic reconstruction might be very time- and money-consuming.

ACKNOWLEDGMENTS

We are grateful to Martin Kočí, Erika Lošáková, and David Opálka for help with field measurements in species-rich grasslands and Milan Chytrý for providing data from the Czech National Phytosociological Database. We thank Jan Mládek and Hanno Schaefer for comments on earlier drafts of the manuscript. This work was supported by Palacký University [IGA_PrF_2018_020 and IGA_PrF_2019_021]. Data collection in species-rich grasslands was supported by Ministry of the Environment of the Czech Republic [RPV SP/2D3/179/07], Grant Agency of the Czech University of Life Sciences [CIGA 42110/1313/3114], and Ministry of Agriculture of the Czech Republic [5400/2013-MZE-13212]. MB and MD were supported by Ministry of Education, Youth and Sports of the Czech Republic [INTER-EXCELLENCE, LTC18056, COST action 16212], and MB was also supported by the Scholarship of the city of Ostrava. RJP was supported by the Strategic Research Programme of the Scottish Government's Rural and Environment Science and Analytical Services Division.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

MB and PM conceived the ideas, designed the study, and analyzed the data. MB conducted phylogenetic analysis and wrote the manuscript with help from RJP and MD. All authors discussed the results, contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

All data supporting the results (accession numbers, alignment matrices, BEAST.xml file, phylogenetic trees, plot data, species lists, and simulation results) are archived in the Mendeley Data depository (https://data.mendeley.com/datasets/gbv472pxsb/1).

ORCID

Martin Bitomský https://orcid.org/0000-0002-0022-8473
Robin J. Pakeman https://orcid.org/0000-0001-6248-4133
Martin Duchoslav https://orcid.org/0000-0002-8553-8226

REFERENCES

Allen, J. M., Germain-Aubrey, C. C., Barve, N., Neubig, K. M., Majure, L. C., Laffan, S. W., … Soltis, P. S. (2019). Spatial phylogenetics of Florida vascular plants: The effects of calibration and uncertainty on diversity estimates. *iScience*, *11*, 57–70. https://doi.org/10.1016/j.isci.2018.12.002

Barker, G. M. (2002). Phylogenetic diversity: A quantitative framework for measurement of priority and achievement in biodiversity conservation. *Biological Journal of the Linnean Society*, *76*, 165–194. https://doi.org/10.1111/j.1095-8312.2002.tb02081.x

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2017). GenBank. *Nucleic Acids Research*, *41*, D37–D42. https://doi.org/10.1093/nar/gkw1070

Cadotte, M. W. (2015). Phylogenetic diversity–ecosystem function relationships are insensitive to phylogenetic edge lengths. *Functional Ecology*, *29*, 718–723. https://doi.org/10.1111/1365-2435.12429

Cadotte, M. W., Davies, T. J., Regetz, J., Kembel, S. W., Cleland, E., & Oakley, T. H. (2010). Phylogenetic diversity metrics for ecological communities: Integrating species richness, abundance and evolutionary history. *Ecology Letters*, *13*, 96–105. https://doi.org/10.1111/j.1461-0248.2009.01405.x

Campbell, N., Neat, F., Burns, F., & Kunzlik, P. (2010). Species richness, taxonomic diversity, and taxonomic distinctness of the deep-water demersal fish community on the Northeast Atlantic continental slope (ICES Subdivision Via). *ICES Journal of Marine Science*, *68*, 365–376. https://doi.org/10.1093/icesjms/fsq070

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, *17*, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

Chytrý, M., Pyšek, P., Tichý, L., Knollová, I., & Danihelka, J. (2005). Invasions by alien plants in the Czech Republic: A quantitative assessment across habitats. *Preslia*, *77*, 339–354.

Chytrý, M., & Rafajová, M. (2003). Czech National Phytosociological Database: Basic statistics of the available vegetation-plot. *Preslia*, *75*, 1–15.

Clarke, K. R., & Warwick, R. M. (2001). A further biodiversity index applicable to species lists: Variation in taxonomic distinctness. *Marine*

Ecology Progress Series, *216*, 265–278. https://doi.org/10.3354/meps216265

Dobeš, C., Rossa, J., Paule, J., & Hülber, K. (2013). Morphology, DNA-molecular variation, karyology, ecogeography, and phytosociology suggest allopatric differentiation and species rank for *Potentilla rigoana* (Rosaceae). *Taxon*, *62*, 733–745. https://doi.org/10.12705/624.8

Drori, M., Rice, A., Einhorn, M., Chay, O., Glick, L., & Mayrose, I. (2018). OneTwoTree: An online tool for phylogeny reconstruction. *Molecular Ecology Resources*, *18*, 1492–1499. https://doi.org/10.1111/1755-0998.12927

Durka, W., & Michalski, S. G. (2012). Daphne: A dated phylogeny of a large European flora for phylogenetically informed ecological analyses. *Ecology*, *93*, 2297. https://doi.org/10.1890/12-0743.1

Elliott, T. L., Waterway, M. J., & Davies, T. J. (2016). Contrasting lineage-specific patterns conceal community phylogenetic structure in larger clades. *Journal of Vegetation Science*, *27*, 69–79. https://doi.org/10.1111/jvs.12345

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*, 1–10.

Gerhold, P., Carlucci, M. B., Proches, S., & Prinzing, A. (2018). The deep past controls the phylogenetic structure of present, local communities. *Annual Review of Ecology, Evolution, and Systematics*, *49*, 477–497. https://doi.org/10.1146/annurev-ecolsys-110617-062348

Graham, C. H., Storch, D., & Machac, A. (2018). Phylogenetic scale in ecology and evolution. *Global Ecology and Biogeography*, *27*, 175–187. https://doi.org/10.1111/geb.12686

Hall, S. J., & Greenstreet, S. P. (1998). Taxonomic distinctness and diversity measures: Responses in marine fish communities. *Marine Ecology Progress Series*, *166*, 227–229.

Heard, S. B. (1992). Patterns in tree balance among cladistic, phenetic, and randomly generated phylogenetic trees. *Evolution*, *46*, 1818–1826. https://doi.org/10.1111/j.1558-5646.1992.tb01171.x

Hinchliff, C. E., Smith, S. A., Allman, J. F., Burleigh, J. G., Chaudhary, R., Coghill, L. M., … Cranston, K. A. (2015). Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences of the United States of America*, *13*, 12764–12769. https://doi.org/10.1073/pnas.1423041112

Iv, A. P. G. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, *181*, 1–20. https://doi.org/10.1111/boj.12385

Jantzen, J. R., Whitten, W. M., Neubig, K. M., Majure, L. C., Soltis, D. E., & Soltis, P. S. (2019). Effects of taxon sampling and tree reconstruction methods on phylodiversity metrics. *Ecology and Evolution*, *9*, 9479–9499. https://doi.org/10.1002/ece3.5425

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., … Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, *26*, 1463–1464. https://doi.org/10.1093/bioinformatics/btq166

Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, *34*, 1812–1819. https://doi.org/10.1093/molbev/msx116

Li, D., Trotta, L., Marx, H. E., Allen, J. M., Sun, M., Soltis, D. E., … Baiser, B. (2019). For common community phylogenetic analyses, go ahead and use synthesis phylogenies. *Ecology*, *100*, e02788. https://doi.org/10.1002/ecy.2788

Miller, E. T., Farine, D. R., & Trisos, C. H. (2017). Phylogenetic community structure metrics and null models: A review with new methods and software. *Ecography*, *40*, 461–477. https://doi.org/10.1111/ecog.02070

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. *Proceedings of the gateway computing environments workshop (GCE)* (pp. 1–8).

Ndiribe, C., Pellissier, L., Antonelli, S., Dubuis, A., Pottier, J., Vittoz, P., … Salamin, N. (2013). Phylogenetic plant community structure along elevation is lineage specific. *Ecology and Evolution*, *3*, 4925–4939.

Park, D. S., Worthington, S., & Xi, Z. (2018). Taxon sampling effects on the quantification and comparison of community phylogenetic diversity. *Molecular Ecology*, *27*, 1296–1308. https://doi.org/10.1111/mec.14520

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D., & R Core Team (2019). *nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-139*. Retrieved from https://CRAN.R-project.org/package=nlme

Qian, H., & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology*, *9*, 233–239. https://doi.org/10.1093/jpe/rtv047

R Core Team (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rambaut, A., Drummond, A. J., Xie, D., Baele, D., & Suchard, M. A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*, 901–904. https://doi.org/10.1093/sysbio/syy032

Redding, D. W., & Mooers, A. O. (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology*, *20*, 1670–1678. https://doi.org/10.1111/j.1523-1739.2006.00555.x

Sandel, B. (2018). Richness-dependence of phylogenetic diversity indices. *Ecography*, *41*, 837–844. https://doi.org/10.1111/ecog.02967

Scheiner, S. M., Kosman, E., Presley, S. J., & Willig, M. R. (2017). The components of biodiversity, with a particular focus on phylogenetic information. *Ecology and Evolution*, *7*, 6444–6454. https://doi.org/10.1002/ece3.3199

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, *4*, vey016. https://doi.org/10.1093/ve/vey016

Swenson, N. G. (2009). Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PLoS ONE*, *4*, e4390. https://doi.org/10.1371/journal.pone.0004390

Swenson, N. G. (2014). *Functional and phylogenetic ecology in R*. New York, NY: Springer.

Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., … Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, *92*, 698–715. https://doi.org/10.1111/brv.12252

Vellend, M., Cornwell, W. K., Magnuson-Ford, K., & Mooers, A. O. (2011). Measuring phylogenetic diversity. In A. E. Magurran, & B. McGill (Eds.), *Biological diversity: Frontiers in measurement and assessment* (pp. 193–206). Oxford, UK: Oxford University Press.

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. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Reviews of Ecology, Evolution, and Systematics*, *33*, 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

Yguel, B., Jactel, H., Pearse, I. S., Moen, D., Winter, M., Hortal, J., … Prinzing, A. (2016). The evolutionary legacy of diversification predicts ecosystem function. *The American Naturalist*, *188*, 398–410. https://doi.org/10.1086/687964

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., … Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, *506*, 89–92. https://doi.org/10.1038/nature12872

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Bitomský M, Mládková P, Pakeman RJ, Duchoslav M. Clade composition of a plant community indicates its phylogenetic diversity. *Ecol Evol*. 2020;00:1–11. https://doi.org/10.1002/ece3.6170

Variation in mycorrhizal status is much larger within than among sites: a phylogenetically-informed assessment across grassland and coastal floras

Journal:	*Oikos*
Manuscript ID	Draft
Wiley - Manuscript type:	Research
Keywords:	arbuscular mycorrhiza, community mycorrhizal dispersion, community mycorrhization, phylogenetic diversity, plant-fungal interactions, resource-acquisition strategies
Abstract:	Plant community mycorrhization is a measure of the functional composition of mycorrhizal status in vegetation. It is based on a community weighted mean (CWM) approach which, however, is controversial because it ignores within-site trait variation, might be sensitive to phylogenetic autocorrelation and can have unacceptable significance rates. Here, we propose a concept of community mycorrhizal dispersion, i.e. a rate of within-site differences in mycorrhizal status, and assess its ecological relevance. Using data from Central European grasslands, Scottish coastal habitats and two approaches to assign plant mycorrhizal status to species lists (empirical vs. taxonomic), we decomposed the variation in mycorrhizal status (obligatory, facultative and non-mycorrhizal plants) to within- and among-site components. Additionally, we examined how community mycorrhization and mycorrhizal dispersion are linked to phylogenetic diversity. Within-site variation in mycorrhizal status was always at least 2.2 times (but mostly nine times) larger than among site variation regardless of quadrat size or mycorrhizal status assignment approach. Community mycorrhization and mycorrhizal dispersion were usually tightly and non-linearly related to phylogenetic diversity in both systems. The tightness of this relationship could be explained by strong phylogenetic conservatism of mycorrhizal status, while the non-linearity is probably the result of stabilising selection that favours mycorrhizal symbiosis over mutualism abandonment. However, mycorrhizal status assignment approaches produced different results in some situations. We encourage future studies to also focus on within-site variability in mycorrhizal status and consider phylogenetic effects, because both aspects appear to be highly relevant in herbaceous systems.

SCHOLARONE™
Manuscripts

ORIGINAL RESEARCH

Variation in mycorrhizal status is much larger within than among sites: a phylogenetically-informed assessment across grassland and coastal floras

DATA ACCESSIBILITY

Should the manuscript be accepted, all data supporting the results (vegetation records, mycorrhizal status data, phylogenetic tree, quadrat data) will be archived in the Mendeley Data depository ((https://data.mendeley.com) and the data DOI will be included in the paper.

ABSTRACT

Plant community mycorrhization is a measure of the functional composition of mycorrhizal status in vegetation. It is based on a community weighted mean (CWM) approach which, however, is controversial because it ignores within-site trait variation, might be sensitive to phylogenetic autocorrelation and can have unacceptable significance rates. Here, we propose a concept of community mycorrhizal dispersion, i.e. a rate of within-site differences in mycorrhizal status, and assess its ecological relevance. Using data from Central European grasslands, Scottish coastal habitats and two approaches to assign plant mycorrhizal status to species lists (empirical vs. taxonomic), we decomposed the variation in mycorrhizal status (obligatory, facultative and non-mycorrhizal plants) to within- and among-site components. Additionally, we examined how community mycorrhization and mycorrhizal dispersion are linked to phylogenetic diversity. Within-site variation in mycorrhizal status was always at least 2.2 times (but mostly nine times) larger than among site variation regardless of quadrat size or mycorrhizal status assignment approach. Community mycorrhization and mycorrhizal dispersion were usually tightly and non-linearly related to phylogenetic diversity in both systems. The tightness of this relationship could be explained by strong phylogenetic conservatism of mycorrhizal status, while the non-linearity is probably the result of stabilising selection that favours mycorrhizal symbiosis over mutualism abandonment. However, mycorrhizal status assignment approaches produced different results in some situations. We encourage future studies to also focus on within-site variability in mycorrhizal status and consider phylogenetic effects, because both aspects appear to be highly relevant in herbaceous systems.

Keywords: arbuscular mycorrhiza, community mycorrhizal dispersion, community mycorrhization, phylogenetic diversity, plant-fungal interactions, resource-acquisition strategies

INTRODUCTION

Plant community mycorrhization is a relatively novel trait-based concept useful for hypothesis testing related to the prevalence of mycorrhizal symbiosis in plant communities (Moora 2014). Plant species can be distinguished according to several mycorrhizal traits; but from a plant-centric view, mycorrhizal status is often considered as the most important. Mycorrhizal status denotes whether the roots of a plant species are consistently (obligatory mycorrhiza, OM), sometimes, but not always (facultative mycorrhiza, FM) or never (non-mycorrhiza, NM) colonised by mycorrhizal fungi in nature (Smith & Read 2008, Hempel et al. 2013, Moora 2014). So far, changes of community mycorrhization or distributions of OM, FM and NM plants have been linked to plant diversity (Gerz et al. 2016), soil conditions and age (Read 1991, Peat & Fitter 1993, Lambers et al. 2008, Craine 2009, Hempel et al. 2013, Gerz et al. 2016), habitat type (Peat & Fitter 1993, Brundrett & Tedersoo 2018), land use (Gerz et al. 2019), climate (Menzel et al. 2016), elevation or latitude (Bueno et al. 2017). Quantitative community studies addressing the prevalence of mycorrhizal symbioses are rapidly emerging, however, there is no study, which examines potential analytical shortcomings.

Community mycorrhization is an estimate of functional composition based on the CWM approach (Lavorel et al. 2008), i.e. community means of mycorrhizal status weighted by plant species abundances (Moora 2014). However, the CWM approach has been recently questioned as it can produce inconsistent parameter estimation and inflated Type I error rates (Peres-Neto et al. 2017), but note that this also depends on underlying assumptions about the link of species or sample attributes to species composition (i.e. the hypothesis in question; Zelený 2018). In other words, some studies might have provided too optimistic statistical support for the relationship between community mycorrhization and environmental gradients when, in fact, there is no relationship. More importantly, current research has so far focused solely on differences of community mycorrhization among sites, but there is little evidence regarding the drivers of within-site variation of mycorrhizal traits. This component of trait variation could be very important as there are systems, such as phosphorus-impoverished old soils in Australia (Lambers et al. 2008) where both mycorrhizal and NM plants can coexist via facilitation of P acquisition by mycorrhizal plants by neighbouring non-mycorrhizal, cluster-rooted plants (Lambers et al. 2018), or tundra assemblages with coexisting OM plants (*Betula* and *Vaccinium*) and NM sedges (Craine 2009). In general, within-site trait variability can be relatively higher than the trait variability among sites (Gallagher & Leishman 2012, Pescador et al. 2015, Peres-Neto et al. 2017). With respect to nutrient-acquisition strategies, Zemunik et al. (2015) found high functional diversity especially in nutrient-poor soils. Thus, even if there is a significant relationship between mycorrhizal status and the environment, in fact, it can only describe a portion of the total trait variability as it ignores the within-site component.

Phylogenetic autocorrelation is an important aspect of community-level analyses that should be considered prior to all trait-environment relationship testing (de Bello et al. 2017, Duarte et al. 2018). This is especially relevant for mycorrhizal traits, which are usually strongly phylogenetically structured. For example, orchid and ericoid mycorrhizal types are unique for most species in the Orchidaceae and Ericaceae, respectively (Brundrett & Tedersoo 2019). Mycorrhizal status tends to be usually conserved at the family level, but there are some exceptions (see Brundrett & Tedersoo 2019, Bueno et al. 2019a,b). This phylogenetic dependency might inflate type I errors in the mycorrhizal status-environment association when species distributions also show a phylogenetic signal even though mycorrhizal status may not be ecologically related to the studied environmental gradient (Duarte et al. 2018). Thus, analyses addressing patterns of community mycorrhization should be phylogenetically informed because the factors of interest might actually drive changes of phylogenetic structure and changes in community mycorrhization might be just a side effect. In addition, decoupling trait and phylogenetic information can uncover some hidden signals underlying species coexistence and turnover (de Bello et al. 2017). Nevertheless, the trait dependency on phylogeny is still challenging to interpret because it is dependent on spatial scale, number of traits measured, mode and rate of trait evolution or variation in species richness and evenness (Pavoine et al. 2013, Tucker et al. 2018), and the evidence on correlations between functional (FD) and phylogenetic diversity (PD), i.e. the PD-FD relationship, is mixed (Cadotte et al. 2019).

Given the described inherent limitations of the CWM approach when studying mycorrhiza at the plant community level, we aim to draw attention to community mycorrhizal dispersion, i.e. a concept that attempts to indicate a rate of within-site dissimilarity in mycorrhizal status (Table 1). To demonstrate this approach, we used two geographically, environmentally and ecologically different herbaceous systems: semi-natural grasslands (Central Europe) and coastal habitats (dunes and machair, a type of calcareous coastal grassland, in Scotland). The overall mycorrhization in grasslands is expected to be relatively high as these systems are usually dominated by OM plants (Hartnett & Wilson 2002). Additionally, OM plants are more frequent in Central Europe (Bueno et al. 2017). On the other hand, OM plants are less frequent in Scottish coastal habitats, possibly due to higher latitude, disturbance and moistness (Peat & Fitter 1993, Hempel et al. 2013, Bueno et al. 2017).

First, we evaluated the ecological relevance of the within-site variability in mycorrhizal status at various spatial scales in both systems. To do so, we computed within- and among-site variability in mycorrhizal status using the partitioning framework proposed by Peres-Neto et al. (2017). Second, we examined how community mycorrhization and mycorrhizal dispersion are related to phylogenetic diversity in both systems. For community mycorrhizal dispersion, the nature of the PD-FD relationship depends on the mode and rate of trait evolution (Tucker et al. 2018), thus, we also estimated transition rates among OM, FM and NM states to examine potential evolutionary persistence of mycorrhizal status and system-dependent differences in mycorrhizal status evolution in both regional species pools. Natural selection is more likely to favour the ability to form mycorrhiza rather than mutualism abandonment because evolutionary shifts (measured as transition rates) towards OM or FM are more frequent in seed plants (Maherali et al. 2016). Due to this stabilising selection, we expected a non-linear saturating curve (similar to logarithmic increase) to fit the PD-FD relationship (as in the case of the continuous trait divergence constrained by stabilising selection, Tucker et al. 2018).

MATERIALS AND METHODS

Vegetation recording

For the semi-natural grasslands, we collected data during two vegetation surveys in 12 experimental sites (Beskydy Mountains, N 49.45°, E 18.33°; and White Carpathian Mountains, N 48.97°, E 17.82°) in the Czech Republic. Sites are characterised as mesic meadows and pastures, and broad-leaved semi-dry and submontane grasslands (alliances: *Arrhenatherion elatioris, Cynosurion cristati, Bromion erecti and Violion caninae*). During the first survey (summer 2013), we recorded 240 permanent quadrats. We specifically designed data collection to address spatial scale effects. To do so, we used three quadrat sizes (a nested-plot design): 0.33 m x 0.33 m, 0.67 m x 0.67 m and 1 m x 1m. Quadrats were under four management treatments (abandoned, spring grazing, grazing + spring burning, mowing in mid-July). During the second survey (summer 2018), we recorded 110 mown quadrats (1 m x 1 m) in 11 experimental sites. In both surveys, we visually estimated species' percentage cover. The second dataset consisted of vegetation records from a 2009-2011 resurvey of coastal systems around Scotland, UK. Data collection is fully explained in Pakeman et al. (2015). The quadrat size was 5 m x 5 m, the number of plots was 1462 from these habitat classes (following Pakeman et al. 2015; number of quadrats in parentheses): acid grassland (64), cliff (35), fixed dune (619), heath (86), improved grassland (26), mobile dune (80), salt marsh (43), scrub (26), semi-fixed dune (35), slack (117), strand (13), tall grass mire (45), unimproved grassland (53), wet grassland (61), wet heath (23) and 136 quadrats were not assigned to this classification. In total, both datasets consisted of 457 species, 171 for semi-natural grasslands and 386 species for coastal habitats. Both datasets shared 100 species.

Phylogenetic analysis

To get DNA sequence data, we used an online tool OneTwoTree (Drori et al. 2018), which yielded a set of 17 markers (data DOI). The sequences were aligned in MAFFT (Katoh & Standley 2013) and the alignment was cured to eliminate poorly aligned and divergent regions using GBlocks (Castresana 2000). Then, we performed a maximum likelihood (ML) analysis using the RaxML-HPC2 program (Stamatakis 2014) on the CIPRES portal (Miller et al. 2010). We selected a moss *Physcomitrella patens* as an outgroup, set the GTRGAMMA model and provided a backbone tree created at the www.timetree.org website (Kumar et al. 2017) prior to the analysis. The tree with the best likelihood score was then selected as the optimal tree.

To obtain the dated phylogenetic tree, we translated genetic distances into absolute times in BEAST v1.10.4 (Suchard et al. 2018) using the best ML tree as a starting tree. We set an uncorrelated relaxed clock as the clock model, Yule process as a speciation model and GTR+G+I (with 4 gamma categories) as a nucleotide substitution model. To translate genetic distances into absolute times, we exploited the TimeTree database (Kumar et al. 2017) and set several time priors with normally distributed errors (median and standard deviation computed from all studies available in the TimeTree database reporting a given divergence time estimate). Then we performed three independent runs (with different starting seeds) for 150 million generations each. Then we checked convergence in Tracer v1.7.1 (Rambaut et al. 2018) and combined all runs (10% of generations used as a burn-in). The dated maximum clade credibility tree was sampled from 30 000 trees (10% trees as a burn-in). Originally, the above described phylogenetic analysis was for the purposes of a different analysis that was based on 506 species, thus we pruned this larger phylogeny to get the regional phylogeny of Scottish coastal habitats (386 species). We took advantage of this broader taxon sampling in order to mitigate the effect of under sampling that can lead to an underestimation of divergence times and phylogenetic diversity (Park et al 2018). The phylogeny of 171 species occurring in the grassland dataset was taken from Anonymous (2020), which followed very similar phylogenetic analysis protocol but was based on 20 DNA markers, hence, we prioritised this tree. Both phylogenies with depicted NM and FM families are shown in Figs. S1, S2.

Mycorrhizal status data

There is an ongoing discussion about certain concepts underlying mycorrhizal traits (Brundrett & Tedersoo 2019; Bueno et al. 2019a,b). Brundrett & Tedersoo (2019) recently pointed out that host species lists contain errors, such as reported mycorrhiza in families with well-documented NM status (e.g. Brassicaceae, Caryophyllaceae or Juncaceae). Mycorrhizal status is usually consistent at the family level (taxonomic approach, Table 1), but plants from NM families can become colonised by mycorrhizal fungi under some conditions (e.g. in Brassicaceae, Cosme et al. 2018). Some authors argue that these inconsistencies are the result of technical misdiagnosis and empirical data should be checked against the standard reference based on family mycorrhizal status (Brundrett & Tedersoo 2018, 2019). On the other hand, the taxonomic approach assumes strong phylogenetic conservatism and may conflict with empirical data or neglect relevant evolutionary and ecological processes at the species level (Cosme et al. 2018, Bueno et al. 2019a,b). These conceptual differences can be summarised to two main methods: the approach based on empirical information and the approach based on taxonomy (Bueno et al. 2019b). Both approaches (hereafter referred to as the empirical or taxonomic approach, respectively) have some pros and cons and assume that either all accumulated empirical data are correct or that mycorrhizal status is completely conserved at the chosen taxonomic level (discussed in detail in Bueno et al. 2019b). Unfortunately, both approaches can provide very different mycorrhizal status estimates, especially in the case of FM (91% of species mismatched in a European dataset, Bueno et al. 2019b). Since there is clearly no agreement on what approach is likely more precise, we believe it is better to evaluate results using both approaches to avoid biased conclusions.

We grouped plant species into the three categories: OM, FM and NM plants. This classification has proved to be useful in various plant-centric ecological or biogeographical studies of mycorrhiza (e.g. Hempel et al. 2013, Menzel et al. 2016, Bueno et al. 2017). In grasslands, further classification according to mycorrhizal type is not particularly helpful, because arbuscular mycorrhiza (AM) dominates here. Additionally, AM is considered as the only mycorrhizal type where FM can appear, while all other mycorrhizal types, such as orchid and ericoid mycorrhiza, are considered as OM plants (Moora 2014). The most dominant grasses are usually OM plants (Hartnett & Wilson 2002). Both empirical and taxonomic mycorrhizal status assignments for all species were taken from Bueno et al. (2019b). For several missing species we obtained the empirical mycorrhizal statuses from the MycoFlor database (Hempel et al. 2013) using the *TR8* package (Bocci 2015), while the taxonomic mycorrhizal statuses were taken from Brundrett and Tedersoo (2019). In total, the empirical approach indicated 198 OM, 216 FM and 43 NM plants, while the taxonomic approach indicated 336 OM, 41 FM and 80 NM plants in our species list (data DOI).

The mismatch rate between both assignments was 48%. As expected, almost all of the 220 mismatched species were linked to facultative mycorrhiza (FM-OM and FM-NM mismatches). Because of the substantial presence of mismatched species, we tested whether the mismatches (described as a binary variable: 0/1) are phylogenetically correlated, which would indicate the presence of some problematic phylogenetic groups where both assignment approaches provide different results. To do so, we estimated phylogenetic signal as the alpha parameter of the phylogenetic logistic regression (Ives & Garland 2010) using the *phyloglm* function (the *phylolm* package, Ho & Ané 2014) in both regional species pools. In grasslands, the alpha estimate reached its upper bound, suggesting negligence of phylogenetic correlation (Ives & Garland 2010) and, therefore, random occurrence of mismatches across the tree. In coastal habitats, we detected low phylogenetic signal (alpha = 0.12). For example, the mismatch rates were over 70% in some more abundant families, such as Brassicaceae, Campanulaceae, Chenopodiaceae, Juncaceae or Polygonaceae (Table S1).

Data processing

We did all analyses in R version 3.6.0. (R Core Team 2019). We followed the usual numerical coding of mycorrhizal status: NM plants = 0, FM plants = 0.5 and OM plants = 1 (Moora 2014, Gerz et al. 2016, 2019). To some extent, this numerical coding could be an oversimplification as the trait distances between NM-FM and FM-OM plants are assumed to be equal (0.5), while the real functional (and genetic) distance between FM and OM plants is probably smaller because both strategies possess traits (and genes) required for a mutualistic association. Some studies focusing on the species level response use categorical coding (e.g. Hempel et al. 2013). Nevertheless, the numerical coding is convenient for the quantification of functional dispersion and for the community-level focus of this study.

To examine the relevance of within- and among-sites variation in mycorrhizal status in all studied systems, we used the *TraitEnvCor* function provided by Peres-Neto et al. (2017), which enabled us to obtain percentage values of both components. We defined within-site variability as the trait variability within the surveyed quadrats and computed percentage variances using community and trait matrices for all quadrat sizes and occurring species and both mycorrhizal status assignment approaches. We used both abundance (relative cover) and presence/absence community matrices.

To examine how community mycorrhization and community mycorrhizal dispersion are related to phylogenetic diversity, we calculated empirically based and taxonomically based community mycorrhization index (CWM) for each plot, following Moora (2014). To estimate community mycorrhizal dispersion (empirically and taxonomically based) and phylogenetic diversity, we computed abundance-weighted mean functional (MFD) and mean phylogenetic (MPD) distances (Weiher et al. 1998, Webb et al. 2002, Swenson 2014) using the *mpd* function (the *picante* package, Kembel et al. 2010). Both CWM and MFD scale from 0 to 1. As MFD and MPD average across all species pairs, they both describe basal functional and phylogenetic components that do not reflect distances among the closest relatives, i.e. the terminal diversity (Swenson 2014). Nevertheless, the basal component is appropriate for our data because shifts in mycorrhizal status tend to appear at least at the family level (Brundrett & Tedersoo 2019). For this analysis, we combined data on 1 m x 1 m quadrats from both grassland surveys (350 plots in total) and 5 m x 5 m quadrats from the survey of coastal habitats (1462 plots in total).

Visual data inspection revealed non-linearities, thus, we fitted generalised additive models (GAMs) using the *mgcv* package (Wood 2017) with restricted maximum likelihood (REML) and with applied weighting on each quadrat in order to account for uncertainty of the quantification of community mycorrhization and mycorrhizal dispersion caused by mismatched species (empirical vs. taxonomic approach). Weights were simply defined as 1 – proportion of mismatched species in a quadrat. Models were checked visually and by the *gam.check* function, and were fitted separately for each mycorrhizal assignment approach and dataset (four models in total). Additionally, we compared community mycorrhization index and MFD values based on the empirical and taxonomic approach using paired t-tests.

We also examined the evolutionary flexibility with which plants from the given regional species pool can switch between different mycorrhizal statuses. Considering alternative models of trait evolution can generate testable predictions and new insights into the interactions between ecology and evolution, such as the nature of the PD-FD relationship (Cadotte et al. 2018, Tucker et al. 2018). Mycorrhizal status is a discrete trait so we considered a macroevolutionary ARD (“all rates different”) Mk model (Lewis 2001, Harmon 2019) as the transition rates among mycorrhizal statuses tend to be highly asymmetric at various phylogenetic scales (Maherali et al. 2016). To reduce the number of parameters that need to be estimated, we a priori set the NM-OM transition rate to zero because this transition is very rare and unlikely due to the loss of genes responsible for the plant-mycorrhiza symbiosis in NM plants (Delaux et al. 2014, Maherali et al. 2016). Therefore, we estimated five out of six transition rates (instantaneous rates of change among all possible states; Pennell et al. 2014, Harmon 2019) possible: FM-OM, OM-FM, NM-FM, FM-NM and OM-NM. The switches are in descending order according to their commonness in seed plants (Maherali et al. 2016). We also evaluated the adequacy of the Pagel's transformations of branch lengths (lambda, delta and kappa; Pagel 1999, Harmon 2019) by an AIC comparison. We used the *fitDiscrete* function in *geiger* (Pennell et al. 2014) and set 500 iterations. For the lambda model, we set 1100 iterations because a half of the iterations tended to fail in several independent runs. The transition rates were compared between the empirical and taxonomic approach and between both regional species pools.

RESULTS

Within-site variation in mycorrhizal status was at least 2.2 times higher than among sites (Table 2) and mostly at least nine times higher for the taxonomic approach (Table 2b). For the empirical approach, within-site variation tended to increase with increasing quadrat size in grasslands; while for the taxonomic approach, the highest within-site variation was in quadrats of the size of 0.67 m x 0.67 m. When considering the presence/absence data, the within-site variability was always higher than 94% regardless of quadrat size, dataset or mycorrhizal status assignment approach (Table 2).

Community mycorrhization and mycorrhizal dispersion were related to phylogenetic diversity non-linearly in both systems, except for the taxonomic approach in semi-natural grasslands where we found a very weak linear relationship (Table S2). For community mycorrhization (CWM) and MPD, we observed a unimodal-like relationship (Fig. 1a). At both ends of the CWM gradients, MPD tended to be lower, while quadrats with higher MPD tended to retain a narrow range of CWM values (approximately 0.65–0.85). Phylogenetically diverse quadrats were unlikely to have either extremely low or high community mycorrhization. For community mycorrhizal dispersion (MFD), we found a positive non-linear relationship with MPD, similar to a logarithmic increase (Fig. 1b). In particular, MPD rapidly increased when MFD was low (approximately 0.0–0.2), while it tended to be relatively constant when MFD was high.

Importantly, the mycorrhizal indices were highly affected by the input mycorrhizal status data. On average and across all quadrat sizes, the taxonomic approach produced 1.45 times (by 0.28 units) higher values of community mycorrhization than the empirical approach, while the empirical approach produced 4.90 times (by 0.26 units) higher values of community mycorrhizal dispersion than the taxonomic approach (Table S3). For grasslands, the taxonomic approach produced a very narrow range of CWM and MFD values. This dataset was not suitable for modelling and the overall explanatory power was very low (Table S2).

For the empirical approach, lambda transformed Mk models had the best fit in both grasslands ($\lambda = 0.76$) and coastal habitats ($\lambda = 0.88$; Table S4). For the taxonomic approach, non-transformed models with constant rate (equivalent to $\lambda = 1.00$) performed best in both regional species pools (Table S4). Both mycorrhizal status assignment approaches suggested very different transition rates within regional species pools (Fig. 2a,c). However, the NM-FM, FM-OM and OM-FM transition rates tended to be, on average, higher than the rest of mycorrhizal status changes, consistently in both regional phylogenies and for both mycorrhizal status assignment approaches. Even though the estimated transition rates among mycorrhizal statuses were generally lower in coastal habitats (Fig. 2a,c), probably due to higher tree height, it did not affect the nature of the relationships between CWM, MFD and MPD as those relationships exhibited similar non-linear trends (Fig. 1). Both species pools had similar proportions of OM and FM plants, while NM plants were approximately 1.85 times more frequent in coastal habitats (Fig. 2b,d).

DISCUSSION

We have shown that the within-site variability in mycorrhizal status is ecologically more relevant that the variability among sites in grassland and coastal habitats (and most likely in other herbaceous systems), regardless of quadrat size or mycorrhizal assignment approach. The trait variation decomposition (Peres-Neto et al. 2017) consistently indicated that the CWM approach, which exclusively focuses on the among-sites trait variation, only describes at least a 2.2 times (but mostly at least a nine times) smaller portion of the total variation in mycorrhizal status (Table 2), while the majority of mycorrhizal status variation lay within sampled quadrats, i.e. sites (Table 2). At first glance, this finding seems to be surprising, because the majority of studies has attributed variability of mycorrhizal symbiosis to environmental factors that play a role at larger spatial scales. For example, the ability to form mycorrhiza is considered as an adaptation to low nutrient supply (Read 1991, Craine 2009) and is expected to be associated with habitat or nutritional specialization, e.g. FM plants are disturbance opportunists, halophytes or aquatic plants, while NM plants are epiphytes or parasites (Brundrett & Tedersoo 2018). However, there is a clear bias in the ecological literature towards studying changes of community mycorrhization along various ecological gradients (Lambers et al. 2008, Gerz et al. 2016, 2019) or distribution patterns of mycorrhizal status categories (Hempel et al. 2013, Menzel et al. 2016, Bueno et al. 2017). These approaches only focus on the among-sites component and ignore the within-site trait variation, which can be higher as has been shown for various plant traits (e.g. Gallagher & Leishman 2012, Pescador et al. 2015, Peres-Neto et al. 2017). This perspective opens new research possibilities, such as testing the importance of environmental factors and biotic interactions (competition or facilitation) in shaping within-site variability in mycorrhizal status, or resource-acquisition strategies in general (see Zemunik et al. 2015, Lambers et al. 2018).

Based on these findings, we encourage the consideration of the concept of community mycorrhizal dispersion in ecological studies. Community mycorrhizal dispersion indicates a within-site rate of mycorrhiza-based resource-acquisition strategies and can be estimated via MFD (this study) or other various functional dispersion measures (Zemunik et al. 2015). This concept also has some analytical advantages (besides focusing on key trait variability component) that are rarely used for the CWM approach. In our data, we mostly observed decreasing heteroscedasticity (the scale parameter computed by the *gamlss* function, Rigby & Stasinopoulos 2005) of CWM and MFD with increasing species richness (Table S5, Fig. S3). In species-rich quadrats, only a small range of CWM or MFD values is possible, suggesting that null model corrected versions of CWM and MFD might be more appropriate to treat for the species richness dependence (Swenson 2014). Further, trait distances can be decoupled from phylogeny (de Bello et al. 2017). Decoupling can give more importance to trait differentiation within a set of closely related species (de Bello et al. 2017), this was the case in grasslands where standardised effect sizes (SES) of decoupled MFD tended to produce higher values than SES of MFD (Fig. S4). For some analytical purposes, evolutionary legacy in non-decoupled trait distances might be considered as a confounding factor and given that community mycorrhization and mycorrhizal dispersion seem to be (non-linearly) linked to phylogenetic diversity (Fig. 1), a decoupling framework might prove to be useful.

Both measures of within- and among-site variation of mycorrhizal status were non-linearly linked to phylogenetic diversity. For community mycorrhization, we observed a plateau of maximal MPD values within a narrow interval of CWM (Fig. 1a), while both ends of the CWM gradient tended to have lower phylogenetic diversity. This suggests possible bias in phylogenetically diverse habitats and at very large quadrat sizes, where the among-site variability of community mycorrhization can be relatively low and narrowly distributed around some convergence value. This could be the reason for lower among-site variability in our datasets. For community mycorrhizal dispersion, we found a rapid non-linear increase of MPD when MFD values were relatively low (Fig. 1b), but the effect decreased (and sometimes fluctuated) when MFD was high. In general, this result is in concordance with most studies testing the links between phylogenetic and functional diversity, as positive relationships are more frequent in nature (Cadotte et al. 2019). Tucker et al. (2018) examined how a rate of increase of functional richness per unit of phylogenetic richness (measured as Faith's PD; Faith 1992) is dependent on tempo and rate of trait evolution, and showed that positive saturating curves fit the PD-FD relationship for simulated continuous traits according to the Orstein-Uhlenbeck (OU) model. As expected, our models in both datasets (except for the taxonomic approach in grasslands) suggest similar curves. Stabilising selection, which favours mycorrhizal symbiosis over mutualism abandonment as suggested by estimated transition rates in this study (Fig. 2a,c) and in Maherali et al. (2016), could be the reason of the observed non-linear PD-FD relationship.

The largest source of variation of tested patterns in this study were clearly differences between mycorrhizal status assignment approaches (Table S3). The empirical approach probably overestimated the number of FM plants (Table S1). It seems to be unlikely that 47% of all plants have FM status because such high proportion of FM is not usually reported in seed plants (Maherali et al. 2016, Brundrett & Tedersoo 2018, 2019). The FM strategy is usually associated with habitat specialisation (Brundrett & Tedersoo 2018) but our dataset consisted of mostly common herbaceous plants. The taxonomic approach assigned OM status to 74% of all plants, which is close to the recent estimate of global proportion of vascular plants forming the arbuscular mycorrhizal symbiosis (72%, Brundrett & Tedersoo 2018). Unfortunately, there is no agreement on which host species lists are more precise, and the whole topic is currently under hot debate (Brundrett & Tedersoo 2018, 2019, Bueno et al. 2019a,b), needless to say that the mismatch rate between both approaches is very disturbing (Table S1). In addition, both approaches suggest quite different evolution of mycorrhizal status in our regional species pools in terms of transition rates and the evolutionary model (Fig. 2, Table S4). Fortunately, both assignment approaches support the key findings of this study that the within-site variation of mycorrhizal status matters more than the among-sites component, as well as both approaches mostly suggest similar curves (except for the taxonomic approach in grasslands) fitting the relationships between community mycorrhization, mycorrhizal dispersion and phylogenetic diversity.

REFERENCES

Anonymous (2020). Details omitted for double-blind reviewing.

de Bello F., Šmilauer P., Diniz-Filho J.A.F., Carmona C.P., Lososová Z., Herben T., & Götzenberger L. (2017). Decoupling phylogenetic and functional diversity to reveal hidden signals in community assembly. Methods in Ecology and Evolution 8, 1200–1211. doi: 10.1111/2041-210X.12735

Bocci G. (2015). TR8: an R package for easily retrieving plant species traits. Methods in Ecology and Evolution 6, 347–350. doi: 10.1111/2041-210X.12327

Brundrett M.C., & Tedersoo L. (2018). Evolutionary history of mycorrhizal symbioses and global host plant diversity. New Phytologist 220, 1108–1115. doi: 10.1111/nph.14976

Brundrett M.C., & Tedersoo L. (2019). Misdiagnosis of mycorrhizas and inappropriate recycling of data can lead to false conclusions. New Phytologist 221, doi: 10.1111/nph.15440

Bueno C.G., Aldrich-Wolfe L., Chaudhary V.B., Gerz M., Helgason T., Hoeksema J.D., … Moora M. (2019a). Misdiagnosis and uncritical use of plant mycorrhizal data are not the only elephants in the room. New Phytologist 224, 1415–1418. doi: 10.1111/nph.15976

Bueno C.G., Gerz M., Zobel M., & Moora M. (2019b). Conceptual differences lead to divergent trait estimates in empirical and taxonomic approaches to plant mycorrhizal trait assignment. Mycorrhiza 29, 1–11. doi: 10.1007/s00572-018-0869-1

Bueno C.G., Moora M., Gerz M., Davison J., Öpik M., Pärtel M., … Zobel M. (2017). Plant mycorrhizal status, but not type, shifts with latitude and elevation in Europe. Global Ecology and Biogeography 26, 690–699. doi: 10.1111/geb.12582

Castresana J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334

Cadotte M.W., Campbell S.E., Li S., Sodhi D.S., & Mandrak N.E. (2018). Preadaptation and naturalization of nonnative species: Darwin's two fundamental insights into species invasion. Annual Review of Plant Biology 69, 661–684. doi: 10.1146/annurev-arplant-042817-040339

Cadotte M.W., Carboni M., Si X., & Tatsumi S. (2019). Do traits and phylogeny support congruent community diversity patterns and assembly inferences? Journal of Ecology 107, 2065–2077. doi: 10.1111/1365-2745.13247

Cosme M., Fernández I., Van der Heijden M.G.A., & Pieterse C.M.J. (2018). Non-mycorrhizal plants: the exceptions that prove the rule. Trends in Plant Science 23, 577–587. doi: 10.1016/j.tplants.2018.04.004

Craine J.M. (2009). Resource strategies of wild plants. New Jersey, USA: Princeton University Press.

Delaux P.M., Varala K., Edger P.P., Coruzzi G.M., Pires J.C., & Ané J.M. (2014). Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. PLoS Genetics 10, e1004487. doi: 10.1371/journal.pgen.1004487

Drori M., Rice A., Einhorn M., Chay O., Glick L., & Mayrose I. (2018). OneTwoTree: an online tool for phylogeny reconstruction. Molecular Ecology Resources 18, 1492–1499. doi: 10.1111/1755-0998.12927

Duarte L.D.S., Debastiani V.J., Carlucci M.B., & Diniz-Filho J.A.F. (2018). Analyzing community-weighted trait means across environmental gradients: should phylogeny stay or should it go? Ecology 99, 385–398. doi: 10.1002/ecy.2081

Faith D.P. (1992). Conservation evaluation and phylogenetic diversity. Biological Conservation 61, 1–10.

Gallagher R.V., & Leishman M.R. (2012). Contrasting patterns of trait-based community assembly in lianas and trees from temperate Australia. Oikos 121, 2026–2035. doi: 10.1111/j.1600-0706.2012.20025.x

Gerz M., Bueno C.G., Ozinga W.A., Zobel M., & Moora M. (2019). Responses of plant community mycorrhization to anthropogenic influence depend on the habitat and mycorrhizal type. Oikos 128, 1565–1575. doi: 10.1111/oik.06272

Gerz M., Bueno C.G., Zobel M., & Moora M. (2016). Plant community mycorrhization in temperate forests and grasslands: relations with edaphic properties and plant diversity. Journal of Vegetation Science 27, 89–99. doi: 10.1111/jvs.12338

Harmon L.J. (2019). Phylogenetic comparative methods: learning from trees. Self published under a CC-BY-4.0 license. https://lukejharmon.github.io/pcm/

Hartnett D.C., & Wilson G.W.T. (2002). The role of mycorrhizas in plant community structure and dynamics: lessons from grasslands. Plant and Soil 244, 319–331. doi: 10.1023/A:1020287726382

Hempel S, Götzenberger L., Kühn I., Michalski S.G., Rillig M.C., Zobel M., & Moora M. (2013). Mycorrhizas in the Central European flora: relationships with plant life history traits and ecology. Ecology 94, 1389–1399.

Ho L.S.T., & Ané C. (2014). A linear-time algorithm for gaussian and non-gaussian trait evolution models. Systematic Biology 63, 397–408. doi: 10.1093/sysbio/syu005

Ives A.R., & Garland T. (2010). Phylogenetic logistic regression for binary dependent variables. Systematic Biology 59, 9–26. doi: 10.1093/sysbio/syp074

Katoh K., & Standley D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30, 772–780. doi: 10.1093/molbev/mst010

Kembel S.W., Cowan P.D., Helmus M.R., Cornwell W.K., Morlon H., Ackerly D.D., ... Webb C.O. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

Kumar S., Stecher G., Suleski M., & Hedges S.B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. Molecular Biology and Evolution 34, 1812–1819. doi: 10.1093/molbev/msx116

Lambers H., Albornoz F., Kotula L., Laliberté E., Ranathunge K., Teste F.P., & Zemunik G. (2018). How belowground interactions contribute to the coexistence of mycorrhizal and non-mycorrhizal species in severely phosphorus-impoverished hyperdiverse ecosystems. Plant and Soil 424, 11–33. doi: 10.1007/s11104-017-3427-2

Lambers H., Raven J.A., Shaver G.R., & Smith S.E. (2008). Plant nutrient-acquisition strategies change with soil age. Trends in Ecology & Evolution 23, 95–103. doi: 10.1016/j.tree.2007.10.008

Lavorel S., Grigulis K., McIntyre S., Williams N.S.G., Garden D., Dorrough J., … Bonis A. (2008). Assessing functional diversity in the field – methodology matters! Functional Ecology 22, 134–147. doi: 10.1111/j.1365-2435.2007.01339.x

Lewis P.O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. Systematic Biology 50, 913–925. doi: 10.1080/106351501753462876

Maherali H., Oberle B., Stevens P.F., Cornwell W.K., & McGlinn D.J. (2016). Mutualism persistence and abandonment during the evolution of the mycorrhizal symbiosis. The American Naturalist 188, E113-E125. doi: 10.1086/688675

Menzel A., Hempel S., Manceur A.M., Götzenberger L., Moora M., Rillig M.C., Zobel M., & Kühn I. (2016). Distribution patterns of arbuscular mycorrhizal and non-mycorrhizal plant species in Germany. Perspectives in Plant Ecology, Evolution and Systematics 21, 78–88. doi: 10.1016/j.ppees.2016.06.002

Miller M.A., Pfeiffer W., & Schwartz T. (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE): 1–8.

Moora M. (2014). Mycorrhizal traits and plant communities: perspectives for integration. Journal of Vegetation Science 25, 1126–1132. doi: 10.1111/jvs.12177

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

Pakeman R.J., Alexander J., Beaton J., Brooker R., Cummins R., Eastwood A., … Lewis R. (2015). Species composition of coastal dune vegetation in Scotland has proved resistant to climate change over a third of a century. Global Change Biology 21, 3738–3747. doi: 10.1111/gcb.12999

Park D.S., Worthington S., & Xi Z. (2018). Taxon sampling effects on the quantification and comparison of community phylogenetic diversity. Molecular Ecology 27, 1296–1308. doi: 10.1111/mec.14520

Pavoine S., Gasc A., Bonsall M.B., & Mason N.W.H. (2013). Correlations between phylogenetic and functional diversity: mathematical artefacts or true ecological and evolutionary processes? Journal of Vegetation Science 24, 781–793. doi: 10.1111/jvs.12051

Peat H.J., & Fitter A.H. (1993). The distribution of arbuscular mycorrhizas in the British flora. New Phytologist 125, 845–854. doi: 10.1111/j.1469-8137.1993.tb03933.x

Pennell M.W., Eastman J.M., Slater G.J., Brown J.W., Uyeda J.C., FitzJohn R.G., Alfaro M.E., & Harmon L.J. (2014). geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. Bioinformatics 30, 2216–2218. doi: 10.1093/bioinformatics/btu181

Peres-Neto P.R., Dray S., & ter Braak J.F. (2017). Linking trait variation to the environment: critical issues with community-weighted mean correlation resolved by the fourth-corner approach. Ecography 40, 806–816. doi: 10.1111/ecog.02302

Pescador D.S., de Bello F., Valladares F., & Escudero A. (2015). Plant trait variation along an altitudinal gradient in Mediterranean high mountain grasslands: controlling the species turnover effect. PLOS ONE 10, e0118876. doi: 10.1371/journal.pone.0118876

Rambaut A., Drummond A.J., Xie D., Baele D., & Suchard M.A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67, 901–904. doi: 10.1093/sysbio/syy032

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Read D.J. (1991). Mycorrhizas in ecosystems. Experientia 47, 376–391.

Rigby R.A., & Stasinopoulos D.M. (2005). Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society: Applied Statistics Series C 54, 507–554. doi: 10.1111/j.1467-9876.2005.00510.x

Smith S.E., & Read D.J. (2008). Mycorrhizal symbiosis. London, UK: Academic Press.

Stamatakis A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatic*s* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Suchard M.A., Lemey P., Baele G., Ayres D.L., Drummond A.J., & Rambaut A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evolution 4, vey016. doi: 10.1093/ve/vey016

Swenson N.G. (2014). Functional and phylogenetic ecology in R. New York, USA: Springer.

Tucker C.M., Davies T.J., Cadotte M.W., Pearse W.D. (2018). On the relationship between phylogenetic diversity and trait diversity. Ecology 99, 1473–1479. doi: 10.1002/ecy.2349

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. doi: 10.1146/annurev.ecolsys.33.010802.150448

Weiher E., Clarke G.D.P., & Keddy P.A. (1998). Community assembly rules, morphological dispersion, and the coexistence of plant species. Oikos 81, 309–322. doi: 10.2307/3547051

Wood S.N. (2017). Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC.

Zelený D. (2018). Which results of the standard test for community-weighted mean approach are too optimistic? Journal of Vegetation Science 29, 953–966. doi: 10.1111/jvs.12688

Zemunik G., Turner B.L., Lambers H., & Laliberté E. (2015). Diversity of plant nutrient-acquisition strategies increases during long-term ecosystem development. Nature Plants 1, 15050. doi: 10.1038/NPLANTS.2015.50

Table 1 Overview of the terminology used in this study.

Term	Explanation	References
a) data related		
Empirical mycorrhizal status assignment approach (referred here as the empirical approach)	Uses available empirical information on mycorrhizal status	Bueno et al. (2017, 2019b)
Taxonomic mycorrhizal status assignment approach (referred here as the taxonomic approach)	Mycorrhizal status based on affiliation to a family	Brundrett & Tedersoo (2019) Bueno et al. (2019b)
b) analysis related		
Community mycorrhization	Community mean of mycorrhizal status weighted by plant species relative abundances: a measure of functional composition of a community. Mycorrhizal status coded as 0 for non-mycorrhizal (NM), 0.5 for facultative mycorrhizal (FM) and 1 for obligatory mycorrhizal (OM) plants.	Moora (2014)
Community mycorrhizal dispersion	Community dissimilarity (e.g. mean distance) of mycorrhizal status: a measure of functional dispersion of a community. Same coding as for community mycorrhization. This concept reflects the within-site variation in mycorrhizal status and can be used to test various hypotheses related to co-occurrence or coexistence of mycorrhiza-based resource acquisition strategies.	This study Zemunik et al. (2015)
MFD, MPD	Mean functional distance and mean phylogenetic distances: measures of trait dispersion and phylogenetic divergence, respectively.	Weiher et al. (1998) Webb et al. (2002) Swenson (2014)

Table 2 Relative importance (%) of the within- and among-sites components of the mycorrhizal status variability, based either on a) the empirical or b) taxonomic approach, across quadrats of different sizes (0.33 m x 0.33 m, 0.67 m x 0.67 m, 1 m x 1 m and 5 m x 5 m). We computed the variance components using abundance and presence/absence community matrices from Czech semi-natural grasslands and Scottish coastal habitats.

Scale	Percentage cover		Presence/absence	
	Within (%)	Among (%)	Within (%)	Among (%)
a) Empirical mycorrhizal status				
0.33 m x 0.33 m (grasslands 1st survey)	80.15	19.85	96.07	3.93
0.67 m x 0.67 m (grasslands 1st survey)	84.07	15.93	97.58	2.42
1 m x 1 m (grasslands 1st survey)	86.50	13.50	98.35	1.65
1 m x 1 m (grasslands, 2nd survey)	88.51	11.49	97.03	2.97
5 m x 5 m (coastal habitats)	69.20	30.80	96.00	4.00
b) Taxonomic mycorrhizal status				
0.33 m x 0.33 m (grasslands 1st survey)	90.83	9.17	94.81	5.19
0.67 m x 0.67 m (grasslands 1st survey)	95.58	4.42	97.23	2.77
1 m x 1 m (grasslands 1st survey)	93.58	6.42	98.14	1.86
1 m x 1 m (grasslands, 2nd survey)	84.69	15.31	96.46	3.54
5 m x 5 m (coastal habitats)	79.47	20.53	95.21	4.79

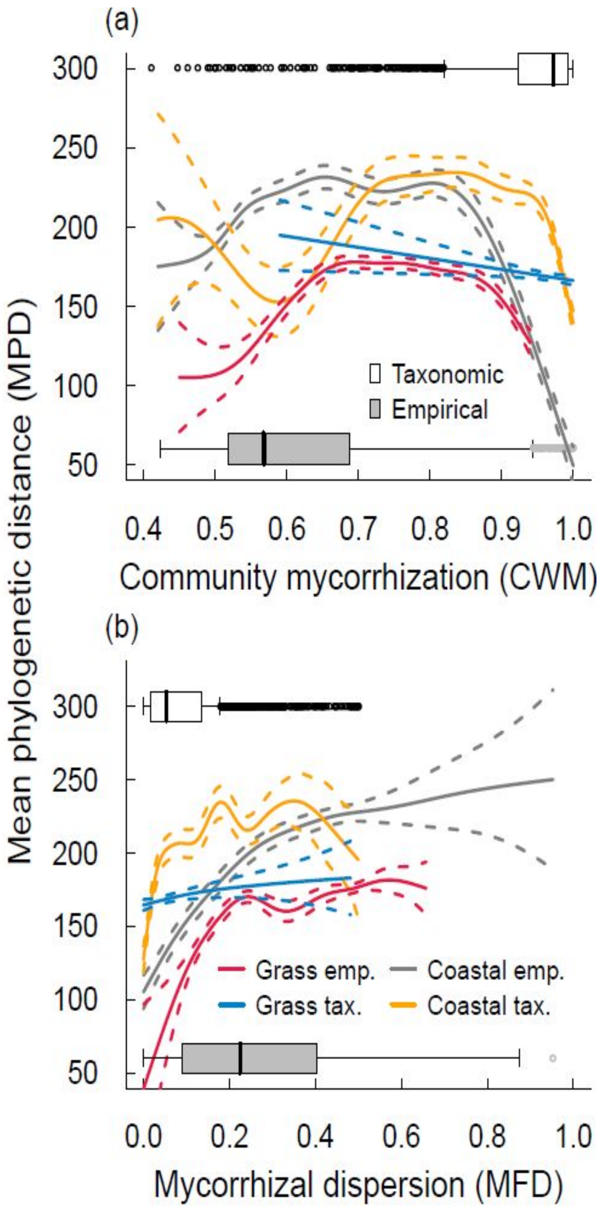


Fig. 1 Generalised additive models fitting the relationships between phylogenetic diversity (MPD) and a) plant community mycorrhization (CWM of mycorrhizal status) and b) community mycorrhizal dispersion (MFD of mycorrhizal status). The CWM and MFD values were estimated using both empirical and taxonomic approaches to mycorrhizal status assignment. The boxplots indicate the distributions of the CWM and MFD values based either on the empirical (grey boxplots) or taxonomic (white boxplots) approach. Models were fitted for each dataset (Czech grasslands and Scottish coastal habitats) and each mycorrhizal status assignment approach.

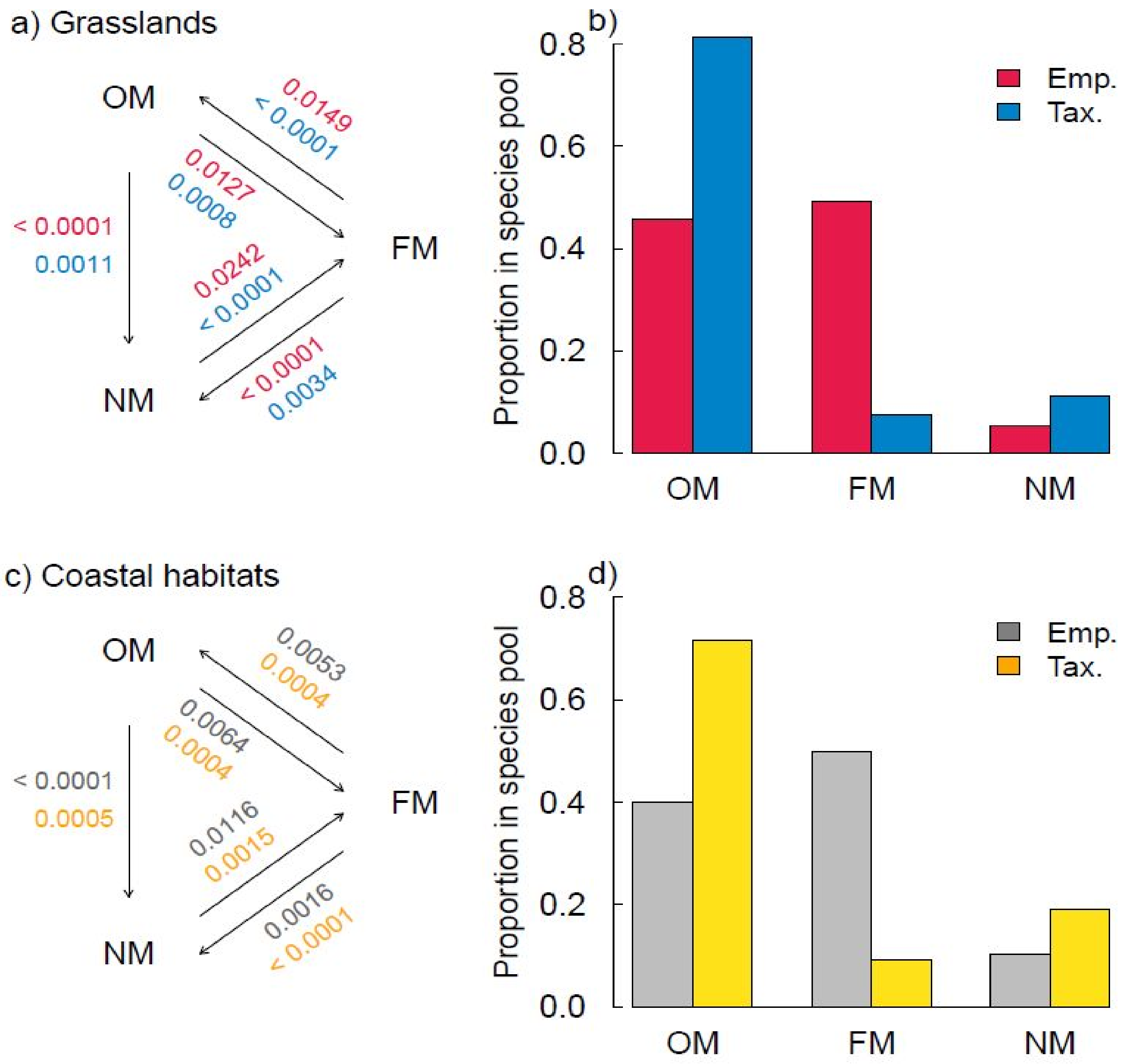


Fig. 2 Estimated transition rates between mycorrhizal statuses in a) grasslands and c) coastal habitats. b, d) Proportions of OM, FM and NM plants in both regional species pools. Colour coding follows Fig. 1 and indicates results based either on the empirical or taxonomic approach of mycorrhizal status assignment.

SUPPORTING INFORMATION

Table S1 Comparison of the empirical and taxonomic mycorrhizal status assignment approaches at the family level.

Table S2 Summary of models fitting the relationships between a) phylogenetic diversity and community mycorrhization, and b) phylogenetic diversity and community mycorrhizal dispersion at all spatial scales.

Table S3 Results of paired t-tests comparing estimates of a) community mycorrhization (CWM) and b) community mycorrhizal dispersion (MFD) based on different mycorrhizal status assignment approaches at all quadrat sizes.

Table S4 Sample-size corrected AIC comparison of several macroevolutionary models fitted to both regional phylogenies (grassland and coastal habitats).

Table S5 Summary of models testing the heteroscedasticity of CWM and MFD values along the species richness gradient in both datasets (1 m x 1 m quadrats in semi-natural grasslands and 5 m x 5 m quadrats in coastal habitats).

Fig. S1 Phylogeny of the species pool of Czech semi-natural grasslands with depicted FM and NM families.

Fig. S2 Phylogeny of the species pool of Scottish coastal habitats with depicted FM and NM families.

Fig. S3 Community mycorrhization (CWM) and community mycorrhizal dispersion (MFD) plotted against species richness in both datasets.

Fig. S4 Mean differences (+/- 95% CI) between SES values of MFD based on trait distances decoupled from phylogeny (dcMFD) and raw trait distances (MFD).

Table S1 Comparison of the empirical and taxonomic mycorrhizal assignment approach at the family level. Family species richness and mycorrhizal statuses are indicated. Species from both grasslands and coastal habitats were combined (457 unique species in total).

Family (APG IV)	n	Taxonomic	Empirical			Mismatch rate (%)
			NM	FM	OM	
Adoxaceae	1	OM		1		100
Alismataceae	2	FM		1	1	50
Amaryllidaceae	2	OM		1	1	50
Apiaceae	17	OM	1	7	9	47
Araceae	1	OM	1			100
Araliaceae	2	OM		1	1	50
Asteraceae	42	OM		14	28	33
Athyriaceae	1	OM		1		100
Balsaminaceae	1	OM		1		100
Betulaceae	4	OM		1	3	25
Blechnaceae	1	OM		1		100
Boraginaceae	6	OM		3	3	50
Brassicaceae	14	NM	4	10		71
Campanulaceae	7	OM		5	2	71
Caprifoliaceae	6	OM		1	5	17
Caryophyllaceae	22	NM	9	12	1	59
Celastraceae	1	OM		1		100
Chenopodiaceae	8	NM	2	4	2	75
Cistaceae	1	OM			1	0
Colchicaceae	1	OM			1	0
Convolvulaceae	1	NM		1		100
Crassulaceae	2	FM		2		0
Cupressaceae	1	OM			1	0
Cyperaceae	30	FM	13	15	2	53
Dennstaedtiaceae	1	OM		1		100
Droseraceae	1	NM		1		100
Dryopteridaceae	3	OM		2	1	67
Elaeagnaceae	1	OM		1		100
Equisetaceae	4	OM	1	3		100
Ericaceae	6	OM			6	0
Euphorbiaceae	4	OM		3	1	75
Fabaceae	31	OM		7	24	23
Fagaceae	2	OM			2	0
Gentianaceae	3	OM			3	0

Geraniaceae	6	OM		4	2	67
Hyacinthaceae	1	OM			1	0
Hydrophyllaceae	1	FM			1	100
Hypericaceae	4	OM	1	1	2	50
Iridaceae	1	OM		1		100
Isoetaceae	1	OM		1		100
Juncaceae	14	NM	3	11		79
Lamiaceae	18	OM		8	10	44
Lentibulariaceae	1	NM	1			0
Linaceae	1	OM		1		100
Lycopodiaceae	1	OM		1		100
Menyanthaceae	1	FM		1		0
Montiaceae	1	OM	1			100
Myricaceae	1	NM		1		100
Nartheciaceae	1	OM		1		100
Oleaceae	1	OM			1	0
Onagraceae	5	OM		1	4	20
Ophioglossaceae	2	OM			2	0
Orchidaceae	9	OM			9	0
Orobanchaceae	6	NM	4	2		33
Oxalidaceae	1	OM		1		100
Papaveraceae	1	FM		1		0
Pinaceae	4	OM			4	0
Plantaginaceae	13	OM		7	6	54
Plumbaginaceae	1	FM		1		0
Poaceae	47	OM		29	18	62
Polygalaceae	3	OM			3	0
Polygonaceae	11	NM	1	10		91
Polypodiaceae	1	OM		1		100
Potamogetonaceae	1	FM	1			100
Primulaceae	9	OM		4	5	44
Ranunculaceae	10	OM		4	6	40
Rosaceae	20	OM		7	13	35
Rubiaceae	10	OM		5	5	50
Salicaceae	5	OM		1	4	20
Sapindaceae	1	OM		1		100
Saxifragaceae	2	FM		2		0
Selaginellaceae	1	OM			1	0
Solanaceae	3	OM		2	1	67
Urticaceae	1	FM		1		0
Violaceae	6	OM		4	2	67

Table S2 Summary of GAMs fitting the relationships between a) phylogenetic diversity (MPD) and community mycorrhization (CWM), and b) MPD and community mycorrhizal dispersion (MFD) in both datasets, including both the empirical and taxonomic approach (see Fig. 1). Approximate significance of the smooth terms (effective degrees of freedom, F and P-values) and adjusted R^2 values are given. For the MPD ~ CWM_{tax} relationship, the term was forced to be linear (edf = 1).

	edf	F-value	P-value	R^2 (%)
a) Grasslands				
MPD ~ CWM_{emp}	6.8	36.1	< 0.001	44.5
MPD ~ CWM_{tax}	1.0	-2.4 (t-value)	0.018	1.4
MPD ~ MFD_{emp}	6.5	25.4	< 0.001	35.5
MPD ~ MFD_{tax}	1.6	4.4	0.014	2.2
b) Coastal habitats				
MPD ~ CWM_{emp}	7.1	140.8	< 0.001	44.1
MPD ~ CWM_{tax}	10.7	34.5	< 0.001	22.3
MPD ~ MFD_{emp}	4.6	90.8	< 0.001	26.7
MPD ~ MFD_{tax}	8.5	38.3	< 0.001	20.6

Table S3 Results of paired t-tests comparing estimates of a) community mycorrhization (CWM) and b) community mycorrhizal dispersion (MFD) based on different mycorrhizal status assignment approaches at all quadrat sizes. Estimates indicate the average difference (and 95% confidence interval) of the taxonomic approach from the empirical approach.

Quadrat size	Dataset	Estimate	t-value	df	P-value
a) CWM					
0.33 m x 0.33 m	Grasslands	0.30 (0.282, 0.312)	39.3	239	< 0.001
0.67 m x 0.67 m	Grasslands	0.29 (0.274, 0.301)	42.7	239	< 0.001
1.00 m x 1.00 m	Grasslands (1st + 2nd survey)	0.26 (0.246, 0.269)	44.4	349	< 0.001
5.00 m x 5.00 m	Coastal habitats	0.33 (0.326, 0.342)	81.3	1461	< 0.001
b) MFD					
0.33 m x 0.33 m	Grasslands	-0.33 (-0.347, -0.310)	-34.9	239	< 0.001
0.67 m x 0.67 m	Grasslands	-0.15 (-0.161, -0.142)	-31.6	239	< 0.001
1.00 m x 1.00 m	Grasslands (1st + 2nd survey)	-0.30 (-0.312, -0.279)	-35.7	349	< 0.001
5.00 m x 5.00 m	Coastal habitats	-0.12 (-0.129, -0.110)	-25.4	1461	< 0.001

Table S4 Sample-size corrected AIC comparison of several macroevolutionary models fitted to both regional phylogenies (grassland and coastal habitats). White is a non-phylogenetic model (a star phylogeny); lambda, kappa and delta are Pagel-style transformation parameters that attempt to estimate eventual rate differences across the tree (Pagel 1999, Pennell et al. 2014, Harmon 2019). We also fitted a model of rate constancy through time. Models with the lowest AICc are in bold.

Branch length transformation model	AICc (empirical)	AICc (taxonomic)
a) Grasslands		
White	298.9	212.2
Lambda	**277.7**	74.4
Kappa	291.6	74.4
Delta	289.6	73.1
None (rate constancy)	294.8	**72.3**
b) Coastal habitats		
White	736.6	604.5
Lambda	**663.8**	160.0
Kappa	676.3	160.0
Delta	713.7	159.9
None (rate constancy)	715.7	**157.9**

Table S5 Summary of models testing the heteroscedasticity of CWM and MFD values along the species richness gradient in both datasets (1 m x 1 m quadrats in semi-natural grasslands and 5 m x 5 m quadrats in coastal habitats). Negative estimates indicate the negative relationship between log-transformed sigma (the scale parameter, Rigby & Stasinopoulos 2005) and species richness, i.e. decreasing overall heteroscedasticity with increasing species richness. We used the *gamlss* function to estimate the slope of the scale paramater (Rigby & Stasinopoulos 2005).

Dataset/parameter	Estimate	SE	t-value	P-value
a) Grasslands				
$\text{CWM}_{\text{empirical}}$	-0.03	0.005	-5.9	< 0.001
$\text{CWM}_{\text{taxonomic}}$	-	-	-	n.s.
$\text{MFD}_{\text{empirical}}$	-0.02	0.005	-5.5	< 0.001
$\text{MFD}_{\text{taxonomic}}$	-	-	-	n.s.
b) Coastal habitats				
$\text{CWM}_{\text{empirical}}$	-0.04	0.002	-17.3	< 0.001
$\text{CWM}_{\text{taxonomic}}$	-0.02	0.002	-9.7	< 0.001
$\text{MFD}_{\text{empirical}}$	-0.01	0.002	-3.4	< 0.001
$\text{MFD}_{\text{taxonomic}}$	-0.01	0.002	-6.1	< 0.001

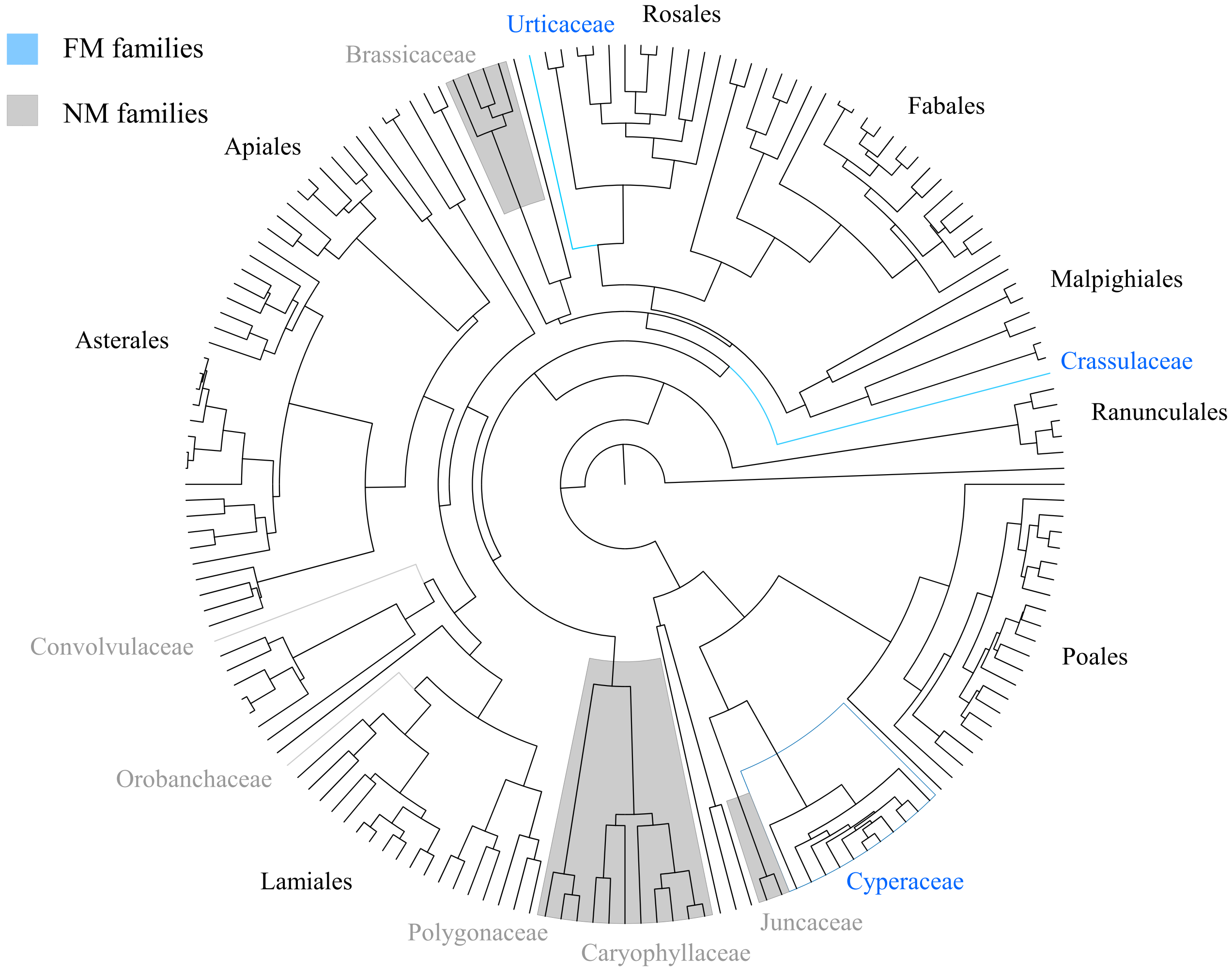


Fig. S1 Phylogeny of the species pool (171 species) of Czech semi-natural grasslands with depicted FM and NM families (following Brundrett & Tedersoo 2019) and several angiosperm orders (to allow orientation). We assigned FM status to *Sedum sexangulare* (Crassulaceae) as this mycorrhizal status is supported by the empirical approach (Bueno et al. 2019b). Even though the taxonomic approach suggests NM status, it is not clear whether species from this family behave uniformly (Brundrett & Tedersoo 2019).

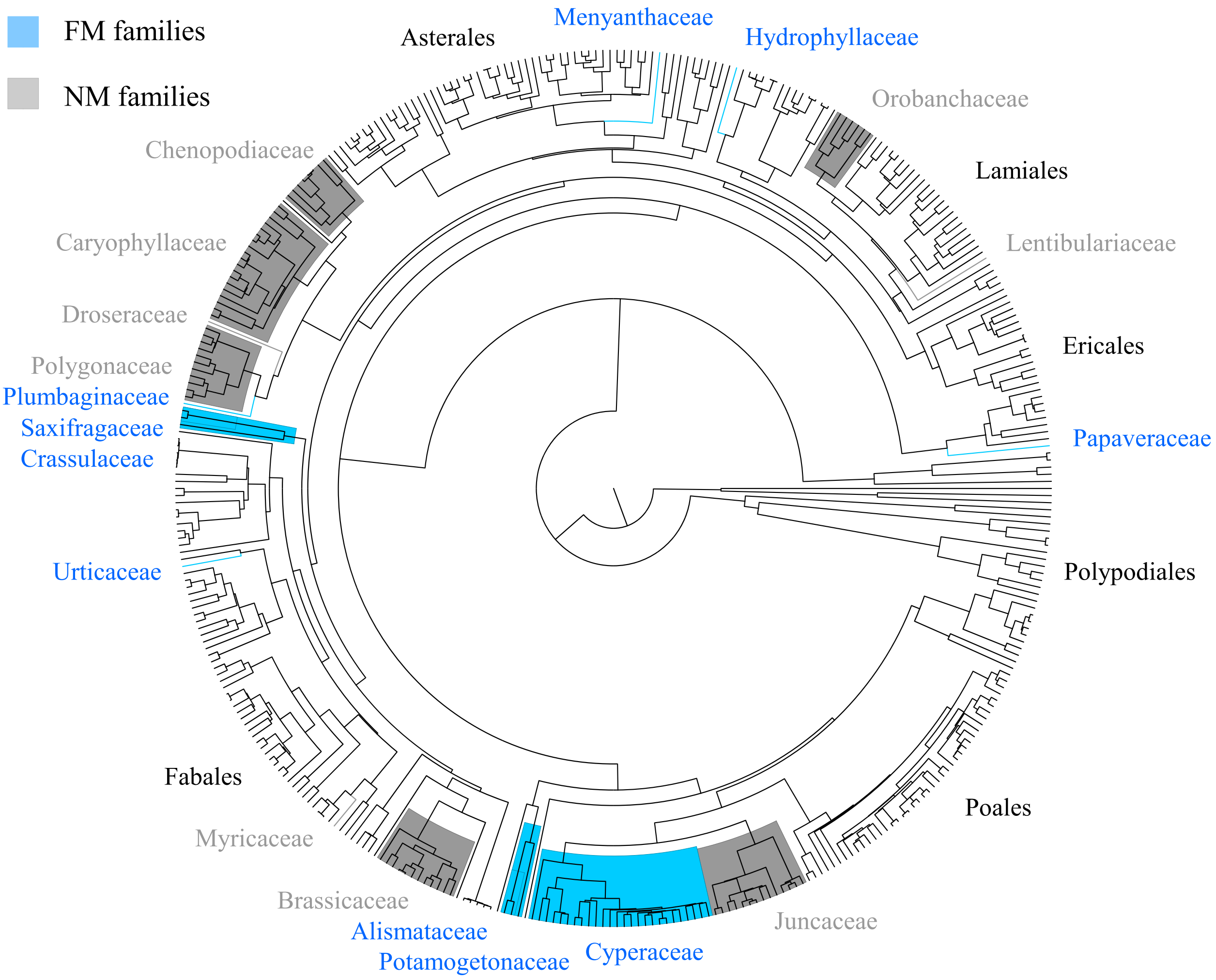


Fig. S2 Phylogeny of the species pool (386 species) of Scottish coastal habitats with depicted FM and NM families (following Brundrett & Tedersoo 2019) and several orders (to allow orientation). We assigned FM status to *Sedum sexangulare* (Crassulaceae) as this mycorrhizal status is supported by the empirical approach (Bueno et al. 2019b). Even though the taxonomic approach suggests NM status, it is not clear whether species from this family behave uniformly (Brundrett & Tedersoo 2019).

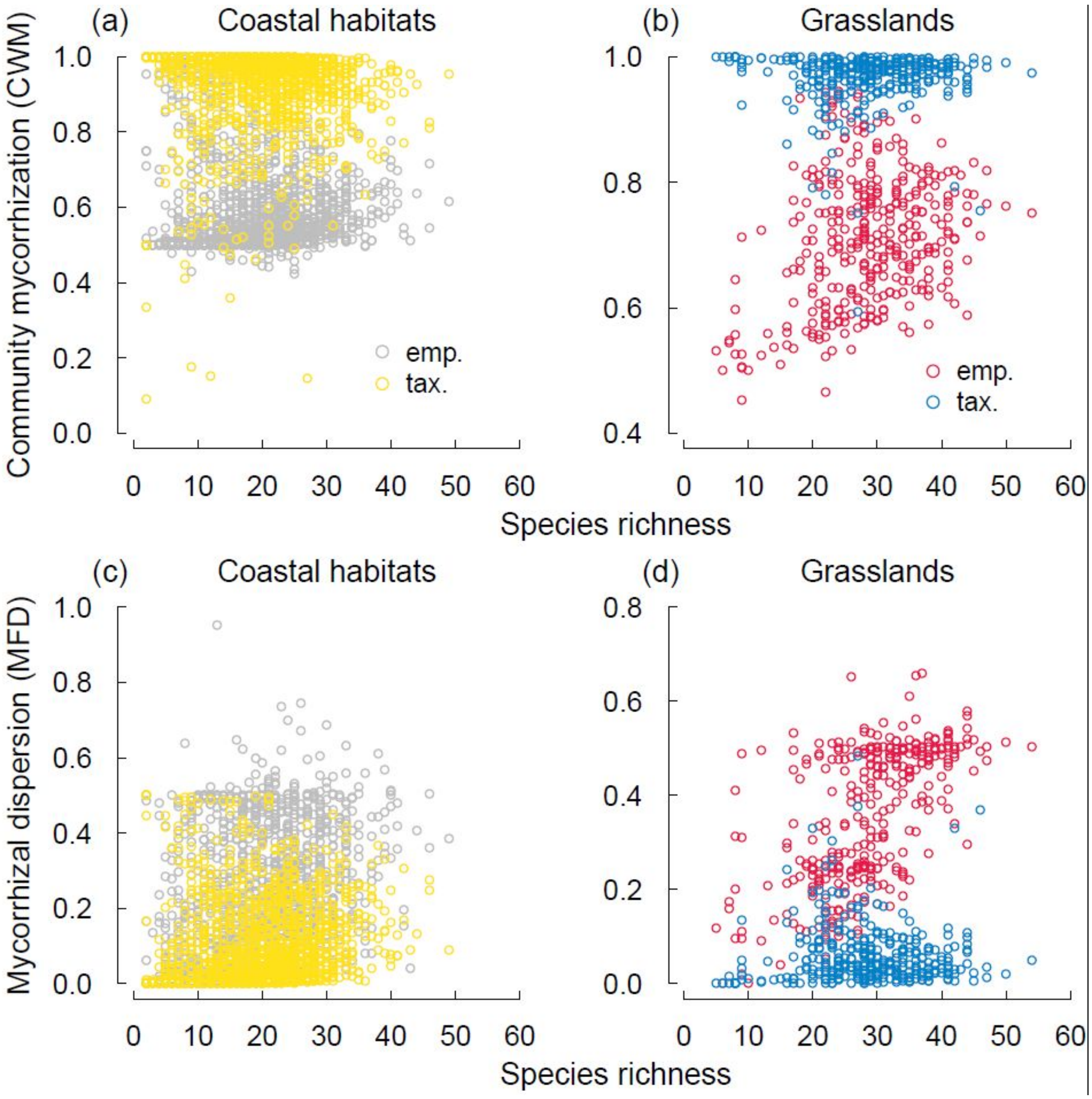


Fig. S3 a-b) Community mycorrhization (CWM) and c-d) community mycorrhizal dispersion (MFD) plotted against species richness in both datasets. Colour coding follows Fig. 1, and indicates estimates based either on the empirical or taxonomic approach. Tests of the heteroscedasticity are given in Table S5.

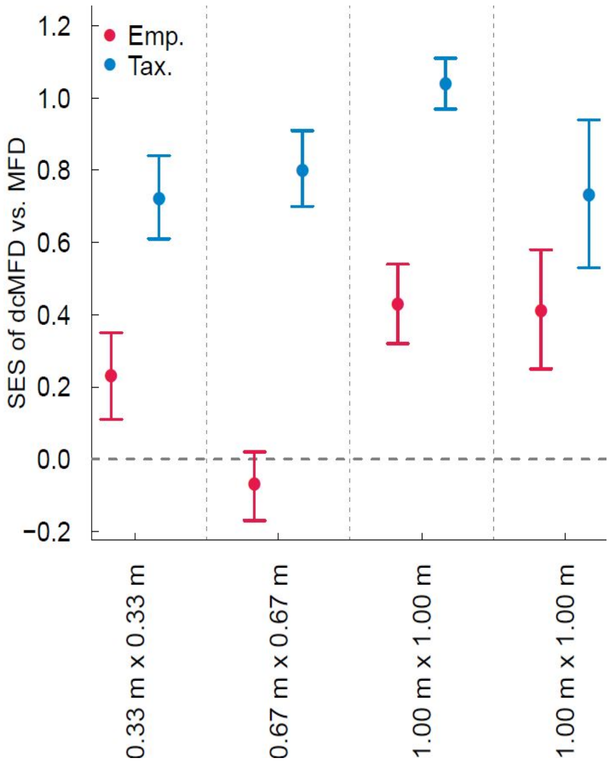


Fig. S4 Mean differences (+/- 95% CI) between SES values of MFD based on trait distances decoupled from phylogeny (dcMFD) and raw trait distances (MFD). SES value were computed using the *ses.mpd* function (*picante*, Kembel et al. 2010) with the "taxa.labels" null model and 999 permutations. Results based on both mycorrhizal assignment aproaches are shown. Non-overlap to zero indicates non-zero difference at the significance level < 0.05 based on paired t-tests. Quadrats were weighted according to the proportion of mismatched species in a quadrat between the empirical and taxonomic approach. Comparison is plotted for all quadrat sizes considered in the grassland dataset (0.33 m x 0.33 m, 0.67 m x 0.67 m and 1.00 m x 1.00 m). The last pair of estimated differences (1.00 m x 1.00 m) is based on the data from the second vegetation survey in grasslands.

Ecological processes in vegetation driven by functional traits and phylogeny
Ph.D. thesis summary

Candidate: Martin Bitomský, Department of Ecology and Environmental Sciences, Faculty of Science, Palacký University in Olomouc

Supervisor: prof. MVDr. Emil Tkadlec, CSc.

Reviewers: doc. RNDr. Zdeňka Lososová, Ph.D.
doc. Mgr. Jiří Doležal, Ph.D.

Defense (place, time and date): Department of Ecology and Environmental Sciences, Šlechtitelů 241/27, 783 71 Olomouc – Holice; Time and Date:

Ph.D. thesis stored in: Department of Ecology and Environmental Sciences, Šlechtitelů 241/27, 783 71 Olomouc - Holice

ABSTRACT

Phylogenies are a crucial (but still overlooked) source of information in ecology. Accompanied with trait data, phylogenetic information can be used to study the most important ecological questions and processes that have implications for nature conservation and community ecology. This recent scientific field is often called community phylogenetics. In this thesis, I summarise our current knowledge about the concept of phylogenetic diversity, the relationship between phylogenetic and functional diversity, and functional and phylogenetic patterns in herbaceous vegetation in response to productivity. I primarily focus on analytical approaches and also provide an up-to-date critical review of some of limitations and issues with popular methods, while I also highlight some important considerations related to every ecological analysis with a phylogenetic and trait framework. Finally, I encourage to routinely implement phylogenetic information (and phylogenetic comparative methods) in ecology because it is needed due to both biological and statistical reasons.

ABSTRAKT

Fylogeneze jsou stěžejním (ale pořád stále přehlíženým) zdrojem informací v ekologii. Společně s daty o funkčních znacích může být fylogenetická informace použita pro studium těch nejdůležitějších ekologických otázek, které mají důsledky pro ochranu přírody a komunitní ekologii. Tento čerstvý vědecký obor se často označuje jako komunitní fylogenetika. V této práci shrnuji naše dosavadní znalosti ohledně konceptu fylogenetické diverzity, vztahu mezi fylogenetickou a funkční diverzitou, či funkčními a fylogenetickými procesy ve vegataci s ohledem na změny v její produktivitě. Primárně se zaměřuji na analytické přístupy a také předkládám

aktuální kritický souhrn limitací a poblémů, které některé populární metody mají, zatímco také zdůrazňuji některé okolnosti vyžadující pozornost v případě všech ekologických analýz s fylogenetickým a funkčním rámcem. Na závěr povzbuzuji k tomu, aby se fylogenetická informace (a fylogenetické komparativní metody) rutinně používala v ekologii, protože je jí potřeba z biologických a statistických důvodů.

PUBLICATIONS

Bitomský M, Mládek J, & Cimalová Š (2018). Light limitation shapes the community seed mass of annual but not of perennial weeds. *Perspectives in Plant Ecology, Evolution and Systematics* 31: 1–6.

Bitomský M, Mládková P, Cimalová Š, & Mládek J (2019). Herbaceous climbers in herbaceous systems are shade-tolerant and magnesium-demanding. *Journal of Vegetation Science* 30: 799–808.

Bitomský M, Mládková P, Pakeman RJ, & Duchoslav M (2020). Clade composition of a plant community indicates its phylogenetic diversity. *Ecology and Evolution* (early view). doi: 10.1002/ece3.6170

Bitomský M, Schaefer H, Pakeman RJ, & Duchoslav M (in review). Variation in mycorrhizal status is much larger within than among sites: a phylogenetically-informed assessment across grassland and coastal floras.

FUNDING

- IGA_PrF_2016_019, IGA_PrF_2017_015, IGA_PrF_2018_020, IGA_PrF_2019_021, IGA_PrF_2020_020 (Palacký University in Olomouc)
- the Scholarship of the city of Ostrava (Ostrava City Hall)
- LTC 18056 (COST action CA16212, Ministry of Education, Youth and Sports)

Introduction

Phylogenies have revolutionised research in plant ecology as they help to understand a mechanistic basis of ecological processes at various spatio-temporal scales (Webb et al. 2002). Thanks to rapidly increasing availability of synthesis-based phylogenies, DNA sequences, computing power and bioinformatic tools, community phylogenetics has become one of the leading fields in ecology despite its quite recent origin (the nineties and the beginning of this millennium).

Chapter One: Phylogenetic Diversity – Key Points

1) Estimation of phylogenetic diversity (PD) is based on phylogenetic distances (the amount of time since the most common ancestor of a pair of species), which are derived from phylogenetic trees. For the purposes of nature conservation or community ecology, phylogenetic distances can be then used to set conservation priorities or to inform our understanding of the assembly mechanisms that structure communities, respectively.

2) PD measures reflect three dimensions of phylogenies: richness, divergence and regularity. This classification comprises more than 70 existing metrics and each dimension can be proxied by a leading measure, namely, Faith's PD (Faith 1992) for the richness dimension, mean phylogenetic distance (MPD; Webb et al. 2002) for the divergence dimension, and variation of pairwise distances (VPD; Clarke & Warwick 2001) for the regularity dimension (Tucker et al. 2017).

3) For the purposes of study questions in community ecology, it is usually recommended to perform a null model analysis in order to standardise a PD metric and account for its dependence on species richness.

4) We propose simple estimates of PD based on taxonomic information (Bitomský et al. 2020).

5) The whole process of PD estimation is affected by important methodological decisions and biases that need to be deeply consider.

6) Inferring regional phylogenies using only species from the species pool is biased in its taxon sampling, which leads to underestimation of PD (Park et al. 2018). Phylogenies with broad taxon sampling are recommended.

7) Null models are not enough to treat for the species richness dependence of PD measures under some conditions. Rarefaction is recommended in some situations (Sandel 2018).

8) PD effects can be confounded by phylogenetic identity effects when assemblages with low or high PD tend to be dominated by a single clade (Hipp et al. 2018). The confounding effect of phylogenetic identity should be visually checked.

9) When opposing patterns in two clades cancel out, it might result in seemingly random or even opposite phylogenetic processes compared to when considering both clades together. Approaches, which explore lineage-specific patterns that allow for opposing mechanisms across the tree topology or specifically evaluate lineage-specific co-occurrence patterns, might be a good idea (Ndiribe et al. 2013, Elliott et al. 2016).

10) Chronograms (dated phylogenetic trees with branch lengths in units of evolutionary time) can provide different estimates of PD than phylograms (branch lengths in units of substitutions per site) (Elliott et al. 2018, Jantzen et al. 2019).

Chapter Two: Phylogeny-Trait Links – Key Points

1) Traits evolve steadily through time, thus, branching patterns of phylogenetic trees, such as topology, branch lengths or tree balance, should also produce comparable estimates of niche space to those obtained by traits. This logic advocates the use of phylogenetic measures in the absence of relevant trait data when traits are sufficiently phylogenetically conserved. Phylogenetic information is easier to obtain, therefore, many studies discuss whether PD serves as a reliable proxy for functional diversity (FD).

2) I summarise two common types of questions regarding the phylogeny-trait links asked in ecological studies, which are tested by different analytical approaches. Question 1: Are there phylogenetic signals in the measured traits in communities? Phylogenetic signal (PS) is the tendency of related species to resemble each other more than species drawn at random from the same tree (Münkemüller et al. 2012).

3) There are various measures of PS that differ in their computation and properties. About 76% of studies in the ecological literature report phylogenetic signal in the measured traits (Cadotte et al. 2019).

4) Question 2: Are PD and FD correlated across multiple sites, i.e. is there the PD-FD relationship? Overall, the PD-FD relationship tends to be tight and positive, which is the expected and intuitive result. However, there is an increasing evidence that interpretation based on parametric correlations can be misleading.

5) Simple correlations are not enough to infer the nature of the PD-FD relationship, while the consideration of alternative evolutionary models and non-linear statistical approaches is necessary (Tucker et al. 2018). Many models of trait evolution have been developed so far to address various evolutionary processes.

6) Inconsistencies in the PD-FD relattionship question some of the most prominent arguments why to use PD in nature conservation and community ecology (Mazel et al. 2018).

7) Some properties of the tree (e.g. taxon sampling, size and shape) affect phylogenetic signal estimates. Functional distances could correlated with phylogeny only for a short distance along the tree (Kelly et al. 2014).

8) The shape of the tree affects phylogenetic distances among species. Trait and phylogenetic distances might have different distributions, which affects statistical properties (Cadotte et al. 2019).

9) Some phylogenies can lack reasonable power to detect relevant patterns of trait evolution (Boettiger et al. 2012). AIC or BIC might favour overly complex (or overly simplistic) evolutionary scenarios (Boettiger et al. 2012). Robustness should be tested.

10) In summary, research efficiency in community phylogenetics would gain from routine correct using of PCM to understand how traits and phylogenetic relationships are linked together. We demonstrated this approach in our study (Bitomský et al. in review).

Phylogeny-Trait-Productivity links – Key Points

1) In the face of the current environmental crisis, phylogenetic relatedness and relevant traits might be very useful to predict changes in the vegetation threatened by eutrophication (nutrient enrichment). Nutrient addition in ecosystems usually increases productivity, i.e. a gradient indicating light limitation and the intensity of competition for light.

2) Higher productivity inevitably results in increased light limitation and the intensity of competition for light, which is one of the most important drivers of plant biodiversity loss (Hautier et al. 2009).

3) Functional differences among plants species occurring along the gradient of productivity can be explained by several key axes of specialisation that include the leaf economy spectrum, plant size and dispersal ability. A very influential approach to proxy these axes is the use of the leaf-height-seed (LHS) plant ecology strategy scheme (Westoby 1998), which is based on three plant traits: specific leaf area (SLA), canopy height and seed mass. Canopy height is a trait representing shade avoidance and the ability to compete for light, while both high SLA and large seeds are trait-states linked to shade tolerance as they increase survival under shady conditions thanks to enhanced carbon gain and well-provisioned seedlings (Bitomský et al. 2018, 2019).

4) Measures of functional diversity can be summarised via an analogous framework as in PD, i.e. functional richness, divergence and regularity (evenness). Additionally, one of the most popular approaches is the concept of community-weighted mean (CWM), i.e. community mean of a trait weighted by species abundances, which is a measure of functional composition.

5) Several recent studies have pointed out several critical issues with some routinely used methods to test the trait-environment and phylogeny-environment relationships. Any metric (e.g. functional or phylogenetic), which is generated at the assemblage/community level by assigning values to species and averaging them for the species present within a cell/plot, can have internal statistical relationships of no biological significance across communities (Hawkins et al. 2017). Many studies might have provided too optimistic statistical support for the relationship between FD, PD measures and environmental gradients when, in fact, there is no relationship. This is a very disturbing fact given the commonness of all criticised methods.

6) For the trait-environment relationship, there are some solutions to control for the type I error, such as the fourth-corner correlation (Peres-Neto et al. 2017), the parametric max test (ter Braak et al 2018) or a method to control for phylogenetic autocorrelation (Duarte et al. 2018).

7) Ecological relevancy of a trait can differ among functional groups or clades. The trait-environment relationship can provide unreliable estimates of the adaptive value of a trait (Bitomský et al. 2018, Laughlin et al. 2018).

8) One should check whether the trait of interest has higher within- or among-site variation in order to focus on ecologically more relevant trait component (Peres-Neto et al. 2017, Bitomský et al. 2019, in review).

9) Assumptions of conventional correlation and regression analyses fundamentally conflict with the ecological concept of limiting factors. Estimating a range of regression quantiles of FD and PD metrics provides a comprehensive description of biological response patterns (Cade et al. 1999).

Synthesis

1) Incorporating phylogenies into ecology is usually due to two reasons: either one wants to use it to study phylogenetic patterns in communities, or one wants to filter its confounding effect. An ecological unifying framework will not work without phylogenetic information because it is needed due to both biological and statistical reasons.

2) First of all, we need to rely on rigorous phylogenetic analyses to generate best phylogenies possible with correct phylogenetic distances among studied species. Second, generated phylogenies need to be examined to check whether the PD-FD patterns truly reflect evolutionary processes but not just some inherent properties of the tree. Third, statistical and phylogenetic comparative methods can be used to infer the PD-FD relationship or trait evolutionary trajectories. Finally, if one is interested in the role of environmental gradients, I would suggest to use a method that has a better performance and controls for the type I error because some popular approaches are a bit outdated and can produce too optimistic results.

REFERENCES

Boettiger C et al. (2012). Is your phylogeny informative? Measuring the power of comparative methods. *Evolution* 66: 2240–2251.

Cade BS et al. (1999). Estimating effects of limiting factors with regression quantiles. *Ecology* 80: 311–323.

Cadotte MW et al. (2019). Do traits and phylogeny support congruent community diversity patterns and assembly inferences? *Journal of Ecology* 107: 2065–2077.

Clarke KR & Warwick RM (2001). A further biodiversity index applicable to species lists: Variation in taxonomic distinctness. *Marine Ecology Progress Series* 216: 265–278.

Duarte LDS et al. (2018). Analyzing community-weighted trait means across environmental gradients: should phylogeny stay or should it go? *Ecology* 99: 385–398.

Elliott MJ et al. (2018). Choice between phylogram and chronogram can have a dramatic impact on the location of phylogenetic diversity hotspots. *Journal of Biogeography* 45: 2190–2201.

Elliott TL et al. (2016). Contrasting lineage-specific patterns conceal community phylogenetic structure in larger clades. *Journal of Vegetation Science* 27: 69–79.

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

Hautier Y et al. (2009). Competition for light causes plant biodiversity loss after eutrophication. *Science* 324: 636–638.

Hawkins BA et al. (2017). Structural bias in aggregated species-level variables driven by repeated species co-occurrences: a pervasive problem in community and assemblage data. *Journal of Biogeography* 44: 1199–1211.

Hipp AL et al. (2018). Effects of phylogenetic diversity and phylogenetic identity in a restoration ecology experiment. In: Scherson RA & Faith DP (eds.). *Phylogenetic diversity: applications and challenges in biodiversity science*. pp. 189–210, Springer.

Jantzen JR et al. (2019). Effects of taxon sampling and tree reconstruction methods on phylodiversity metrics. *Ecology and Evolution* 9: 9479–9499.

Kelly S et al. (2014). Phylogenetic trees do not reliably predict feature diversity. *Diversity and Distributions* 20: 600–612.

Laughlin DC et al. (2018). Survival rates indicate that correlations between community-weighted mean traits and environments can be unreliable estimates of the adaptive value of traits. *Ecology Letters* 21: 411–421.

Mazel F et al. (2018). Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nature Communications* 9: 28888.

Münkemüller T et al. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution* 3: 743–756.

Ndiribe C et al. (2013). Phylogenetic plant community structure along elevation is lineage specific. *Ecology and Evolution* 3: 4925–4939.

Park DS et al. (2018). Taxon sampling effects on the quantification and comparison of community phylogenetic diversity. *Molecular Ecology* 27: 1296–1308.

Peres-Neto PR et al. (2017). Linking trait variation to the environment: critical issues with community-weighted mean correlation resolved by the fourth-corner approach. *Ecography* 40: 806–816.

Sandel B (2018). Richness-dependence of phylogenetic diversity indices. *Ecography* 41: 837–844.

ter Braak CJF et al. (2018). Simple parametric tests for trait–environment association. *Journal of Vegetation Science* 29: 801–811.

Tucker CM et al. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews* 92: 698–715.

Tucker CM et al. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews* 92: 698–715.

Webb CO et al. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics* 33: 475–505.

Westoby M (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil* 199: 213–227.