Universidade Federal do Rio Grande do Sul

Instituto de Biociências

Programa de Pós-Graduação em Ecologia

Tese de Doutorado

Sinal filogenético e conservação filogenética de nicho: integrando métodos aos conceitos ecológicos

Vanderlei Júlio Debastiani

Porto Alegre, Fevereiro de 2016

Sinal filogenético e conservação filogenética de nicho: integrando métodos aos conceitos ecológicos

Vanderlei Júlio Debastiani

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ecologia, do Instituto de Biociências da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Doutor em Ecologia.

Orientador: Prof. Dr. Leandro da Silva Duarte

Comissão Examinadora

Prof. Dr. Valério De Patta Pillar (UFRGS)

Prof. Dr. Nelson Jurandi Rosa Fagundes (UFRGS)

Prof. Dr. Luis Mauricio Bini (UFG)

Porto Alegre, fevereiro de 2016

"Temos uma chance muito maior de realizar algo significativo quando seguimos nossos interesses apaixonados e trabalhamos em áreas de profundo significado pessoal"

Stephen Jay Gould (1941 - 2002)

AGRADECIMENTOS

A tese aqui apresentada é resultado de quatro anos de dedicação e da colaboração de muitas pessoas que contribuíram diretamente ou indiretamente na minha formação. Durante o período do doutorado vivenciei inúmeros momentos que levarei comigo por toda a minha vida. Tive a oportunidade de ler e aprender muitos conceitos e metodologias essenciais para a pesquisa, de participar de discussões importantes para minha formação como ecólogo, de estabelecer várias parcerias com colegas e outros pesquisadores que trouxeram o desafio de resolver problemas práticos e que, desta forma, acabaram também reforçando os laços de amizades. A tese apresentada aqui é apenas uma parte do doutorado que me proporcionou um gigantesco crescimento pessoal e profissional. Dessa forma, quero aproveitar e expressar meus agradecimentos:

Primeiramente aos meus pais, Albino Roque Debastiani e Leni Lenger Debastiani, que sempre me incentivaram na busca das minhas pequenas conquistas. À minha irmã e amiga Jéssica Debastiani, aos meus tios Maria Judita Debastiani e Valdir da Silva e ao meu cunhado Charles Samuel Arend. À minha namorada Bibiane Armiliato de Godoy pelos incentivos, amizade, confiança e ajuda em todos os momentos.

Ao meu orientador, Leandro da Silva Duarte, pelo incentivo à pesquisa, oportunidades proporcionadas e pela amizade construída ao longo destes anos. Também pela orientação, sugestões e críticas que muito contribuíram para a conclusão desta tese.

Aos meus amigos e colegas Fernanda Thiesen Brum, Guilherme Dubal dos Santos Seger, Marcos Bergmann Carlucci, Leandra de Paula Cardoso Pinheiro, Fernando Joner, e especialmente a Larissa Oliveira Gonçalves, Gabriel Nakamura de Souza e Vinícius Augusto Galvão Bastazini pelas discussões e ajuda nas questões que contribuíram significativamente para a melhora da tese.

Aos professores José Alexandre Felizola Diniz Filho, Sergio Camiz, Valério De Patta Pillar, Nelson Jurandi Rosa Fagundes e Luis Mauricio Bini por aceitarem o convite para participar do meu exame de qualificação e banca de defesa.

Aos colegas do Laboratório de Ecologia Filogenética e Funcional, Laboratório de Ecologia Quantitativa, Laboratório de Ecologia Vegetal, Laboratório de Ecologia de Populações e Comunidades, Laboratório de Ecologia de Paisagem, colegas do doutorado e meus amigos. Especialmente a Elisa Viana Salengue, Dirleane Ottonelli Rossato, Lucas Marafina Vieira Porto, Rômulo Silveira Vitória, Thiago Alves Lopes de Oliveira, Jorge Sebastião Bernardo Silva, Rodrigo Bergamin, Grasiela Casas, Rafael Engelman Machado, Daniela Hoss da Silva, André Luís Luza, Taís de Fatima Ramos Guimarães e demais pessoas que contribuíram de alguma forma para a conclusão desta tese.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela oportunidade da pesquisa e bolsa e ao Programa de Pós-graduação em Ecologia da Universidade Federal do Rio Grande do Sul (UFRGS) pela oportunidade de cursar o doutorado.

SUMÁRIO

RESUMO

Compreender os fatores que afetam a distribuição das espécies tem sido um dos principais objetivos dos ecólogos. Atualmente, sabe-se que os processos ecológicos e evolutivos moldam a dinâmica de especiação e extinção de espécies, e determinam a distribuição e abundância das mesmas. Ao longo dos últimos anos, tem havido um aumento no número de estudos que utilizam informação filogenética para explicar as dinâmicas populacionais e as distribuições de espécies, e que buscam identificar os mecanismos responsáveis pela montagem das comunidades. Interações das espécies, sejam elas intraespecífica, interespecífica ou com o ambiente, ocorrem baseadas nas diferenças e semelhanças fenotípicas. Essas variações fenotípicas tem origem na evolução das espécies, e com isso espera-se que as espécies proximamente relacionadas tendam a ser ecologicamente mais semelhantes entre si do que as espécies distantemente relacionadas. Esta concepção tem dado origem a um conceito importante, com implicações para estudos tanto ecológicos quanto evolutivos: o conceito de conservação filogenética de nicho, isto é, quando as espécies relacionadas mantêm seus nichos ancestrais ao longo do tempo evolutivo. Esse padrão tem importância para diversas áreas de ecologia, permitindo a ligação das espécies aos processos ecológicos e auxiliando na maior compreensão da ecologia evolutiva das diferentes linhagens. Devido à sua importância, é fundamental o desenvolvimento de métodos estatísticos adequados para quantificar esses padrões e inferir os processos que o subjazem. Atualmente, os métodos utilizados para inferir conservação filogenética de nicho são, em sua maioria, incompatíveis com determinados conceitos ecológicos e não abrangem todos os tipos de dados e esse fato explica uma visão

incompleta dos processos presentes nas comunidades e conflitante com o objetivo de muitos estudos ecológicos e conservacionistas que buscam vincular as espécies aos processos ecológicos e evolutivos. Desta forma, o principal objetivo desta tese é propor novos métodos para quantificar o sinal filogenético que integrem diferentes aspectos do conceito de nicho ecológico. Apresentamos aqui os novos métodos em detalhes e avaliamos suas propriedades estatísticas (erro tipo I e poder estatístico) por meio de dados simulados. No capítulo 1, nós propomos um método para medir sinal filogenético utilizando o teste de Mantel, incorporando modelos evolutivos para testar hipóteses específicas da evolução dos atributos. No capítulo 2, descrevemos um conjunto de funções e um novo pacote estatístico para explorar os padrões filogenéticos no nível de metacomunidade. Este pacote permite explorar a distribuição de linhagens filogenéticas através de gradientes ecológicos, a análise de sinal filogenético no nível da metacomunidade e explorar a associação entre clados e gradientes ecológicos. No capítulo 3, investigamos a relação entre sinal filogenético dos atributos com os padrões de coocorrência das espécies nos níveis da comunidade. Esta abordagem permite testar se espécies filogeneticamente relacionadas que coocorrem expressam as suas dimensões de nicho com maior semelhança do que seria esperado por modelos neutros de evolução. Por fim, testamos as propriedades estatísticas destes métodos em relação dois modelos nulos, que incorporam diferentes aspectos da estrutura da comunidade e evolução dos atributos das espécies. Os três capítulos representam diferentes trabalhos que se interconectam no sentido de elucidar o conceito de sinal filogenético e conservação filogenética de nicho.

Palavras-chave: modelos evolutivos, evolução de atributos, modelos nulos, padrões filogenéticos, propriedades estatísticas, erro tipo I, poder estatístico.

ABSTRACT

Understanding the factors that can affect species distributions has been a main goal of ecologists. Currently, it is known that evolutionary and ecological processes shape the speciation dynamics, species extinction and determine the distribution and abundance of species. Over the last years, there has been an increase in the number of studies using phylogenetic information to explain the dynamics of population, species distribution and identifying the mechanisms of community assembly. Species interactions – intraspecific, interspecific or with the environment – occur based on their phenotypic differences and similarities. As phenotypic variation has a basis in evolutionary history, it is expected that closely related species tend to be more ecologically similar to each other than distantly related ones. This notion has given rise to an important concept, with implications for both evolutionary and ecological studies: the concept of phylogenetic niche conservatism, that is, when related species maintain their ancestral niches over evolutionary time. This pattern is important for several areas of ecology, and allows to link species to ecological processes and to understand the evolutionary ecology of different lineages. Despite its importance, it is crucial the development of appropriate statistical method to measure this pattern and to infer the processes behind it. The methods currently available to infer phylogenetic niche conservatism are sometimes incompatible with some ecological concepts and do not cover all kind of data, this fact leads to an incomplete view of the process acting in the currents communities and conflict with the goal of many ecological and conservation studies that need to link species to ecological and evolutionary processes. The main goal of this dissertation is to propose novel methods to measure

phylogenetic signal incorporating different aspects of ecological niche. We introduce novel methods in detail and evaluate its statistical properties (type I error and statistical power) by means of simulated data with known structure. In chapter 1 we propose a method to measure phylogenetic signal using the Mantel test, incorporating evolutionary model to test specific hypothesis of trait evolution. In chapter 2, we describe a set of function and a new statistical package for exploring the phylogenetic patterns at the metacommunity level. This package allows the exploration of distribution of phylogenetic lineages across ecological gradients, the analysis of phylogenetic signal at metacommunity level and to explore the association between clades and ecological gradients. In the chapter 3, we access the relationship between phylogenetic signal in traits and species co-occurrence patterns in the community levels. This approach allows one to test whether phylogenetic close related species co-occurring in metacommunities express their niche dimensions more similarly than would be expected by neutral expectation. We tested the statistical properties of these methods in relation to two null models, which incorporate these different aspects of the community structure and evolution of species traits. The three chapters represent different works that are interconnected in order to elucidate the concept of phylogenetic signal and phylogenetic niche conservatism.

Key-words: evolutionary model, trait evolution, null model, phylogenetic patterns, statistical properties, type I error, statistical power.

ÍNDICE DE TABELAS

PCPS modeling the CWM following the sequential approach (PCPS1, PCPS1+PCPS2, and so on). TS null model shuffle terminal tips across the phylogenetic tree and BM null model estimate neutral phenotypic diffusion across the metacommunity. The rejection rate is calculated independently for each model, considering the target level of $\alpha = 0.05$, considering the proportion of curves that are above the null curves under both null models. PhyLow are results with traits without phylogenetic signal (Grafen's $\rho = 0.0001$); PhyBM are results with traits under Brownian motion (Grafen's $\rho = 1.0$); PhyHigh are results with traits conserved with high phylogenetic signal for the pool of species (Grafen's $\rho = 5.0$); FilterLow are metacommunity without any habitat filter (Filter = 0); FilterMedium shows results with medium habitat filter (Filter = 0.5); and FilterHigh shows results with high phylogenetic habitat filter in the metacommunity (Filter = 0.9). More details can be found in main text……116

Table S2. Rejection rate with intermediate levels of phylogenetic signal, for each null model. Each model (λ 1, λ 1:2, λ 1:3, …) shows the results for each PCPS modeling the CWM following the sequential approach (PCPS1, PCPS1+PCPS2, and so on). TS null model shuffle terminal tips across the phylogenetic tree and BM null model estimate neutral phenotypic diffusion across the metacommunity. The rejection rate is calculated independently for each model, considering the targed level of $\alpha = 0.05$, considering the proportion of curves that are above the null curves under both null models. Phy_1.2, Phy_1.5, Phy_2.0 with traits conserved with phylogenetic signal for the pool of species (Grafen's $\rho = 1.2$, Grafen's $\rho = 1.5$, Grafen's $\rho = 2.0$, respectively); FilterLow are metacommunity without any habitat filter (Filter = 0); FilterMedium shows results with medium

ÍNDICE DE FIGURAS

INTRODUÇÃO GERAL

Entender como as espécies se distribuem no ambiente é um dos principais objetivos da ecologia. Para isso são consideradas tanto as interações intraespecíficas e interespecíficas quanto interação das espécies com o ambiente. Nesse contexto, o conceito de nicho ecológico talvez seja um dos mais básicos e importantes. O conceito está ligado aos requerimentos dos organismos e aos efeitos que os organismos têm sobre o ambiente. O nicho ecológico permite ligar as características morfológicas, fisiológicas e comportamentais das espécies conforme elas interagem com outras espécies e com o ambiente, as distribuições das populações e o funcionamento dos ecossistemas.

Ao longo do tempo muitas definições do termo nicho ecológico foram propostas sendo que a primeira descrição formal do termo é atribuída a Grinnell (1917), relacionando nicho ao requerimentos dos indivíduos que permitem a sobrevivência da espécie. Segundo Grinnell, nicho se refere ao lugar em um ambiente onde uma espécie ocupa, ao conjunto de condições necessárias para a espécie existir, como habitat e alimentação, limitações morfológicas e interação com outros organismos. Uma outra definição de nicho ecológico, descrita por Elton (1927), refere-se ao papel funcional da espécie dentro do ecossistema e seu impacto sobre o ambiente, ou seja, aos efeitos das espécies em um determinado habitat. As primeiras definições de nicho geram uma distinção entre os conceitos, Elton destaca os efeitos da espécies no ambiente enquanto Grinnell destaca os efeitos do ambiente sobre as espécies. Uma das definições mais conhecidas de nicho foi elaborada por Hutchinson (1957). Hutchinson destaca nicho como o conjunto de n-dimensões de um

hipervolume de condições ambientais que a espécie precisa para viver, crescer e se reproduzir, salientando também diferenças entre o conjunto de condições na presença ou ausência de interações interespecíficas. MacArthur et al. (1967), na teoria de nicho, destacam ainda a competição intraespecífica com um importante papel para estruturar as comunidades. Uma abordagem mais moderna de nicho foi elaborada por Chase e Leibold (2003), definindo nicho como a descrição conjunta das condições ambientais que permitem a uma espécie satisfazer os requerimentos mínimos para que a taxa de nascimento de uma população seja igual ou maior que a taxa de mortalidade, junto com o conjunto dos efeitos per capita da referida espécie nessas condições ambientais. Essa abordagem é uma tentativa de integrar tanto o conceito de nicho elaborado por Elton quanto o elaborado por Grinnell, ou seja, considerar tanto os efeitos quantos os requerimentos das espécies no ambiente.

O conceito de nicho traz consigo a concepção de diferença e equivalência entre as espécies. Considerando que as espécies não vivem isoladas e interagem dinamicamente com as demais, o nicho ecológico tem sido usado para explicar os padrões no nível de comunidade, como coocorrência e distribuição das abundâncias. Essas características que acompanham as espécies embasam a concepção de equivalência funcional das espécies e o princípio de exclusão competitiva (Gause 1934), fazendo com que algumas espécies possam ser mais fortemente limitadas do que outras dependendo da similaridade em ternos de nicho observada entre elas. A teoria contemporânea de coexistência entre espécies enfatiza que a coexistência depende tanto das diferenças de nicho quanto das diferenças de *fitness* (HilleRisLambers et al. 2012). Essas diferenças promovem crescimento populacional diferencial e dependem de condições ambientais específicas bem como da composição de espécies únicas para a comunidade, de modo que diferenças de nicho

poderiam ser importantes para explicar os mecanismos de estruturação de comunidades (Webb et al. 2002; Cavender-Bares et al. 2009; Gerhold et al. 2015). Embora os efeitos de nicho sejam importantes, alguns padrões de diversidade de espécies e distribuição de abundância podem ser explicados sem invocar nicho ecológico (Hubbell 2001), assumindo que as espécies que ocupam uma determinada comunidade são ecologicamente equivalentes.

O nicho ecológico não tem apenas importância nos processos ecológicos, mas também importância nos processos evolutivos (Wiens et al 2010). As espécies vivem no ambiente e com o passar do tempo podem passar por vários processos que levam à evolução das populações ou a processos de especiação, gerando novas espécies. Evolução é um processo onde as espécies se modificam a partir de um ancestral comum, sendo que vários mecanismos podem resultar na evolução das espécies como mutação, fluxo gênico, deriva genética e seleção natural. Embora nem todos os processos evolutivos tenham como base processos ecológicos, alguns destes processos são determinados por heterogeneidade ambiental e heterogeneidade geográfica, bem como por variáveis ambientais que atuam sobre as populações locais. A interação entre os processos ecológicos e evolutivos pode gerar e manter padrões de diversidade de espécies, moldando a dinâmica de especiação, extinção, bem como os padrões de distribuição e de abundância das espécies (Webb et al. 2002; McPeek 2008; Pelletier et al. 2009; Pillar e Duarte 2010; Schoener 2011). As espécies são originadas a partir de um ancestral comum o que explica por que boa parte da informação genética é compartilhada entre as espécies descendentes. Como consequência, é esperado que espécies filogeneticamente próximas sejam similares em termos de atributos, sejam eles atributos morfológicos, fisiológicos ou comportamentais (Harvey e Pagel 1991). Assim, espera-se que espécies

filogeneticamente relacionadas também sejam semelhantes em relação às suas exigências ecológicas e de nicho (Prinzing et al. 2001). Esse padrão de espécies filogeneticamente próximas apresentarem atributos similares é conhecido como sinal filogenético (Blomberg et al 2003). O sinal filogenético tem implicações em vários campos da biologia, como biogeografia, estrutura de comunidades, ecologia de ecossistemas, resposta das espécies às mudanças climáticas e biologia da conservação.

Portanto, se espécies mais aparentadas tendem a apresentar atributos similares, e visto que os atributos estão intimamente relacionados com o ambiente, esperamos que os nichos de espécies aparentadas também sejam semelhantes. Um fenômeno relacionado ao sinal filogenético compreende a tendência de espécies reterem seu nicho ao longo do tempo evolutivo, isso é chamado de conservação de nicho ou conservação filogenética de nicho (Wiens e Graham 2005). Um longo debate tem sido feito para saber se as espécies retêm ou não seu nichos ecológicos ao longo do tempo (Losos et al. 2003; Wiens e Graham 2005; Losos 2008; Crisp et al. 2009; Cooper et al. 2010; Wiens et al. 2010; Losos 2011). O significado desse conceito é bastante ambíguo pois é usado tanto para descrever um padrão de alta similaridade de nicho em espécies filogeneticamente relacionadas quanto para o conjunto de processos compreendendo os mecanismos evolutivos responsáveis por essa similaridade (Pyron et al. 2015). Ainda, o padrão de sinal filogenético pode não representar conservação filogenética de nicho já que diferentes processos evolutivos podem produzir sinal filogenético semelhante e processos evolutivos semelhantes podem produzir diferentes assinaturas em termos de sinal filogenético (Revell et al. 2008), de modo que o padrão por si só não revela o processo evolutivo envolvido. Wiens e Graham (2005) citam quatro processos que podem gerar conservação de nicho, sendo eles a

seleção estabilizadora, fluxo gênico, falta de variação genética e restrições genéticas que impedem as populações de se adaptarem a novos nichos. Desta forma, o processo de conservação filogenética de nicho é a soma dos efeitos das restrições endógenas dos organismos com variáveis ambientais e geográficas que levam as populações a conservar seus nichos ao longo do tempo evolutivo (Pyron et al. 2015). Portanto conservação filogenética de nicho não envolve apenas uma correlação entre os atributos das espécies e filogenia, existem conceitos ecológicos que devem ser incorporados com o objetivo de relacionar os processos evolutivos e ecológicos presentes nas comunidades.

Para quantificar sinal filogenético e inferir conservação filogenética de nicho devem ser analisados conjuntos de espécies ao invés de uma única espécie. Sendo assim um requerimento das análises é a disponibilidade de estimativas filogenéticas, preferencialmente com datações dos ramos, para o conjunto de espécies a ser considerado. Essa informação é essencial pois é necessário conhecer as relações filogenéticas entre as espécies. Além das filogenias, é necessário estimar valores que indiquem o nicho ecológico de cada espécie (Rosado et al 2016). Geralmente é muito difícil de estimar diretamente o nicho das espécies, por isso são consideradas uma ou mais características das espécies que indiretamente reflitam ou estejam relacionadas com as características de nicho. Uma das abordagens para quantificar o sinal filogenético consiste em comparar a semelhança observada nos atributos das espécies aparentadas com um conjunto de atributos extraído aleatoriamente de uma árvore filogenética (Blomberg et al. 2003). Outra abordagem considera como conservação de nicho apenas quando as espécies são mais similares que o esperado por modelos evolutivos neutros, ou seja, que não pressupõem mecanismos de seleção extra (ex. seleção estabilizadora ou divergente) (Losos 2008). Um dos modelos de evolução

neutra mais simples é o movimento Browniano (Felsenstein 1985), onde as diferenças entre os atributos das espécies são acumuladas ao longo do tempo evolutivo (Freckleton e Harvey 2006). Neste modelo as diferenças são acumuladas por deriva genética sem forte seleção estabilizadora, sem seleção divergente ou mudanças na taxa evolutiva ao longo a árvore filogenética. Outra abordagem consiste em considerar conservação filogenética de nicho apenas quando as espécies filogeneticamente relacionadas tendem a ter exigências de habitat semelhantes e coocorrendo assim, em habitats semelhantes (Pillar e Duarte 2010; Ulrich et al. 2012; Duarte et al. submetido). Apesar dos avanços, tanto teóricos quanto analíticos sobre conservação filogenética de nicho, um ponto pouco contemplado é a coocorrência das espécies na comunidade e sua influência nos padrões de sinal filogenético. Nesse contexto, as coocorrências atuais na comunidade são um aspecto importante a ser considerado para acessar conservação filogenética de nicho. Assim, conservação de nicho seria definido quando há sinal filogenético nas espécies e uma elevada relação entre a estrutura filogenética das comunidades e a variação dos atributos ao nível da comunidades (Duarte et al. submetido).

Para quantificar o sinal filogenético é preciso ter métodos e ferramentas apropriadas (Pagel 1999; Blomberg et al. 2003; Pillar e Duarte 2010; Fritz e Purvis 2010; Boettiger et al. 2012; Diniz-Filho et al. 2012; Adams 2014). Mesmo com vários métodos disponíveis, eles não abrangem todas as situações, tipos de dados ou aspectos relacionados a avaliação de sinal filogenético. Desta maneira, esta tese tem como objetivo propor novos métodos para quantificar o sinal filogenético, de modo que incorporem conceitos não contemplados ou raramente utilizados nos métodos existentes, tais como a coocorrência das espécies e a comparação com modelos evolutivos e, com isso, reduzir algumas das limitações existentes nos métodos atuais.

Além de apresentar novas metodologias, usamos simulações numéricas para testar suas propriedades estatísticas. Com base nos dados simulados é possível estimar importantes propriedades estatísticas, como taxa de erro tipo I (falsos positivos) e poder dos métodos propostos. Ainda, tivemos como objetivo disponibilizar de uma maneira acessível os algoritmos dos métodos para que sejam facilmente utilizados por outros pesquisadores. As funções foram escritas para o ambiente R, um software livre e bastante utilizado na área de ecologia. A tese encontra-se estruturada em três capítulos, cada um correspondente a um artigo científico, formatado de acordo com as normas dos periódicos para onde foram ou serão submetidos.

No primeiro capítulo propomos um método para medir sinal filogenético usando o teste de Mantel. Embora o uso deste teste não seja uma novidade, propomos a inclusão de modelos evolutivos para avaliar hipótese específica de evolução dos atributos. A inclusão de distintos modelos evolutivos permite verificar se os atributos em questão evoluem mais ou menos que o esperado por tais modelos. Uma das principais críticas ao uso do teste de Mantel para quantificar o sinal filogenético é o fato de não incorporar modelos neutros de evolução, portanto, nossa proposta resolve essa importante crítica ao incorporar os modelos evolutivos. Uma das vantagens na utilização do teste de Mantel é o fato de possibilitar o acesso ao sinal filogenético utilizando atributos categóricos e binários, sendo que estas duas últimas categorias, muitas vezes, são a principal forma de caracterizar o nicho de inúmeros grupos taxonômicos.

O segundo capítulo é resultado da criação de um conjunto de funções que permitem explorar padrões filogenéticos no nível de metacomunidade. Neste capítulo descrevemos as funcionalidades do pacote de funções, chamado de PCPS, para o

ambiente R. Este permite calcular gradientes filogenéticos (também denominados Coordenadas Principais de Estrutura Filogenética, PCPS), analisar sinal filogenético no nível de metacomunidade, bem como associar diferentes linhagens filogenéticas a gradientes ecológicos.

No terceiro capítulo investigamos a relação entre sinal filogenético e distribuição de espécies e clados em metacomunidades. Essa abordagem permite verificar se espécies filogeneticamente próximas tendem a apresentar exigências de habitat semelhantes ocorrendo, assim, em habitats semelhantes. Poucos estudos têm incorporado a coocorrência de espécies para quantificar sinal filogenético, sendo essa uma das grandes novidades para inferir conservação filogenética de nicho. Essa abordagem também incorpora a comparação com modelos de evolução dos atributos com o objetivo de quantificar se espécies coocorrentes nas comunidades que compõem a metacomunidade, e próximas filogeneticamente, expressam suas dimensões de nicho de forma mais semelhante do que seria esperado por tais modelos evolutivos. Para isto, foram testadas as propriedades estatísticas dos métodos em relação a dois modelos nulos, estes que incorporam diferentes aspectos da estrutura de comunidade e da evolução dos atributos.

CAPÍTULO 1 - Evolutionary models and phylogenetic signal assessment via Mantel test*

Abstract

Phylogenetically closely related species tend to be more similar to each other than to more distantly related ones, a pattern called phylogenetic signal. Appropriate tests to evaluate the association between phylogenetic relatedness and trait variation among species are employed in a myriad of eco-evolutionary studies. However, most tests available to date are only suitable for datasets describing continuous traits, and are most often applicable only for single trait analysis. The Mantel test is a useful method to measure phylogenetic signal for multiple (continuous, binary and/or categorical) traits. However, the classical Mantel test does not incorporate any evolutionary model (EM) in the analysis. Here, we describe a new analytical procedure, which incorporates explicitly an evolutionary model in the standard Mantel test (EM-Mantel). We run numerical simulations to evaluate its statistical properties, under different combinations of species pool size, trait type and number. Our results showed that EM-Mantel test has appropriate type I error and acceptable power, which increases with the strength of phylogenetic signal and with species pool size but depended on trait type. EM-Mantel test is a good alternative for measuring phylogenetic signal in binary and categorical traits and for datasets with multiple traits.

* Manuscrito submetido para o periódico *Oecologia*. Leandro da Silva Duarte é coautor do manuscrito.

Key Words: Trait evolution, Brownian motion, null model, type I error, statistical power

Introduction

Species are not independent biological and statistical units, but share part of their evolutionary history. This circumstance makes closely related species to be more similar to each other than to more distantly related ones (Harvey and Pagel 1991). The similarity between species is manifested in morphological and physiological traits, behavioral characteristics and other characteristics that express species life history patterns. Although phylogenetic conservatism in traits is a tendency observed for different biological groups, appropriate tests to evaluate the association between phylogenetic relatedness and trait variation among species, the so-called phylogenetic signal (Blomberg et al. 2003), are required. Furthermore, evaluating phylogenetic signal in traits for a given set of species is crucial to detect phylogenetic niche conservatism (Wiens and Graham 2005; Losos 2008; Cooper et al. 2010; but see also Revell 2008) and might possibly help to explain species assembly patterns into communities (Webb et al. 2002; Mouquet et al. 2012; but see Gerhold et al. 2015).

The measurement of phylogenetic signal has received much attention over the last decade. Up to date, most of the methods available to measure phylogenetic signal are suitable only for continuous traits, and frequently are applicable only for datasets which contain only one trait, such as Pagel's λ (Pagel 1999), K statistic (Blomberg et al. 2003), and phylogenetic eigenvector regression analysis (PVR) (see Diniz-Filho et al. 2012 for a recent review and expansion). More recently, Fritz and Purvis (2010) developed a measure to evaluate phylogenetic signal in binary traits (D statistic).

Similarly to other methods, D statistic is only capable of estimating phylogenetic signal for single traits. However, the use of a single trait is incompatible with the goal of many ecological (Pillar and Duarte 2010) and conservation studies (Lefcheck et al. 2015), that need to integrate more than one trait in order to link species to ecological processes and to understand the evolutionary ecology of different lineages. Despite of the importance of using multiple traits, insofar, few statistics have been developed to quantify phylogenetic signal in multiple traits. For continuous traits, the generalized K statistic (Adams 2014) and the multivariate extension of PVR (Diniz-Filho et al. 2014) are two possible options. Another option for measuring phylogenetic signal for multiple traits is the Mantel test (Mantel 1967; Losos et al. 2003; Pillar and Duarte 2010), which is a statistical procedure to evaluate the correlation between two resemblance matrices (Mantel 1967). One of the main advantages of the Mantel test, over the other methods, for estimating phylogenetic signal in multiple traits is that it allows the use of multiple types of traits (continuous, binary and/or categorical) simultaneously. In the particular case of evaluating phylogenetic signal, the first resemblance matrix consists of pairwise phylogenetic distances between species and the second is a pairwise trait dissimilarity matrix between species. Although the Mantel test has been criticized in the past years, mainly due to its low power and high rates of type I error (Legendre and Fortin 2010; Guillot and Rousset 2012), it is still useful when the hypotheses are formulated in terms of distances (Legendre and Fortin 2010).

The assessment of phylogenetic signal in traits via Mantel tests has faced some criticism over the last years (Harmon and Glor 2010; Seger et al. 2013). In both studies, authors found that the performance of Mantel test in phylogenetic signal analysis showed lower power when compared with other methods for estimating

phylogenetic signal, such as K statistics (Blomberg et al. 2003) and PVR (Diniz-Filho et al. 1998). Nonetheless, Hardy and Pavoine (2012) have demonstrated that using Abouheif proximities (Abouheif 1999) as pairwise phylogenetic distances between species, instead of the sum of branch lengths between a pair of species in a phylogenetic tree, increase the statistical power of the standard Mantel test. Furthermore, the standard Mantel test does not explicitly assume any evolutionary model in the assessment of phylogenetic signal in traits, and therefore such tests are not useful for inferring phylogenetic niche conservatism (Losos 2008; but see Pillar and Duarte 2010). The inflation of type I error, associated with the absence of an evolutionary model in the Mantel test, could be a complicating factor in evaluating phylogenetic signal. By taking into account an evolutionary model for trait evolution, we might change the original null hypothesis of the Mantel test to a more realistic one stating that the association between phylogenetic distances and trait resemblances do not deviate from the expected by a particular model of evolution. A neutral model commonly used in several measure of phylogenetic signal is the Brownian motion model, which ascertains that evolutionary change in trait expression is a function of a constant mutational rate along time. Therefore, trait variation between any two species is directly proportional to the amount of time since their evolutionary divergence (Freckleton and Harvey 2006). This model is useful because it does not make extra mechanistic assumptions to explain trait evolution, such as stabilizing or divergent selection; species inherit their traits values from ancestors and simply diverge as a function of evolutionary time.

Moving from the classical null hypothesis to a more realistic one involves explicitly incorporating an evolutionary model in the assessment of phylogenetic signal via the Mantel test. To accomplish that goal we incorporate a trait simulation

under a particular evolutionary model (EM) approach to the standard Mantel test in order to estimate the phylogenetic signal in the traits. Hereafter, we call such approach as EM-Mantel. The method differs from the standard Mantel test in relation to the evaluation of the significance of the Mantel statistic. In the original framework (Mantel 1967), the null hypothesis states that association between the pairwise resemblance matrices does not differ from zero. This null hypothesis is naïve for testing the phylogenetic signal, in as much as large part of the evolutionary history of species is shared among them; therefore, it is expected that phylogenetic and trait distances are related at some level in the standard Mantel test. In EM-Mantel test, the null hypothesis states that the association between the pairwise phylogenetic distance matrix and the trait dissimilarity matrix does not differ from that expected from a particular evolutionary model. Such null hypothesis is similar to those implemented for K and D statistic using the Brownian evolutionary model (Ackerly 2009; Fritz and Purvis 2010).

In this paper we describe the analytical procedures used to perform the EM-Mantel tests in detail and evaluate its statistical properties (type I error rate and statistical power) by means of simulated data with known structure, and by changing species richness in the pool, type and number of traits. We also compare the results of the EM-Mantel to those of the standard Mantel test.

Material and methods

EM-Mantel test

The standard Mantel test (1967) is a method conceived to evaluate the association between two distance matrices, in this case the elements of one matrix contain the pairwise evolutionary distance between species and the other their phenotypic distances, and to test a hypothesis about the relationship between these matrices. The null hypothesis is that the distances among objects in a matrix containing response variables are not linearly correlated with another matrix of explanatory variables (Legendre and Legendre 2012). The significance of the test is assessed by permuting the species vector in one of the original matrices and by rearranging the distance matrix according to the permutation of the species vector. This procedure produces a distribution of the Mantel statistic under the null hypothesis to which the observed Mantel statistic is compared.

The Mantel statistic is not restricted to access the relationship between two sets of empirical data. It may also be used to assess an a priori hypothesis (Legendre and Legendre 2012). To do so, a model matrix is constructed in order to represent a specific hypothesis to be tested. The standard Mantel statistic does not have an expected value for an evolutionary model, so the resulting statistic will vary based on the internal topology of the phylogeny, branch lengths, type of traits and the number of species. We can use the model matrix approach to generate a Mantel test, denoted EM-Mantel, which tests if the observed Mantel statistic differs from the expected given a specific evolutionary model. The EM-Mantel test establishes a model matrix where objects are expected to represent trait values as if they evolved under a given evolutionary model (here we considered only Brownian motion, but other evolutionary models might be used instead). For this, a set of traits, with the same characteristics of observed traits (e.g., type and number of traits, structure of correlation between traits, frequency of occurrence for each level, among others), is

simulated under an evolutionary model based on an observed phylogenetic tree. Given a model of evolution the distribution of traits can be predicted by the evolutionary time, given by the branch lengths in the phylogenetic tree (Paradis 2012). For the Brownian motion, for example, branch lengths are required to estimate the variance parameter of a Gaussian distribution, which will be used to generate continuous trait values. If necessary, a continuous trait can be transformed into a binary or a categorical trait using a threshold model, which restrains its frequency of occurrence for each level, as observed in the original trait set (Fritz and Purvis 2010). This simulation is performed *n* times to take into account the variation of the evolutionary model. At each time step, a distance matrix is computed for each set of simulated traits, following the same procedure that generated the distance matrix of observed trait values, then the association between the simulated trait distance matrix and the original phylogenetic distance matrix is assessed based on the Mantel correlation statistic. Finally, the Mantel statistics obtained through simulated traits are compared to the observed Mantel statistic, as in the standard Mantel test. The distribution of all values of Mantel statistics under this scenario produces a one-tailed p-value distribution that expresses the probability of the observed Mantel statistic being greater than or equal to the expected value under a particular evolutionary model (e.g., Brownian motion). Alternatively, the p-value in EM-Mantel could be interpreted as a two-tailed test, expressing the probability of the observed Mantel statistic being greater or smaller to the expected value under a particular evolutionary model.

Significant results in the standard Mantel test show that the phylogenetic distances are related with the distances based on species traits, but do not distinguish such association from that expected given an evolutionary model of trait evolution. If

correlation coefficients from the standard Mantel test show statistical significance, EM-Mantel should be performed for estimating whether coefficients differ from the expected by an evolutionary model of trait evolution. Therefore the EM-Mantel test should be used together with standard Mantel test for assessment the phylogenetic signal.

Estimating type I error and statistical power of the EM-Mantel test

We used numeric simulations to compare the statistical properties of the standard Mantel and EM-Mantel test. We simulated ultrametric phylogenetic trees with purely stochastic birth with homogeneous speciation rates of 0.1 and zero extinction rates using the function *sim.bdtree* of package *geiger* (Harmon et al. 2008). This process produces a conservative distribution of speciation events on the tree from past to present that is intermediate between splitting closer the root and closer the present (Mooers et al. 2012). Based on these simulated phylogenetic trees, we changed several parameters (number of species in the phylogenetic tree, number of traits, type of traits and degree of trait conservatism) to generate different phylogeny-trait association scenarios.

We tested four levels of conservatism in trait evolution: Brownian, labile and two levels of trait conservatism, all compared with the Brownian model. To generate the different patterns of trait evolution we used the Grafen's method (Grafen 1989) to transform the branch lengths of the phylogenetic tree. These transformations were achieved by different exponentiation of branch lengths, denoted by ρ. Grafen's ρ values lower that 1.0 shrink deeper branches and lengthen those near the tips, whereas values higher that 1.0 elongate branch lengths near the root of the tree. Using these

phylogenetic trees with altered branches, we simulated continuous traits evolving under Brownian evolutionary model using the function *rTraitCont* of the *ape* package (Paradis 2004). Accordingly, where trait evolution is constant along time, and the variance between species is proportional to time-calibrated branch length since evolutionary divergence between them (Freckleton and Harvey 2006), the combinations of Grafen's ρ values and Brownian model generate different phylogenetic signal for the simulated trait (Diniz-Filho et al. 2012; Seger et al. 2013). Under Grafen's ρ = 0.0001, a trait does not show phylogenetic signal, that is to say that the traits evolved more rapidly than expected by the Brownian model, and the trait variation between closely related species was very high. When Grafen's ρ = 1.0, traits evolved according to Brownian model. Under Grafen's ρ = 2.0 or 5.0 the traits evolved less than the expected by the Brownian model. In this case, closely related species showed trait values more similar than expected by Brownian evolutionary model traits showed high phylogenetic signal.

To simulate binary and categorical traits, we first simulated a continuous trait and then transformed it into a binary or categorical trait using a simple threshold model (Fritz and Purvis 2010). To do so, the continuous trait was split into percentiles and each interval below or above the thresholds was transformed into a binary (a single threshold at 50th percentile) or categorical trait (two thresholds at 33th and 66th percentiles) (Fig. 1).

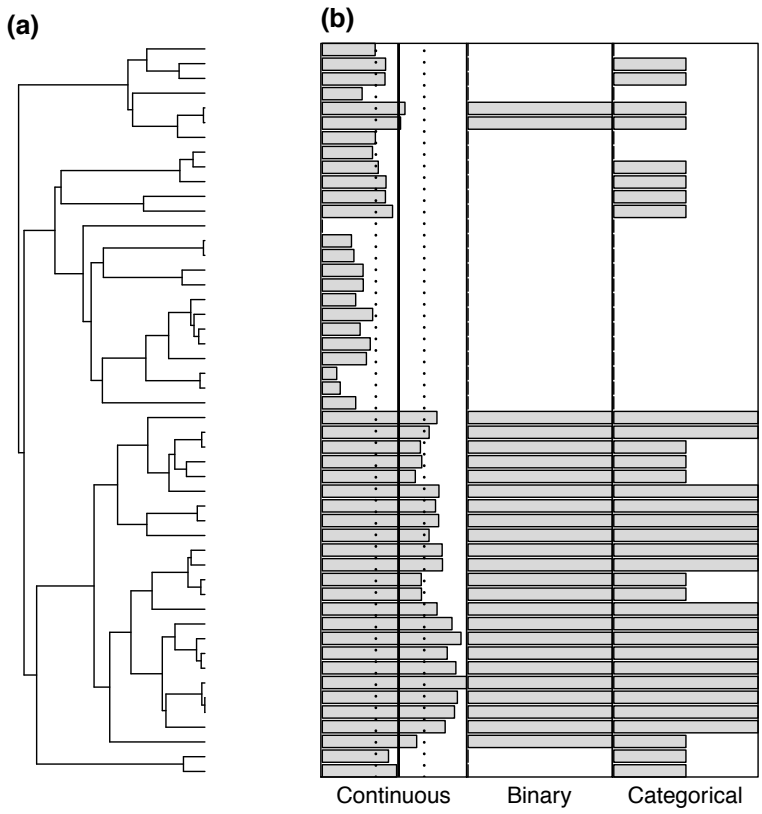

Fig. 1 – Representation of numerical simulation of traits. (a) An ultrametric phylogenetic tree with purely stochastic birth. The traits are simulated based on this phylogeny. (b) Simulation of a continuous, binary and categorical trait under Brownian motion. Each horizontal bar represents the trait value for the species in the phylogenetic tree showed in (a). Binary and categorical traits are transformed using a threshold model. For the binary trait the continuous trait was split in two categories based on a single threshold at 50th percentile (solid line). For the categorical trait the continuous trait was split in three categories based on, two thresholds at 33th and 66th percentiles (dotted lines).

We performed trait simulations considering two different scenarios. The first set of simulations was performed to test the effect of the number of traits used to compute the trait dissimilarity matrix on the phylogenetic signal. For this, we simulated sets with one, two, five or six traits in the matrix in a phylogenetic tree with 100 species. Simulations were performed for each type of trait (continuous, binary or

categorical) separately for two and five traits in the set. The set with six traits was simulated with two traits of each type (binary, continuous and categorical). In the second scenario species pool size varied. This second set of numerical simulations was performed to test influence of the size of the phylogeny on the type I error and statistical power of Mantel tests. We generated phylogenetic trees with 50, 100 and 200 species each. In each phylogenetic tree we generated one continuous trait, one binary trait with a balanced frequency of presences and absences, and a categorical trait with three levels with similar frequencies. Each set of traits was simulated with the same level of one of four levels of phylogenetic signal, as described above. An additional set of simulation was performed with a set of five traits (binary, continuous and categorical) with high phylogenetic signal and one to four traits evolved without phylogenetic signal or under to Brownian motion. This simulation was performed based in a phylogenetic tree with 100 species to access the sensibility of our framework when multivariate traits evolved under a mix of evolutionary models.

We performed each the standard Mantel tests and EM-Mantel tests using Euclidean distances for continuous traits and Gower dissimilarity for categorical, binary or mixed traits (Gower 1971). The evolutionary distance was estimated as the sum of branch length connecting the species in the phylogenetic tree. Additional sets of simulations were performed by square root transforming phylogenetic distances, which is more appropriate for traits evolving under Brownian evolutionary model (Letten and Cornwell 2015). Test significance was accessed using 999 permutations in standard Mantel tests or 999 trait simulations under Brownian motion in EM-Mantel test. In each simulation set, the number and type of traits, and the frequency of occurrence of each level in binary and categorical traits were preserved. The method used to compute the dissimilarity matrix for the simulated trait set was the same of the

observed trait matrix. We recorded the proportion of significant Mantel correlations (i.e., rejection rate) obtained for each test to estimate the statistical power and type I error rate (i.e., rate of false positives). Correlations were considered significant, when p-value were smaller than or equal to 0.05 in the EM-Mantel (one-tailed test), which expresses the probability of the observed Mantel statistic being greater or equal to the expected under the Brownian motion evolutionary model. Rejection rates were recorded from 1,000 repetitions in each parameter combination from standard Mantel and EM-Mantel tests.

For the standard Mantel test type I error was estimated considering the rejection rates of simulations performed using Grafen's $\rho = 0.0001$, where traits did not have any structure related to the phylogenetic tree. For EM-Mantel tests, the type I error was estimated considering the rejection rates of simulations performed using Grafen's $\rho = 0.0001$ (labile traits) and 1.0 (traits generated under Brownian motion). The type I error rate was computed adopting a null hypothesis rejection threshold level of $\alpha = 0.05$. The statistical power was estimated considering the rejection rates of simulations performed using Grafen's $\rho = 1.0$, 2.0 and 5.0 for standard Mantel test and with the rejection rates of simulations performed using Grafen's $\rho = 2.0$ and 5.0 for EM-Mantel tests. All numerical simulations were performed by R (R Development Core Team 2014). The R function implementing the EM-Mantel is provided in the online resource 1.

Results

Our results showed that the power of the standard Mantel test is acceptable for traits evolving under Brownian model and for traits more conserved than expected by Brownian motion. The rejection rates were close to one, irrespective of parameter combination or the type and number of traits (Fig. 2a, 2c and 2e). Further, standard Mantel tests showed acceptable type I error for labile traits, with rejection rates close to the target level of $\alpha = 0.05$ (ranging from 0.04 to 0.063, see also online resource 2 Table S1). Note that for traits evolving under Brownian motion, the standard Mantel test also showed significant results, as expected.

EM-Mantel tests showed are acceptable type I error, with values close to the targeted level of $\alpha = 0.05$ (0.045 to 0.055, see Fig. 2b, d and f, online resource 2 Table S1) for traits evolving by Brownian motion. When traits were less conserved than expected by Brownian motion, the EM-Mantel test did not show any significant result (Fig. 2, online resource 2 Table S1). Further, the statistical power of EM-Mantel tests was weaker under Grafen's $\rho = 2.0$, with rejection rates ranging from 0.226 to 0.339 for one trait, from 0.31 to 0.454 for two traits and from 0.492 to 0.608 for five traits (Fig. 2b, 2d and 2f). Nonetheless, the power of the test was improved for Grafen's $\rho = 5.0$, with rejection rates higher than 0.7 for categorical and continuous traits, with one, two or five traits. For binary traits the test showed smaller power for one trait (rejection rate = 0.423); however, for two or five binary traits the power was improved (rejection rate = 0.646 and 0.797, respectively, see Fig. 2b, d and f, and Table S1). When the set of traits was composed by traits evolving under high levels of phylogenetic signal and by traits without phylogenetic signal or under Brownian motion, the power of EM-Mantel tests were greatly reduced (Online resource 2 Table S2 and S3), e.g. for five continuous traits under Grafen's $\rho = 5.0$ the rejection rate was 0.798 for all traits with high levels of phylogenetic signal and it

decreased to 0.657, 0.529, 0.348 and 0.149 for one, two, three and four traits under Brownian motion respectively (Online resource 2 Table S3). Rejection rates decreased to 0.195, 0.004, 0 and 0 for one, two, three and four traits without any phylogenetic signal respectively (Online resource 2 Table S2).

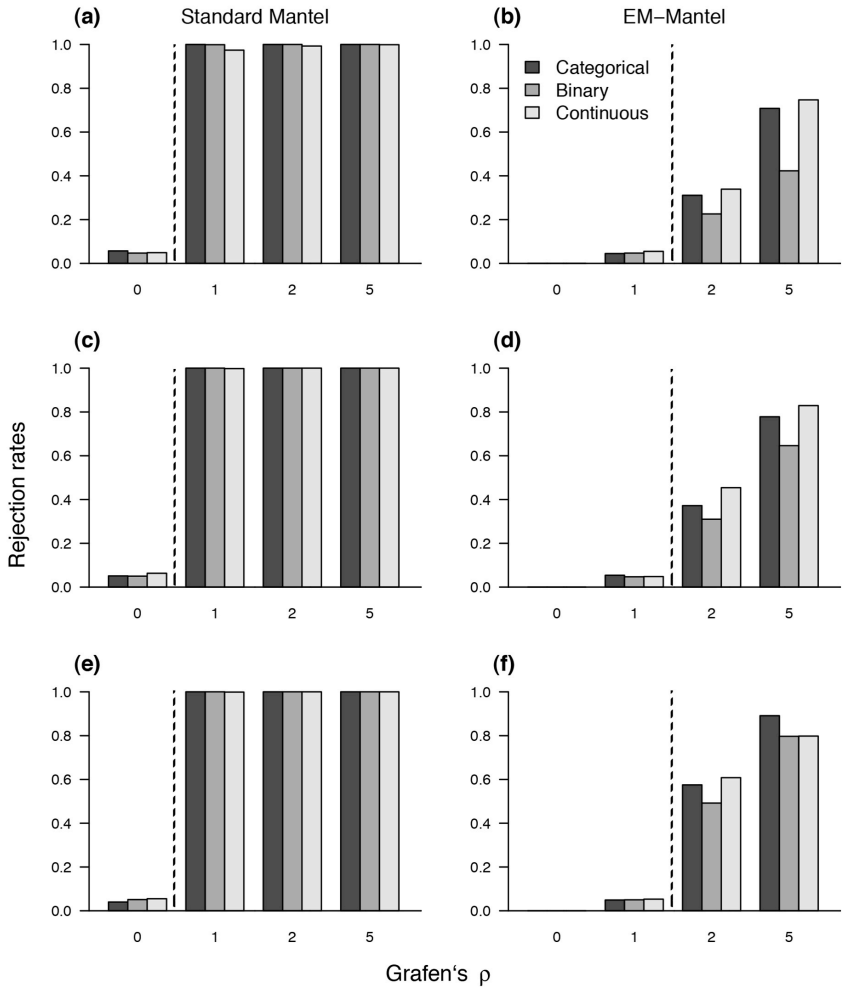

Fig. 2 – Proportion of significant results for standard Mantel and EM-Mantel with simulation of traits under the Brownian motion for different number and type of traits: (a – b) One trait, (c – d) two traits and (e – f) five traits. Left column (a, c and e) shows the power for standard Mantel tests, while right column (b, d and f) shows the power for EM-Mantel tests. Grafen's $\rho = 0.0001$ indicates a trait without phylogenetic signal. Grafen's $\rho = 1.0$ indicates traits evolving under Brownian motion. Grafen's ρ

= 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. Different shades of gray in the bars show the type of traits (continuous, binary, categorical). In all parameter combinations the number of species in the pool was retained constant in 100 species. The dashed line separates the type I error results and the statistical power for each test.

The results for type I error estimates were very similar irrespective of the number of species in the pool, either for standard or EM-Mantel tests (Fig. 3, online resource 2 Table S4). On the other hand, increasing the number of species improved the statistical power of EM-Mantel test. (Fig. 3b, 3d, 3f and online resource 2 Table S4). Furthermore, when traits were of several different types (two continuous, two categorical and two binary), the rejection rate of EM-Mantel test for Grafen's $\rho = 1.0$ was 0.058, showing appropriated type I error, and the rejection rates for Grafen's $\rho = 2.0$ and 5.0, were 0.63 and 0.936, respectively, showing acceptable statistical power (Online resource 2 Table S5).

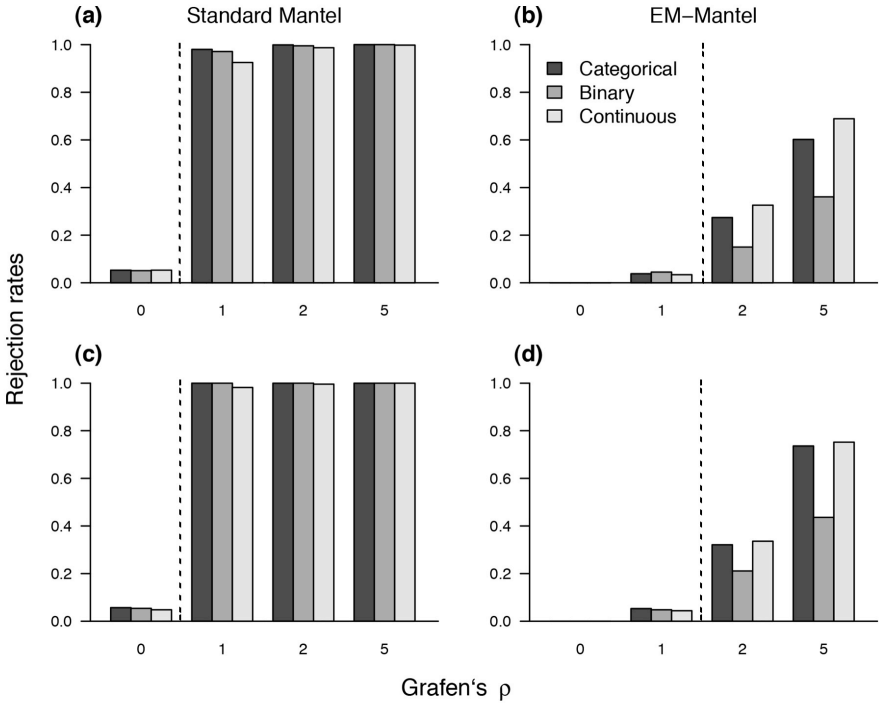

Fig.3 – Proportion of significant results for standard Mantel and EM-Mantel tests with simulation of traits under the Brownian motion for different number of species in the pool: (a – b) 50 species in the pool, (c – d) 200 species in the pool. Left column (a and c) shows the power for standard Mantel test, while right column (b and d) shows the power for EM-Mantel test. Grafen's ρ = 0.0001 indicates a trait without phylogenetic signal. Grafen's ρ = 1.0 indicates traits evolving under Brownian motion. Grafen's ρ = 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. Different shades of gray in the bars show the type of traits (continuous, binary, categorical). The dashed line separates the type I error results and the statistical power for each test.

The results obtained using square root transformed phylogenetic distances were very similar to those without transformation (Online resource 2 Table S6, S7, S8, S9 and S10).

Discussion

The use of Mantel tests for ecological and evolutionary analyses has received critiques in several contexts (Lapointe 1995; Harmon and Glor 2010; Legendre and Fortin 2010; Guillot and Rousset 2012), mainly due to its low power and high rates of type I error. In the phylogenetic context, the Mantel test does not consider any evolutionary model to estimate the phylogenetic signal, which is problematic, since the definition of phylogenetic signal depends on the evolutionary model of interest (Revell et al. 2008, Münkemüller et al. 2015). Lapointe and Garland (2001) offered a solution to the problem of lack of independence among species based on restrictions in phylogenetic permutations, which might be used to assess phylogenetic signal in the traits via the Mantel tests; however, in this case the null hypothesis in Mantel test is the same of the standard Mantel test (i.e, the association between phylogenetic distances and trait dissimilarities among species are not different from zero). Standard Mantel tests with restrictions in the phylogenetic permutations do not explicitly consider any evolutionary model to estimate the phylogenetic signal. In this way, the proposed EM-Mantel is the first attempt to incorporate explicitly an evolutionary model in the Mantel test. Our results clearly show that the phylogenetic signal assessment via the standard Mantel test has an appropriate type I error for labile traits, confirming the results found by Harmon and Glor (2010) and Seger et al. (2013); however, when traits evolve under a Brownian evolutionary model, the proportion of significant results was close to one, indicating that the standard Mantel test does not distinguishes the phylogenetic signal generated by neutral evolution from a trait conservatism higher than expected by neutral evolution. Therefore, using the standard Mantel test for assessing phylogenetic signal in traits should be avoided.

The Brownian motion is one of the simplest evolutionary models for continuous traits (Felsenstein 1985), where the differences between species are merely accumulated over evolutionary time (Freckleton and Harvey 2006) without strong stabilizing selection, strong divergent selection or changes in the evolutionary rate along the phylogenetic tree. The understanding of the association between pairwise phylogenetic distances and trait dissimilarities among species under a very simple model of character evolution is a minimal procedure for assessing phylogenetic niche conservatism (Losos 2008; but see Pillar and Duarte 2010) or for explaining functional and phylogenetic community patterns (Webb at al. 2002). Our results demonstrate that the EM-Mantel have appropriate type I error for traits evolving under Brownian motion and acceptable statistical power for highly phylogenetically conserved traits, especially for categorical and continuous traits. Therefore, the EM-Mantel test may be a useful alternative of estimating trait conservatism for sets of categorical traits. For binary traits we found a smaller power of the EM-Mantel tests in comparison to other trait types. This problem is particularly critical when the analysis is performed using a single binary trait. In this case, the use of the D statistic developed by Fritz and Purvis (2010) might represent a better solution for estimating phylogenetic signal, although we have not evaluated its statistical performance. On the other hand, the statistical power of EM-Mantel tests is enhanced when multiple binary traits are analyzed. Thus, whereas the D statistic is effective only if restricted to single traits, EM-Mantel tests are useful to estimate phylogenetic signal for multiple binary traits. Binary and categorical traits are very commonly found in trait datasets, when important characteristics expressing life history aspects or ecological requirements of species, such as color, type of reproduction, diet or degree of endangerment cannot be described as continuous

variables. Thus, the EM-Mantel test enables a robust estimate of trait conservatism for non-continuous traits or mixed sets of traits, enabling elements for inference of niche conservatism and analyses of trait-mediated assembly patterns even where quantitative assessment of ecological traits of species is not possible.

Hardy and Pavoine (2012) showed that the statistical power of Mantel test depends on the metric used to define pairwise trait and phylogenetic distances between species. For continuous traits, the simplest resemblance measure is the Euclidean distance (Legendre and Legendre 2012). However, for mixed traits, the most flexible dissimilarity index is the Gower index, which allows ecologists to work with continuous, categorical, categorical ordered and binary traits, including asymmetric binary variables and weighted variables (Gower 1971; Podani 1999; Legendre and Legendre 2012). Although we do not compare dissimilarity indexes here, other dissimilarly indices might be used to perform the EM-Mantel tests. For instance, under Brownian model of trait evolution, it is expected that the phenotypic variation between any two taxa is more appropriately represented as a linear function of the square root of time (Letten and Cornwell 2015). Thus, a simple square root transformation of the trait-based Euclidean distances or Gower dissimilarities among species might increase the robustness of EM-Mantel tests. Furthermore, alternative phylogenetic distances, such as node counting or Abouheif proximities (Abouheif 1999) might be used to assess the phylogenetic signal via EM-Mantel tests. Other models of evolution with more complex properties, such as Ornstein–Uhlenbeck model (Hansen 1997), Early Burst model (Harmon et al. 2010), Ornstein–Uhlenbeck model with different selective optima (Butler and King 2010), might be easily incorporated in the EM-Mantel test in order to create confidence intervals for trait-phylogeny correlations using the same species pool. Furthermore confidence intervals

of Mantel statistic can be estimated for two or more models of evolution with the goal of contrasting and testing whether a set of traits follows a particular model of evolution (Boettiger et al. 2012). Of course, using more complex measures of phylogenetic and/or functional distances, or even alternative evolutionary models would rely on sound biological reasoning.

References

Abouheif E (1999) A method for testing the assumption of phylogenetic independence in comparative data. Evol Ecol Res 1:895–909.

Ackerly D (2009) Conservatism and diversification of plant functional traits: Evolutionary rates versus phylogenetic signal. Proc Natl Acad Sci U S A 106 Suppl:19699–706. doi: 10.1073/pnas.0901635106

Adams D (2014) A Generalized K Statistic for Estimating Phylogenetic Signal from Shape and other High-dimentional Multivariate Data. Syst Biol 63:685–697. doi: 10.1093/sysbio/syu030

Blomberg SP, Garland T, Ives AR (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution (N Y) 57:717–745. doi: 10.1111/j.0014-3820.2003.tb00285.x

Boettiger C, Coop G, Ralph P (2012) Is your phylogeny informative? Measuring the power of comparative methods. Evolution (N Y) 66:2240–2251. doi: 10.1111/j.1558-5646.2011.01574.x

Butler MA, King AA (2004) Phylogenetic Comparative Analysis: A Modeling

Approach for Adaptive Evolution. Am Nat 164:683–695. doi: 10.1086/426002

Cooper N, Jetz W, Freckleton RP (2010) Phylogenetic comparative approaches for studying niche conservatism. J Evol Biol 23:2529–2539. doi: 10.1111/j.1420-9101.2010.02144.x

Diniz-Filho JAF, de Sant'Ana CER, Bini LM (1998) An Eigenvector Method for Estimating Phylogenetic Inertia. Evolution (N Y) 52:1247–1262. doi: 10.2307/2411294

Diniz-Filho JAF, Rangel TF, Santos T, Bini LM (2012) Exploring patterns of interspecific variation in quantitative traits using sequential phylogenetic eigenvector regressions. Evolution (N Y) 66:1079–1090. doi: 10.1111/j.1558-5646.2011.01499.x

Diniz-Filho JAF, Bini LM, Sakamoto M, Brusatte, SL (2014) Phylogenetic Eigenvector Regression in Paleobiology. Rev. bras. paleontol. 17(2):107–122. doi: 10.4072/rbp.2014.2.01

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

Freckleton RP, Harvey PH (2006) Detecting non-Brownian trait evolution in adaptive radiations. PLoS Biol 4:e373. doi: 10.1371/journal.pbio.0040373

Fritz SA, Purvis A (2010) Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. Conserv Biol 24:1042–51. doi: 10.1111/j.1523-1739.2010.01455.x

Gerhold P, Cahill JF, Winter M, et al. (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). Funct Ecol 29:600–614.

doi: 10.1111/1365-2435.12425

Gower JC (1971) A General Coefficient of Similarity and Some of Its Properties. Biometrics 27:857–871. doi: 10.2307/2528823

Grafen A (1989) The phylogenetic regression. Philos Trans R Soc Lond B Biol Sci 326:119–157.

Guillot G, Rousset F (2012) Dismantling the Mantel tests. Methods Ecol Evol 4:336–344. doi: 10.1111/2041-210x.12018

Hansen TF (1997) Stabilizing Selection and the Comparative Analysis of Adaptation. Evolution (N Y) 51:1341–1351. doi: 10.2307/2411186

Hardy OJ, Pavoine S (2012) Assessing phylogenetic signal with measurement error: a comparison of Mantel tests, Blomberg et al.'s K, and phylogenetic distograms. Evolution 66:2614–21. doi: 10.1111/j.1558-5646.2012.01623.x

Harmon LJ, Glor RE (2010) Poor statistical performance of the Mantel test in phylogenetic comparative analyses. Evolution 64:2173–8. doi: 10.1111/j.1558-5646.2010.00973.x

Harmon LJ, Losos JB, Jonathan Davies T, et al. (2010) Early bursts of body size and shape evolution are rare in comparative data. Evolution (N Y) 64:2385–2396. doi: 10.1111/j.1558-5646.2010.01025.x

Harmon LJ, Weir JT, Brock CD, et al. (2008) GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–31. doi: 10.1093/bioinformatics/btm538

Harvey PH, Pagel MD (1991) The comparative method in evolutionary biology. Oxford Univ. Press, Oxford, U.K.

Lapointe F-J, Garland TJ (2001) A Generalized Permutation Model for the Analysis of Cross-Species Data. J Classif 18:109–127. doi: 10.1007/s00357-001-0007-0

Lapointe F-J, Legendre P (1995) Comparison Tests for Dendrograms - a Comparative Evaluation. J Classif 12:265–282.

Lefcheck JS, Bastazini V a. G, Griffin JN (2015) Choosing and using multiple traits in functional diversity research. Environ Conserv 42:104–107. doi: 10.1017/S0376892914000307

Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Mol Ecol Resour 10:831–844. doi: 10.1111/j.1755-0998.2010.02866.x

Legendre P, Legendre L (2012) Numerical ecology. 3nd ed. Elsevier.

Letten AD, Cornwell WK (2015) Trees, branches and (square) roots: why evolutionary relatedness is not linearly related to functional distance. Methods Ecol Evol 6:439–444. doi: 10.1111/2041-210X.12237

Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecol Lett 11:995–1007. doi: 10.1111/j.1461-0248.2008.01229.x

Losos JB, Leal M, Glor RE, et al. (2003) Niche lability in the evolution of a Caribbean lizard community. Nature 424:542–5. doi: 10.1038/nature01814

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220.

Mooers A, Gascuel O, Stadler T, et al. (2012) Branch lengths on birth-death trees and the expected loss of phylogenetic diversity. Syst Biol 61:195–203. doi: 10.1093/sysbio/syr090

Mouquet N, Devictor V, Meynard CN, et al. (2012) Ecophylogenetics: advances and perspectives. Biol Rev Camb Philos Soc 87:769–85. doi: 10.1111/j.1469-185X.2012.00224.x

Münkemüller T, Boucher FC, Thuiller W, Lavergne S (2015) Phylogenetic niche conservatism - common pitfalls and ways forward. Funct Ecol 29:627–639. doi: 10.1111/1365-2435.12388

Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401:877–84. doi: 10.1038/44766

Paradis E (2004) APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20:289–290. doi: 10.1093/bioinformatics/btg412

Paradis E (2012) Analysis of Phylogenetics and Evolution with R. 2nd ed. Springer-Verlag, Heidelberg.

Pillar VD, Duarte LS (2010) A framework for metacommunity analysis of phylogenetic structure. Ecol Lett 13:587–596. doi: 10.1111/j.1461-0248.2010.01456.x

Podani J (1999) Extending Gower's General Coefficient of Similarity to Ordinal Characters. Taxon 48:331–340. doi: 10.2307/1224438

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

Revell LJ, Harmon LJ, Collar DC (2008) Phylogenetic signal, evolutionary process, and rate. Syst Biol 57:591–601. doi: 10.1080/10635150802302427

Seger GDS, Duarte LDS, Debastiani VJ, et al. (2013) Discriminating the effects of phylogenetic hypothesis, tree resolution and clade age estimates on phylogenetic signal measurements. Plant Biol 15:858–67. doi: 10.1111/j.1438-8677.2012.00699.x

Webb CO, Ackerly DD, Mcpeek MA, Donoghue MJ (2002) Phylogenies and community ecology. Annu Rev Ecol Syst 33:475–505. doi: 10.1146/annurev.ecolsys.33.010802.150448

Wiens JJ, Graham CH (2005) Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. Annu Rev Ecol Evol Syst 36:519–539. doi: 10.1146/annurev.ecolsys.36.102803.095431

Supplemental Material

Online resource 1

R function for estimating phylogenetic signal in traits using EM-Mantel test.

```
## Required packages

require(vegan)
require(FD)

## Description

# Calculate the phylogenetic signal based in Mantel test, incorporating the Brownian motion evolutionary model.

## Arguments

# tree = Phylogenetic tree, as phylo object.
# traits = Matrix or data frame containing the traits data, with traits as columns and species as rows. Traits can be numeric, ordered, or factor, as the required by the gowdis function. (Symmetric or asymmetric binary variables should be numeric and only contain 0 and 1. character variables will be converted to factor).
# runs = Number of permutations in assessing significance.
# euclidean = Logical argument (TRUE or FALSE) to specify if use transformation of euclidean properties in pairwise trait dissimilarities (Default euclidean = TRUE).
# sqrtPhylo = Logical argument (TRUE or FALSE) to specify if use square root transformation of phylogenetic distance (Default sqrtPhylo = FALSE).
# checkdata = Logical argument (TRUE or FALSE) to check if species sequence in the trait data follows the same order as in phylogenetic tree (Default checkdata = TRUE).
# ... = Parameters for gowdis function.

## Value

# perm.NULL = A vector of permuted Mantel statistic under the null hypothesis that the distances between both matrices are not related.
# perm.BM = A vector of permuted Mantel statistic under the null hypothesis that the traits evolve under a Brownian motion evolutionary model.
# r.Mantel = The Mantel observed statistic.
# p.NULL = The p value from no phylogenetic structure (standard p value in Mantel test).
# p.BM = The p value under simulation of traits from Brownian phylogenetic structure.


EM.mantel<-function(tree, traits, runs = 999, euclidean= TRUE, sqrtPhylo=FALSE, checkdata = TRUE, ...){
 phylo.dist<-cophenetic(tree)
 if(sqrtPhylo){
        phylo.dist<-sqrt(phylo.dist)
 }
 if(checkdata){
        if(is.null(tree$tip.label)){
         stop("\n Error in tip labels of tree\n")
        }
        if(is.null(rownames(traits))){
                stop("\n Error in row names of traits\n")
```

```
        }
        match.names <- match(rownames(traits),rownames(phylo.dist))
        if(sum(is.na(match.names)) > 0){
                stop("\n There are species from traits data that are not on phylogenetic tree\n")
        }
        phylo.dist <- phylo.dist[match.names, match.names]
}
if(length(tree$tip.label) > dim(traits)[1]){
        warning("Tree have more species that species in traits data")
}
if(dim(phylo.dist)[1] != dim(traits)[1] & checkdata == FALSE){
        stop("\n Different number of species in tree and in traits data, use checkdata = TRUE\n")
}
gow.dist<-gowdis(traits, ...)
if(euclidean){
        gow.sim<-1-gow.dist
        gow.dist<-sqrt(1-gow.sim)
}
traits.attr<-attr(gow.dist, "Types", exact = TRUE)
res.mantel<-mantel(phylo.dist,gow.dist,permutations=runs)
res.BM<-matrix(NA,runs,1)
for(k in 1:runs){
        traits_sim<-matrix(NA,length(tree$tip.label),dim(traits)[2])
        rownames(traits_sim)<-tree$tip.label
        for(i in 1:dim(traits)[2]){
                traits_sim[,i]<-rTraitCont(tree,model="BM")
        }
        traits_sim<-decostand(traits_sim,method="standardize",MARGIN=2)
        traits_sim<-as.data.frame(traits_sim)
        for(i in 1:dim(traits)[2]){
                if(traits.attr[i] == "B" | traits.attr[i] == "A"){
                        probs<-sum(traits[,i])/dim(traits)[1]
                        threshold<-quantile(traits_sim[,i],probs=1-probs)
                        traits_sim[,i]<-ifelse(traits_sim[,i]>=threshold,1,0)
                }
                if(traits.attr[i] == "N" | traits.attr[i] == "O"){
                        n.levels<-length(levels(traits[,i]))
                        traits.levels<-levels(traits[,i])
                        probs<-cumsum(table(traits[,i]))/sum(table(traits[,i]))
                        probs<-probs[1:(n.levels-1)]
                        threshold<-quantile(traits_sim[,i],probs=probs)
                        threshold<-c(min(traits_sim[,i]),threshold,max(traits_sim[,i]))
                        temp<-matrix(NA,length(traits_sim[,i]),1)
                        for(j in 1:n.levels){
                                if(j < n.levels){
                                        temp[1:length(traits_sim[,i]),1]<-ifelse(traits_sim[,i]>=threshold[j]
& traits_sim[,i]<threshold[j+1], traits.levels[j],temp)
                                }
                                if(j == n.levels){
                                        temp[1:length(traits_sim[,i]),1]<-ifelse(traits_sim[,i]>=threshold[j]
& traits_sim[,i]<=threshold[j+1], traits.levels[j],temp)
                                }
                        }
                        traits_sim[,i]<-as.factor(temp)
                        if(traits.attr[i] == "O"){
                                traits_sim[,i]<-ordered(temp,levels=levels(traits[,i]))
                        }
                }
        }
        if(checkdata == TRUE){
```

```
            match.names <- match(rownames(traits),rownames(traits_sim))
            traits_sim<-traits_sim[match.names,,drop=FALSE]
        }
        gow.dist.BM<-gowdis(traits_sim, ...)
        if(euclidean){
            gow.sim.BM<-1-gow.dist.BM
            gow.dist.BM<-sqrt(1-gow.sim.BM)
        }
        res.mantel.BM<-mantel(phylo.dist,gow.dist.BM,permutations=0)
        res.BM[k,1]<-res.mantel.BM$statistic
 }
 p.BM<-(sum(ifelse(res.BM[,1]>=res.mantel$statistic,1,0))+1)/(runs+1)
 p.NULL<-res.mantel$signif
 r.Mantel<-res.mantel$statistic
 RES<-
 list(perm.NULL=res.mantel$perm,perm.BM=res.BM[,1],r.Mantel=r.Mantel,p.NULL=p.NULL,p.BM
 =p.BM)
return(RES)
}

## Examples

require(geiger)
tree<-sim.bdtree(b=0.1,d=0,stop="taxa",n=100,extinct=FALSE)
trait<-matrix(rTraitCont(compute.brlen(tree,power=5),model="BM"),100,1)
rownames(trait)<-tree$tip.label
EM.mantel(tree,trait,runs=99)
```

Online resource 2

Tables with summary statistics (mean ± SD) and rejection rates for Mantel and EM-Mantel.

Table S1. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for set with different number and type of traits. Grafen's ρ = 0.0001 indicates a trait without phylogenetic signal. Grafen's ρ = 1.0 indicates traits evolving under Brownian motion. Grafen's ρ = 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. In all parameter combinations the number of species in the pool was retained constant in 100 species. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of $\alpha = 0.05$.

Number of traits	Grafen's ρ	Type of trait	Mean Mantel statistic (SD)	Rejection rates	
				Mantel	EM-Mantel
1	0	Categorical	0.0005 (0.0131)	0.057	0
		Binary	-0.0006 (0.0125)	0.047	0
		Continuous	0.0005 (0.0289)	0.049	0
	1	Categorical	0.2673 (0.1227)	1	0.045
		Binary	0.2632 (0.1381)	0.999	0.047
		Continuous	0.3469 (0.1699)	0.974	0.055
	2	Categorical	0.3983 (0.1291)	1	0.311
		Binary	0.3610 (0.1565)	1	0.226
		Continuous	0.4994 (0.1813)	0.993	0.339
	5	Categorical	0.4934 (0.1131)	1	0.708
		Binary	0.4317 (0.1432)	1	0.423
		Continuous	0.6552 (0.1500)	0.999	0.747
2	0	Binary	0.0000 (0.0130)	0.05	0
		Continuous	0.0002 (0.0308)	0.063	0
		Categorical	0.0004 (0.0168)	0.051	0
	1	Binary	0.3488 (0.1228)	1	0.047

		Continuous	0.4603 (0.1439)	0.998	0.048
		Categorical	0.3224 (0.1074)	1	0.054
	2	Binary	0.4586 (0.1309)	1	0.31
		Continuous	0.6421 (0.1313)	1	0.454
		Categorical	0.4475 (0.1188)	1	0.372
	5	Binary	0.5265 (0.1288)	1	0.646
		Continuous	0.7532 (0.1071)	1	0.829
		Categorical	0.5470 (0.1076)	1	0.778
5	0	Binary	-0.0001 (0.0131)	0.051	0
		Continuous	-0.0009 (0.0323)	0.055	0
		Categorical	-0.0006 (0.0174)	0.04	0
	1	Binary	0.4980 (0.1077)	1	0.05
		Continuous	0.6287 (0.1106)	0.999	0.053
		Categorical	0.4293 (0.0980)	1	0.049
	2	Binary	0.5974 (0.1139)	1	0.492
		Continuous	0.7762 (0.0845)	1	0.608
		Categorical	0.5638 (0.1000)	1	0.575
	5	Binary	0.6376 (0.1138)	1	0.797
		Continuous	0.8229 (0.0721)	1	0.798
		Categorical	0.6421 (0.1062)	1	0.891

Table S2. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for set with five traits evolving under different levels of conservatism. Grafen's ρ = 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. Traits without phylogenetic signal was simulated with Grafen's ρ = 0.0001. In all parameter combinations the number of species in the pool was retained constant in 100 species. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of α = 0.05.

Grafen parameter	Type of trait	Number of traits without signal	Mean Mantel statistic (SD)	Proportion of significant results	
				Mantel	EM-Mantel
2	Categorical	1	0.4676 (0.1071)	1	0.155
		2	0.3700 (0.0973)	1	0.006
		3	0.2619 (0.0859)	0.997	0
		4	0.1357 (0.0616)	0.964	0
	Binary	1	0.5245 (0.1084)	1	0.062
		2	0.4208 (0.1029)	1	0
		3	0.3030 (0.0902)	1	0
		4	0.1633 (0.0685)	0.999	0
	Continuous	1	0.6622 (0.0907)	1	0.043
		2	0.5269 (0.0995)	0.999	0
		3	0.3657 (0.0965)	0.995	0
		4	0.1850 (0.0766)	0.953	0
5	Categorical	1	0.5568 (0.1114)	1	0.528
		2	0.4523 (0.1008)	0.999	0.077
		3	0.3329 (0.0871)	0.999	0
		4	0.1724 (0.0616)	0.988	0
	Binary	1	0.5779 (0.1143)	1	0.284
		2	0.4882 (0.1086)	1	0
		3	0.3608 (0.0965)	1	0
		4	0.1942 (0.0661)	1	0
	Continuous	1	0.7170 (0.0826)	1	0.195

2	0.5992 (0.0892)	1	0.004
3	0.4549 (0.0842)	1	0
4	0.2409 (0.0748)	0.979	0

Table S3. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for set with five traits evolving under different levels of conservatism. Grafen's ρ = 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. Traits under Brownian motion was simulated with Grafen's ρ = 1. In all parameter combinations the number of species in the pool was retained constant in 100 species. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of $\alpha = 0.05$.

Grafen parameter	Type of trait	Number of traits under Brownian motion	Mean Mantel statistic (SD)	Proportion of significant results	
				Mantel	EM-Mantel
2	Categorical	1	0.5306 (0.1049)	1	0.412
		2	0.5113 (0.1074)	1	0.313
		3	0.4816 (0.1001)	1	0.163
		4	0.4587 (0.1012)	1	0.102
	Binary	1	0.5801 (0.1069)	1	0.358
		2	0.5568 (0.1119)	1	0.259
		3	0.5363 (0.1081)	1	0.174
		4	0.5187 (0.1088)	1	0.081
	Continuous	1	0.7546 (0.0853)	1	0.471
		2	0.7299 (0.0910)	1	0.331
		3	0.6974 (0.0924)	1	0.182
		4	0.6649 (0.1059)	1	0.101
5	Categorical	1	0.5980 (0.1096)	1	0.765
		2	0.5685 (0.1016)	1	0.596
		3	0.5279 (0.1059)	1	0.396
		4	0.4899 (0.1033)	1	0.184
	Binary	1	0.6202 (0.1160)	1	0.679
		2	0.6037 (0.1124)	1	0.522
		3	0.5686 (0.1116)	1	0.301
		4	0.5339 (0.1073)	1	0.138
	Continuous	1	0.7923 (0.0794)	1	0.657
		2	0.7706 (0.0806)	1	0.529

3	0.7395 (0.0879)	1	0.348
4	0.6902 (0.0918)	1	0.149

Table S4. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for different number of species in the pool. Grafen's $\rho = 0.0001$ indicates a trait without phylogenetic signal. Grafen's $\rho = 1.0$ indicates traits evolving under Brownian motion. Grafen's $\rho = 2.0$ and 5.0 indicate traits evolving less that expected by Brownian motion. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of $\alpha = 0.05$.

Number of species	Grafen's ρ	Type of trait	Mean Mantel statistic (SD)	Rejection rates	
				Mantel	EM-Mantel
50	0	Categorical	0.0001 (0.0263)	0.053	0
		Binary	0.0007 (0.0275)	0.051	0
		Continuous	-0.0008 (0.0488)	0.053	0
	1	Categorical	0.2684 (0.1311)	0.98	0.038
		Binary	0.2588 (0.1535)	0.971	0.045
		Continuous	0.3327 (0.1767)	0.925	0.034
	2	Categorical	0.3806 (0.1443)	0.999	0.274
		Binary	0.3440 (0.1607)	0.995	0.15
		Continuous	0.5033 (0.1872)	0.987	0.326
	5	Categorical	0.4775 (0.1243)	1	0.602
		Binary	0.4311 (0.1568)	1	0.361
		Continuous	0.6508 (0.1661)	0.998	0.689
200	0	Categorical	-0.0001 (0.0066)	0.057	0
		Binary	-0.0002 (0.0065)	0.054	0
		Continuous	0.0000 (0.0201)	0.048	0
	1	Categorical	0.2794 (0.1184)	1	0.053
		Binary	0.2616 (0.1335)	1	0.048
		Continuous	0.3480 (0.1608)	0.982	0.044
	2	Categorical	0.4059 (0.1197)	1	0.321
		Binary	0.3527 (0.1442)	1	0.211
		Continuous	0.5062 (0.1702)	0.996	0.336
	5	Categorical	0.4989 (0.1048)	1	0.736

	Binary	0.4334 (0.1416)	1	0.436
	Continuous	0.6507 (0.1444)	1	0.752

Table S5. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for set with six mixed traits, two traits of each type (binary, continuous and categorical). Grafen's ρ = 0.0001 indicates a trait without phylogenetic signal. Grafen's ρ = 1.0 indicates traits evolving under Brownian motion. Grafen's ρ = 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. In all parameter combinations the number of species in the pool was retained constant in 100 species. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of α = 0.05.

Grafen's ρ	Mean Mantel statistic (SD)	Rejection rates	
		Mantel	EM-Mantel
0	0.0002 (0.0161)	0.056	0
1	0.4971 (0.1063)	1	0.058
2	0.6303 (0.0984)	1	0.63
5	0.6969 (0.1045)	1	0.936

Table S6. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for set with different number and type of traits. Standard Mantel and EM-Mantel using square root transformation in the phylogenetic distance. Grafen's ρ = 0.0001 indicates a trait without phylogenetic signal. Grafen's ρ = 1.0 indicates traits evolving under Brownian motion. Grafen's ρ = 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. In all parameter combinations the number of species in the pool was retained constant in 100 species. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of α = 0.05.

Number of traits	Grafen's ρ	Type of trait	Mean Mantel statistic (SD)	Rejection rates	
				Mantel	EM-Mantel
1	0	Categorical	0.0004 (0.0137)	0.051	0
		Binary	-0.0001 (0.0127)	0.042	0
		Continuous	0.0007 (0.0268)	0.059	0
	1	Categorical	0.2786 (0.1133)	1	0.055
		Binary	0.2652 (0.1329)	1	0.063
		Continuous	0.3496 (0.1516)	0.993	0.063
	2	Categorical	0.3961 (0.1121)	1	0.318
		Binary	0.3455 (0.1257)	1	0.217
		Continuous	0.4892 (0.1498)	0.998	0.353
	5	Categorical	0.4844 (0.0977)	1	0.711
		Binary	0.4138 (0.1208)	1	0.418
		Continuous	0.6132 (0.1279)	1	0.749
2	0	Binary	-0.0001 (0.0135)	0.05	0
		Continuous	0.0012 (0.0272)	0.052	0
		Categorical	-0.0005 (0.0162)	0.047	0
	1	Binary	0.3482 (0.1085)	1	0.059
		Continuous	0.4615 (0.1265)	0.998	0.053
		Categorical	0.3244 (0.0999)	1	0.04
	2	Binary	0.4482 (0.1156)	1	0.316
		Continuous	0.6171 (0.1213)	1	0.474

		Categorical	0.4574 (0.0990)	1	0.438
	5	Binary	0.5128 (0.1122)	1	0.654
		Continuous	0.7085 (0.1002)	1	0.816
		Categorical	0.5399 (0.0987)	1	0.803
5	0	Binary	0.0004 (0.0136)	0.052	0
		Continuous	0.0007 (0.0285)	0.042	0
		Categorical	0.0011 (0.0175)	0.062	0
	1	Binary	0.4934 (0.0936)	1	0.059
		Continuous	0.6299 (0.0884)	1	0.052
		Categorical	0.4380 (0.0864)	1	0.048
	2	Binary	0.5815 (0.0950)	1	0.526
		Continuous	0.7532 (0.0729)	1	0.613
		Categorical	0.5668 (0.0918)	1	0.641
	5	Binary	0.6161 (0.1012)	1	0.718
		Continuous	0.7635 (0.0825)	1	0.654
		Categorical	0.6344 (0.0886)	1	0.919

Table S7. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for set with five traits evolving under different levels of conservatism. Standard Mantel and EM-Mantel using square root transformation in the phylogenetic distance. Grafen's ρ = 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. Traits without phylogenetic signal was simulated with Grafen's ρ = 0.0001. In all parameter combinations the number of species in the pool was retained constant in 100 species. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of α = 0.05.

Grafen parameter	Type of trait	Number of traits without signal	Mean Mantel statistic (SD)	Proportion of significant results	
				Mantel	EM-Mantel
2	Categorical	1	0.4667 (0.0881)	1	0.137
		2	0.3751 (0.0844)	1	0.004
		3	0.2623 (0.0737)	0.999	0
		4	0.1379 (0.0546)	0.981	0
	Binary	1	0.5111 (0.0912)	1	0.048
		2	0.4170 (0.0905)	1	0
		3	0.2978 (0.0788)	1	0
		4	0.1550 (0.0598)	0.999	0
	Continuous	1	0.6324 (0.0837)	1	0.012
		2	0.5042 (0.0815)	1	0
		3	0.3448 (0.0787)	0.999	0
		4	0.1800 (0.0676)	0.967	0
5	Categorical	1	0.5489 (0.0901)	1	0.551
		2	0.4519 (0.0876)	1	0.053
		3	0.3302 (0.0778)	1	0
		4	0.1755 (0.0558)	0.995	0
	Binary	1	0.5549 (0.0951)	1	0.244
		2	0.4728 (0.0961)	1	0
		3	0.3454 (0.0838)	1	0
		4	0.1872 (0.0574)	1	0

Continuous	1	0.6645 (0.0866)	1	0.067
	2	0.5632 (0.0816)	1	0
	3	0.4176 (0.0733)	1	0
	4	0.2243 (0.0648)	0.991	0

Table S8. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for set with five traits evolving under different levels of conservatism. Standard Mantel and EM-Mantel using square root transformation in the phylogenetic distance. Grafen's ρ = 2.0 and 5.0 indicate traits evolving less that expected by Brownian motion. Traits under Brownian motion. was simulated with Grafen's ρ = 1. In all parameter combinations the number of species in the pool was retained constant in 100 species. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of α = 0.05.

Grafen parameter	Type of trait	Number of traits under Brownian motion	Mean Mantel statistic (SD)	Proportion of significant results	
				Mantel	EM-Mantel
2	Categorical	1	0.5397 (0.0892)	1	0.475
		2	0.5203 (0.0894)	1	0.365
		3	0.4925 (0.0887)	1	0.222
		4	0.4679 (0.0867)	1	0.125
	Binary	1	0.5618 (0.0972)	1	0.371
		2	0.5504 (0.0882)	1	0.289
		3	0.5320 (0.0981)	1	0.187
		4	0.5062 (0.0906)	1	0.086
	Continuous	1	0.7319 (0.0790)	1	0.463
		2	0.7053 (0.0810)	1	0.334
		3	0.6836 (0.0836)	1	0.219
		4	0.6553 (0.0859)	1	0.106
5	Categorical	1	0.6014 (0.0923)	1	0.78
		2	0.5736 (0.0925)	1	0.69
		3	0.5338 (0.0875)	1	0.431
		4	0.4874 (0.0845)	1	0.182
	Binary	1	0.5999 (0.0927)	1	0.682
		2	0.5844 (0.0946)	1	0.504
		3	0.5562 (0.0922)	1	0.277

	4	0.5310 (0.0913)	1	0.137
Continuous	1	0.7496 (0.0765)	1	0.563
	2	0.7344 (0.0773)	1	0.479
	3	0.7136 (0.0744)	1	0.304
	4	0.6743 (0.0801)	1	0.153

Table S9. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for different number of species in the pool. Standard Mantel and EM-Mantel using square root transformation in the phylogenetic distance. Grafen's $\rho = 0.0001$ indicates a trait without phylogenetic signal. Grafen's $\rho = 1.0$ indicates traits evolving under Brownian motion. Grafen's $\rho = 2.0$ and 5.0 indicate traits evolving less that expected by Brownian motion. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of $\alpha = 0.05$.

Number of species	Grafen's ρ	Type of trait	Mean Mantel statistic (SD)	Rejection rates	
				Mantel	EM-Mantel
50	0	Categorical	-0.0003 (0.0264)	0.051	0
		Binary	-0.0009 (0.0258)	0.051	0
		Continuous	-0.0008 (0.0457)	0.051	0
	1	Categorical	0.2651 (0.1208)	0.991	0.049
		Binary	0.2449 (0.1296)	0.987	0.044
		Continuous	0.3264 (0.1569)	0.952	0.042
	2	Categorical	0.3843 (0.1251)	0.999	0.289
		Binary	0.3438 (0.1380)	1	0.179
		Continuous	0.4870 (0.1631)	0.996	0.335
	5	Categorical	0.4831 (0.1082)	1	0.674
		Binary	0.4164 (0.1377)	1	0.35
		Continuous	0.6243 (0.1369)	0.999	0.747
200	0	Categorical	0.0003 (0.0069)	0.059	0
		Binary	0.0002 (0.0069)	0.063	0
		Continuous	-0.0005 (0.0181)	0.05	0
	1	Categorical	0.2835 (0.1076)	1	0.053
		Binary	0.2639 (0.1164)	1	0.05
		Continuous	0.3432 (0.1407)	0.996	0.05
	2	Categorical	0.4023 (0.1077)	1	0.343
		Binary	0.3518 (0.1216)	1	0.224
		Continuous	0.4892 (0.1423)	1	0.369

5	Categorical	0.4904 (0.0943)	1	0.743
	Binary	0.4184 (0.1200)	1	0.463
	Continuous	0.6097 (0.1325)	1	0.743

Table S10. Mean Mantel statistic, standard deviation (SD), and the rejection rates for standard Mantel and EM-Mantel for set with six mixed traits, two traits of each type (binary, continuous and categorical). Standard Mantel and EM-Mantel using square root transformation in the phylogenetic distance. Grafen's $\rho = 0.0001$ indicates a trait without phylogenetic signal. Grafen's $\rho = 1.0$ indicates traits evolving under Brownian motion. Grafen's $\rho = 2.0$ and 5.0 indicate traits evolving less that expected by Brownian motion. In all parameter combinations the number of species in the pool was retained constant in 100 species. Rejection rates were computed after testing 1,000 data sets generated with each parameter combination using a targeted level of $\alpha = 0.05$.

Grafen's ρ	Mean Mantel statistic (SD)	Rejection rates	
		Mantel	EM-Mantel
0	-0.0002 (0.0151)	0.044	0
1	0.5034 (0.0861)	1	0.059
2	0.6141 (0.0871)	1	0.637
5	0.6762 (0.0881)	1	0.94

CAPÍTULO 2 - PCPS – an R-package for exploring phylogenetic eigenvectors across metacommunities[*]

Abstract

PCPS is an R package for exploring phylogenetic eigenvectors across metacommunities. It offers a set of functions for analyzing principal coordinates of phylogenetic structure (PCPS), allowing analysis of phylogenetic signal in ecological traits of species at the metacommunity level, and the association between each PCPS and environmental, spatial and historical factors. The package is a flexible solution for exploring the distribution of major phylogenetic lineages across ecological or biogeographic gradients. The package is freely available on the CRAN official web server for R.

Keywords: analysis; ecophylogenetics; null model; phylogenetic patterns; phylogenetic signal.

[*] Manuscrito publicado no periódico *Frontiers of Biogeography*. Leandro da Silva Duarte é coautor do artigo.

Introduction

Several factors can affect species' distributions; one of them is the phylogenetic relationship among the clades that comprise the metacommunities. The profusion of studies on phylogenetic patterns for different lineages has made it possible for ecologists to explain the evolutionary basis of species' assembly into biological communities and to investigate the major mechanisms underlying it. Early work focused mainly on assessing whether species assembled in local communities are more or less phylogenetically similar than expected given a regional species pool (Webb et al. 2002). More recently, the distribution of different phylogenetic clades across ecological gradients has been explored using different approaches, giving birth to the new discipline of metacommunity phylogenetics (Leibold et al. 2010), a research field interested in unveiling the ecological and historical determinants of phylobetadiversity patterns (Duarte 2011, Peres-Neto et al. 2012, Duarte et al. 2014).

Principal coordinates of phylogenetic structure (PCPS - Duarte 2011, Duarte et al. 2012) constitute a useful tool for exploring phylogenetic patterns across a set of ecological communities. The method involves decomposing the phylogenetic information at the metacommunity level, which is defined using phylogenetic fuzzy weighting (Pillar & Duarte 2010) in several orthogonal eigenvectors. Each eigenvector is a phylogenetic gradient for the set of communities, capturing the variations in the entire phylogeny, from basal to terminal nodes (Duarte et al. 2012). The advantage of this method lies in the possibility of exploring each phylogenetic gradient independently of the others; therefore, it is possible to evaluate which clades are related to each phylogenetic gradient across the metacommunity and to explore the identity of these clades driving phylobetadiversity patterns among the sites

(Duarte et al. 2012). The PCPS are obtained from a species-composition matrix weighted by phylogenetic distances among species, a methodological approach that employs fuzzy set theory to scale pairwise phylogenetic distances among species up to the metacommunity level (Pillar and Duarte 1010, Debastiani and Pillar 2012). PCPS is an R package[1] released under open-source license, and freely available from CRAN[2]. The package has a set of functions for the analysis of PCPS.

Features

The starting point of PCPS analysis is to arrange a set of data matrices: the community matrix, which may contain either species' presence–absence data or abundances; the pairwise phylogenetic distances between species; environmental/spatial/historical variables for each community (optional); and traits describing the species (optional). The PCPS package operates in an integrated manner with the *SYNCSA* package (Debastiani and Pillar 2012), which is used to compute the matrix describing phylogeny-weighted species composition (function *matrix.p*). It can also be used to organize the data matrices (function *organize.syncsa*), given that the species and community sequence in the data matrices must be the same for all data matrices.

Principal coordinates of phylogenetic structure

The core function in the PCPS package is the function called *pcps*. This function performs a principal coordinates analysis (PCoA, Gower 1966) on the matrix

[1] R Development Core Team, http://www.r-project.org/, last accessed 19/09/2014
[2] CRAN package repository, http://cran.r-project.org/, last accessed 19/09/2014

describing phylogeny-weighted species composition, thus generating the phylogenetic eigenvectors called PCPS (Duarte 2011). Each eigenvector represents a single phylogenetic gradient across the metacommunity, which is orthogonal to all other eigenvectors. The PCPS with higher eigenvalues describe wide phylogenetic gradients related to deeper nodes in the phylogeny, while other eigenvalues describe phylogenetic gradients related to shallower nodes (Duarte et al. 2012). Furthermore the *pcps* function computes correlations between each PCPS axis and phylogenetically weighted species abundances/frequencies, thus allowing biplots relating communities and species/clade scores to be built (Fig. 1).

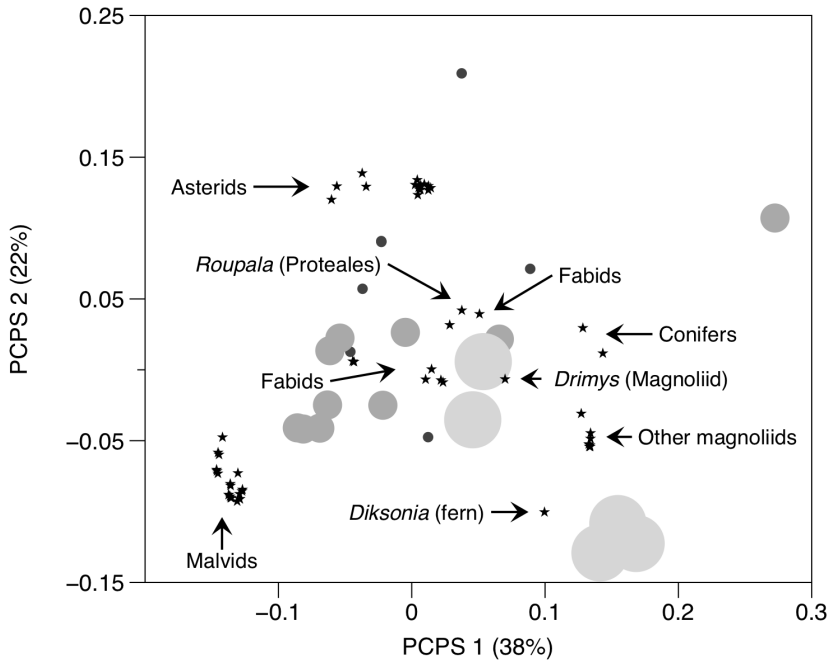

Fig. 1 - Scatter diagram of the first two PCPS axes computed for woody species occurring in forest patches of different sizes in Southern Brazil (modified from Duarte 2011). Circle size represents patch size (small, medium and large). Stars represent species grouped in monophyletic clades in the diagram. The nearness between the clades and the patches belonging to different sizes classes shows the association between them. For example, large patches were associated with *Dicksonia*, conifer trees and magnoliid angiosperms, whereas small patches were related to Asterids.

Phylogenetic signal at metacommunity level

The function *pcps.curve* estimates the phylogenetic signal at the metacommunity level (Pillar and Duarte 2010) for an ecologically relevant species trait. The first step for this analysis consists of scaling species trait information up to the metacommunity level, which is done by averaging trait values across the set of communities (Garnier et al. 2004). Then, the community-averaged trait is taken as the response variable in sequential linear regressions, in which an increasing number of PCPS are taken as predictor variables. Accordingly, in the first regression only the first PCPS axis is used, in the second regression the first two PCPS axes are taken as predictors, and so on. Finally, a curve representing the phylogenetic signal at the metacommunity level is drawn using the proportional accumulation of eigenvalues as new PCPS axes are incorporated into regression (x-axis) and the coefficient of determination of regressions (y-axis) (Diniz-Filho et al. 2012). That curve shows the degree of association between the community-averaged trait and the PCPS axes. A deviation of the curve under or over the 1:1 line, where there is a perfect match between the cumulative phylogenetic variability expressed by the PCPS and the R^2 of the community-averaged trait model, would indicate phylogenetic signal weaker or stronger, respectively, than that expected from the Brownian motion model of evolution (Diniz-Filho et al. 2012), according to which trait variance increases linearly with time.

Nonetheless, PCPS capture not only the phylogenetic signal across the metacommunity, but also the species composition. Therefore, to evaluate whether the curve is representing phylogenetic signal stronger or weaker than expected by

Brownian trait evolution, we should control for the influence of species compositional variation across the metacommunity. For this, the function *pcps.curve* draws null curves, which are generated by shuffling the terminal tips across the phylogenetic tree (Bryant et al. 2008, Kembel et al. 2010) to compute a set of null PCPS. The null PCPS axes are taken as predictors of the linear regressions on the community-averaged trait, and generate curves under the scenario of a random distribution of species across phylogenetic tree. When the observed curve falls above the range of the null curves, it indicates that the phylogenetic signal at the metacommunity level is higher than expected merely by chance. On the other hand, if the observed curve falls under the range of null curves, it indicates that the association between community-averaged traits and the PCPS is lower than expected by chance (Fig. 2).

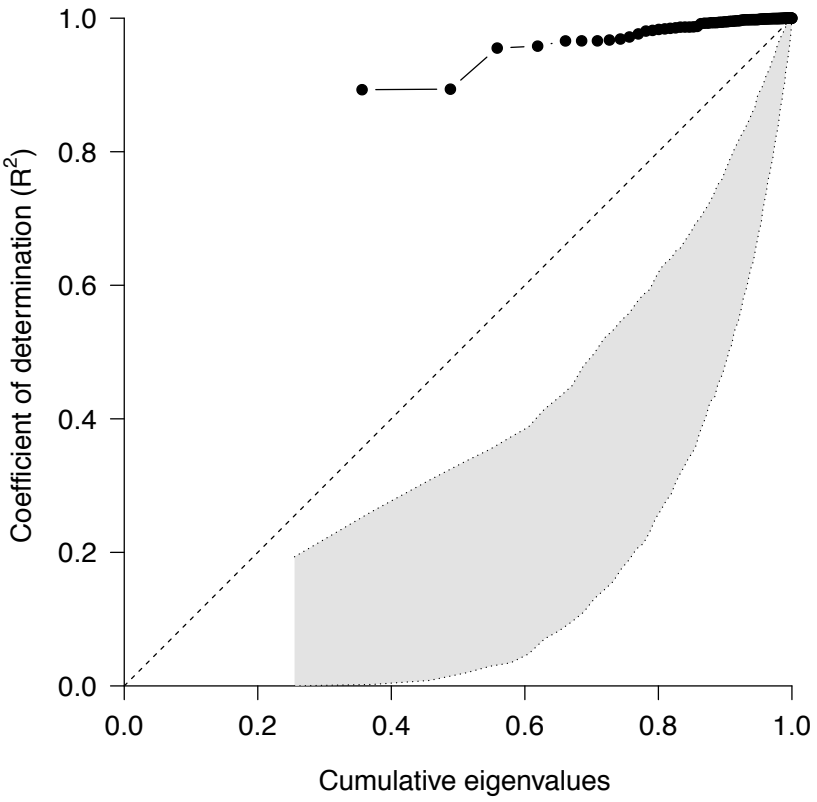

Fig. 2. Hypothetical curve representing phylogenetic signal at the metacommunity level. The diagonal line represents the 1:1 ratio between cumulative eigenvalues of PCPS and coefficient of determination of the community-averaged trait regression on PCPS. The black dots represent the observed curve and the shaded area shows 95% confidence intervals based on null curves generated using a null model that shuffles the terminal tips across the phylogenetic tree.

Association between PCPS axes and environmental factors

The function *pcps.sig* runs a generalized linear model (GLM) with a Gaussian error distribution (Nelder and Wedderburn 1972) to analyze the association between a single PCPS and a set of environmental and/or historical predictors (either categorical, quantitative, or both) (Debastiani et al. unpublished). The significance of the model is obtained by comparing its F-value with null F-values obtained from null models. The null models shuffle terminal tips across the phylogenetic tree to randomize phylogenetic relationships among species, given a tree topology and branch lengths (Bryant et al. 2008, Kembel et al. 2010) and generate sets of null PCPS. The null PCPS are then submitted to a procrustean adjustment (Jackson 1995) and the fitted values between observed PCPS and null PCPS are obtained. Then, the adjusted null PCPS are taken as response variables, the model is rerun, and null F-values are generated. The fraction between the number of null F-values higher than the original F-value and the total number of null F-values computed is taken as the probability of the association between a given PCPS and a set of environmental variables being generated merely by chance.

Questions that can be answered using the PCPS approach

The package PCPS can be used to explore phylogenetic patterns in metacommunities, from finer (communities, assemblages) to broader spatial scales (biomes, continents). At finer scales PCPS analysis has been used to evaluate the association between the distribution of clades across different habitat types for woody plant communities developing over grasslands (Duarte 2011) and along natural grassland–forest ecotones (Debastiani et al. unpublished), and for avian communities distributed across a coastal gradient in Southern Brazil (Gianuca et al. 2014). At wider scales the method allows us to assess the extent to which the distribution of different phylogenetic clades along biogeographic gradients is determined by environmental conditions or spatial gradients. Some examples of this approach are available for woody plants in the Brazilian *Araucaria* forest biome (Duarte et al. 2012) and New World amphibians (Duarte et al. 2014). Furthermore, PCPS analysis enables us to evaluate turnover in phylogenetic composition through time (see Loyola et al. 2014 for an example for amphibians occurring in protected areas in the Brazilian Atlantic Forest) and also to analyze the extent to which the association between average trait values and environmental gradients is influenced by the phylogenetic composition of sites (Brum et al. 2012, 2013).

Conclusions

The field of ecophylogenetics has experienced accelerated development over the last few years. The PCPS package constitutes a flexible way to explore phylogenetic gradients across metacommunities using the same data manipulation ordinarily used to perform multivariate analysis in R. PCPS allows us to describe phylogenetic eigenvectors across metacommunities and to analyze their responses to environmental

factors and the links between phylogenetic patterns and community-averaged traits of species.

References

Brum, F.T., Kindel, A., Hartz, S.M. & Duarte, L.D.S. (2012) Spatial and phylogenetic structure drive frugivory in Tyrannidae birds across the range of Brazilian Araucaria forests. Oikos, 121, 899–906.

Brum, F.T., Gonçalves, L.O., Cappelatti, L., et al. (2013) Land use explains the distribution of threatened New World amphibians better than climate. PloS ONE, 8, e60742.

Bryant, J.A., Lamanna, C., Morlon, H., Kerkhoff, A.J., Enquist, B.J. & Green, J.L. (2008) Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. Proceedings of the National Academy of Sciences of the United States of America, 105, 11505–11511.

Debastiani, V.J. & Pillar, V.D. (2012) SYNCSA-R tool for analysis of metacommunities based on functional traits and phylogeny of the community components. Bioinformatics, 28, 2067–2068.

Diniz-Filho, J.A.F., Rangel, T.F., Santos, T. & Bini, L.M. (2012) Exploring patterns of interspecific variation in quantitative traits using sequential phylogenetic eigenvector regressions. Evolution, 66, 1079–1090.

Duarte, L.D.S. (2011) Phylogenetic habitat filtering influences forest nucleation in grasslands. Oikos, 120, 208–215.

Duarte, L.D.S., Both, C., Debastiani, V.J., et al. (2014) Climate effects on amphibian distributions depend on phylogenetic resolution and the biogeographical history of taxa. Global Ecology and Biogeography, 23, 213–222.

Duarte, L.D.S., Prieto, P. V. & Pillar, V.D. (2012) Assessing spatial and environmental drivers of phylogenetic structure in Brazilian Araucaria forests. Ecography, 35, 952–960.

Garnier, E., Cortez, J., Billes, G., et al. (2004) Plant functional markers capture ecosystem properties during secondary succession. Ecology, 85, 2630–2637.

Gianuca, A.T., Dias, R.A., Debastiani, V.J. & Duarte, L.D.S. (2014). Habitat filtering influences the phylogenetic structure of avian communities across a coastal gradient in southern Brazil. Austral Ecology, 39, 29–38.

Gower, J.C. (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, 54, 325–338.

Jackson, D.A. (1995) Protest: A Procrustean Randomization test of community environment concordance. Ecoscience, 2, 297–303.

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

Leibold, M.A., Economo, E.P. & Peres-Neto, P. (2010) Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. Ecology letters, 13, 1290–1299.

Loyola, R.D., Lemes, P., Brum, F.T., Provete, D.B. & Duarte, L.D.S. (2014). Clade-

specific consequences of climate change to amphibians in Atlantic Forest protected areas. Ecography, 37, 65–72.

Nelder, J.A. & Wedderburn, R.W.M. (1972) Generalized Linear Models. Journal of the Royal Statistical Society Series A, 135, 370–384.

Peres-Neto, P.R., Leibold, M.A. & Dray, S. (2012) Assessing the effects of spatial contingency and environmental filtering on metacommunity phylogenetics. Ecology, 93, S14–S30.

Pillar, V.D. & Duarte, L.D.S. (2010) A framework for metacommunity analysis of phylogenetic structure. Ecology Letters, 13, 587–596.

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.

CAPÍTULO 3 - Phylogenetic signal and clade distribution along environmental gradients[*]

Abstract

Over the past decade, phylogenetic information has been used to explain species distribution and to identify mechanisms of community assembly. Species co-occurrence is an important aspect to consider when accessing phylogenetic niche conservatism, i.e, the tendency of phylogenetically related species to have similar habitat requirements and to occur in similar habitats. We investigate the relationship between phylogenetic signal and the distribution of clades along environmental gradients to access the phylogenetic signal at the metacommunity level. We used numerical simulations to generate metacommunities with different patterns of species co-occurrence and different levels of phylogenetic signal in the species pool. For each simulated metacommunity we generated vectors describing phylogenetic gradients of metacommunity, the so-called Principal Coordinates of Phylogenetic Structure. Furthermore, we described trait variation along simulated metacommunity using trait averages at the community level. These trait averages are modeled by the phylogenetic vectors and the proportion of the variability of traits in the community explained by phylogenetic gradient, representing the phylogenetic signal at the metacommunity level, which is depicted by an accumulation curve. Our results show that the main difference between the results generated by the different parameters used is due to the phylogenetic signal; different species co-occurrence related to clades had similar behaviors of accumulation curves when phylogenetic signal is

[*] Manuscrito a ser submetido para o periódico *Evolutionary Ecology*. Leandro da Silva Duarte é coautor do manuscrito.

equal across the simulated scenarios. Our results show that the phylogenetic structure of community predicts community weighted trait means across metacommunities, even when species are randomly distributed across communities. Thus the phylogenetic signal expressed in traits can be transposed to the community level when the relationship between traits at the metacommunity level and phylogenetic structure of metacommunity is analyzed. Furthermore, traits simulated under Brownian motion or more conserved than it generate curves with a wide range of variation, rarely falls outside the confidence intervals generated by null model that test phylogenetically neutral trait diffusion across communities.

Keywords: phylogenetic niche conservatism; trait evolution; community-weighted means; null model; phylogenetic patterns.

Introduction

Over the last years, the number of studies using phylogenetic information to explain species distribution and to identify mechanisms of community assembly has increased drastically. Evolutionary processes came to be more broadly considered in the interpretation of ecological phenomena and ecological processes are also considered in evolutionary analyses (Webb et al. 2002; Roy and Goldberg 2007; Pennell and Harmon 2013). Currently it is known that the interplay between evolutionary and ecological processes can generate and maintain patterns of species diversity, shaping the dynamics of speciation, extinction of species and the distribution and abundance of species (Webb et al. 2002; McPeek 2008; Pelletier et al. 2009; Pillar and Duarte 2010; Schoener 2011). It is widely expected that closely related species tend to be more similar to each other than distantly related species in terms of their morphological and physiological characteristics (Harvey and Pagel 1991). As these characteristics also reflect how species interact with each other and with the environment, it is also expected that closely related species are also similar in terms of their ecological requirements and niche (Prinzing et al. 2001), defined as the set of conditions that a species needs to survive and persist (Chase and Leibold 2003). The idea that related species maintain their ancestral niches, in other words, that there is similarity in the ecological niche of phylogenetic related species, gave rise to a relatively new concept, called phylogenetic niche conservatism (Wiens and Graham 2005; Wiens et al. 2010). A long debate has been raised among ecologists on whether species retain or not their ancestral ecological niches for long periods of time on the evolutionary scale (Losos et al. 2003; Wiens and Graham 2005; Losos 2008; Crips et al. 2009; Cooper et al. 2010; Wiens et al. 2010; Losos 2011).

One of the proposals largely used to test phylogenetic niche conservatism is the analysis of phylogenetic signal, which test the tendency of phylogenetic closely related species to have more similar traits (Pagel 1999; Blomberg et al. 2003; Willis et al. 2008; Davis et al. 2010). This approach uses only trait information and phylogenetic information, assuming that a single trait can capture the information related to all environmental conditions in which a species can persist and consequently, the entire niche space. The phylogenetic signal shows the existence of a relationship between trait and phylogenetic similarities, and the phylogenetic niche conservatism is defined based only on the relationship between trait and phylogenetic relatedness among species in clades. Following this approach, Losos (2008) define phylogenetic niche conservatism as the phenomenon that closely related species are more ecologically similar than might be expected by a neutral model of evolution. In this way, the definition of phylogenetic niche conservatism is relative to the model of evolution considered (Revell et al. 2008; Cooper et al. 2010; Münkemüller et al. 2015). Each model of trait evolution generates a pattern of trait distribution in extant species, which is an additional restriction compared with these model can define the phylogenetic niche conservatism. Thus a neutral models of trait evolution, e.g. Brownian motion, is relevant because it does not make extra mechanistic assumptions to explain trait evolution, species inherit their traits values from ancestors and simply diverge as a function of evolutionary time (Freckleton and Harvey 2006).

One of the main criticisms related to these approaches is that phylogenetic niche conservatism occurs within species throughout the processes of speciation and has been argued that it should be evaluated together with environmental conditions, where evolutionary processes occur effectively (Holt 2009). Therefore, phylogenetic niche conservatism is revealed when phylogenetically related species tend to have

similar habitat requirements, and so occur in similar habitats (Pillar and Duarte 2010; Ulrich et al. 2012). Thus species co-occurrence is an important aspect to access phylogenetic niche conservatism, as using species co-occurrence approximates the definition of realized niche of Hutchinson (Hutchinson 1957), defined as the set of conditions where a species can exist in the presence of interactions with other organisms.

Few studies have incorporated species co-occurrences to access phylogenetic niche conservatism. Pillar and Duarte (2010) developed a framework to test phylogenetic niche conservatism that considered species co-occurrences, based on the validated of casual models relating, phylogenetic composition, functional composition and environmental conditions. Based on this approach, Duarte et al. (submitted) develop the concept of neutral trait diffusion as the tendency of phylogenetically closely related species to co-occur in communities, expressing their niche dimensions more similar than would be expected by neutral expectation, i.e. the Brownian evolutionary process in traits evolution. This allows the implementation of factors not previously considered in other measures, such as the lack of evolutionary model and especially the species co-occurrence. In this context, phylogenetic niche conservatism would take place when functional composition shows higher phylogenetic conservatism in the metacommunity level than expected by neutral trait diffusion, in other words, there is a high phylogenetic signal in the species pool and a high relationship between the phylogenetic community structure and trait variation at the community level (Duarte et al. submitted). The method has a greater agreement with the Hutchinson niche concept and, insofar, is the one of few approaches known to investigate the phylogenetic signal at metacommunity level.

The aim of this study is to investigate the relationship between phylogenetic signal and the distribution of clades along environmental gradients to access the phylogenetic signal at the metacommunity level. We seek to answer the following questions: (1) will the phylogenetic signal present in the species pool be reflected in the metacommunity level? (2) Will the species co-occurrence pattern in metacommunities affect the detection of phylogenetic signal at the metacommunity level? We expected that the phylogenetic signal in the species pool will be reflected in metacommunity level, i.e. more distinct trait values in close related species could reflect more distinct traits variation in the communities related to phylogenetic community structure. We expected that if phylogenetic close related species co-occurred in the metacommunity this could facilitate the detection of phylogenetic signal at metacommunity level. We used numerical simulation to generate metacommunities with different patterns of species distribution and different levels of trait conservatism in the species pool to test the detection of phylogenetic signal at the metacommunity level and two null models to access its significance. Furthermore, we accessed the rejection rate in each simulated scenario for two null models.

Methods

Simulation of phylogeny

We simulated ultrametric phylogenetic trees with 100 species. The phylogeny is created by purely stochastic birth process, with speciation rate 0.1 and extinction rate 0 (Function *sim.bdtree* package geiger [Harmon et al. 2008]). This function simulates the growth of phylogeny by birth-death process homogeny in all lineages, generating

a conservative distribution of speciation events with the numbers of lineage increasing exponentially over time.

Simulation of phylogenetic signal

Based on the simulated phylogenetic tree we simulated traits with different levels of phylogenetic signal. First of all, we transformed the edge length using Grafen's method (Grafen 1989) (Function *compute.brlen* package ape [Paradis 2004]), based on a power transformation (ρ) of branch lengths (ρ = 0.0001, ρ = 1.0 and ρ = 5.0). Lower values shrink deeper branches and lengthen those near the tips, whereas higher values increase branch lengths near the root of the tree. We simulated only one continuous trait, based on phylogenetic tree with transformed edges lengths (function *rTraitCont*, package ape [Paradis 2004]). In this function, we used a Brownian motion model, a neutral model where trait evolution is constant and the differences between species are accumulated over evolutionary time (Freckleton and Harvey 2006). The combinations of Grafen's values and Brownian model generated three combinations of phylogenetic signal in the traits of species (Diniz-filho et al. 2012; Seger et al. 2013). Grafen's ρ = 0.0001 did not have phylogenetic signal, the traits evolved more than what was expected by the Brownian model and the variation of traits between close related species was very high. Grafen's ρ = 1.0 the traits evolved under the Brownian model and with Grafen's ρ = 5.0 traits evolved less than the expected by the Brownian model, resulting close related species showed very similar values of the functional traits, in this case, the species showed high phylogenetic signal. Additional sets of simulations with intermediate level of phylogenetic signal using Grafen's ρ = 1.2, ρ = 1.5 and ρ = 2.0 were performed. These Grafen's values generated distinct

levels of phylogenetic signal in the pool of species according K statistics (Blomberg et al. 2003). The mean and standard deviation of K statistics (function *Kcal*, package picante [Kembel et al. 2010]) for 1000 simulations was: K = 0.031 (SD ± 0.004), K = 1.006 (SD ± 0.698), K = 1.792 (SD ± 1.224), K = 3.618 (SD ± 2.400), K = 8.423 (SD ± 5.359) and K = 25.064 (SD ± 10.137) for Grafen's $\rho = 0.0001$, $\rho = 1.0$, $\rho = 1.2$, $\rho = 1.5$, $\rho = 2.0$ and $\rho = 5.0$, respectively (Figure S1, Supplementary information S1). Values of K higher that 1.0 indicate high phylogenetic signal, whereas K values lower 1.0 indicate low phylogenetic signal.

Simulation of metacommunities

We simulated three sets of metacommunities, where the communities are linked by dispersal of species that can potentially interact (Leibold et al. 2004). We simulated communities following the species-sorting model (Leibold et al. 2004), where resource gradients or environmental gradients affect the local composition in each community. Usually, under the species-sorting model, species have traits that allow their persistence in the community according to resource availability. We used the same idea in our simulations, however restricting species distribution across communities based on their lineages, and not on an explicit trait sorting related to habitat, assuming that it simulates a response to the environmental gradient. This pattern of species co-occurrence, linking response of species lineage to the gradient, is called phylogenetic habitat filtering (Duarte 2011). Phylogenetic habitat filtering refers to the limitations of some clades to persist in certain habitats, based on certain environmental restriction (Debastiani et al. 2015). Phylogenetic habitat filters were created following one hypothetical environmental gradient, where in one of the

extremes species occurrence in a community is restricted based on their phylogenetic similarity. As the gradient changes, all species could be drawn to the communities without any restriction (Figure 1).

The simulation starts sampling one reference species, this was used to generate the phylogenetic habitat filter. The similarities between species were based in quantiles, with three levels of the force of the habitat phylogenetic filter of 0.9, 0.5 and 0. The filter 0.9 is the most restrictive scenario, where only 10% of the more phylogenetic similar species regarding the reference species, could establish themselves in the extreme of the gradient. The filter was gradually becoming less restrictive until all species in the pool could establish themselves in the communities in the other extreme of the environmental gradient. We call this high phylogenetic filter. The same restriction criteria were established for the other filters, the 0.5 and 0, called medium and low phylogenetic filter. When the filter = 0, species were randomly distributed without any phylogenetic restriction. Once established the phylogenetic restrictions for each species in each local community, 10 species were drawn in each community. The species within the filter, those that could establish themselves in the community, received an equal probability of occurrence in the sampling procedure. To ensure that all communities had 10 species, each and every species in the pool received a very small probability of occurrence in the sampling procedure. Each metacommunity was composed by 100 communities, and characterized only by species incidence.

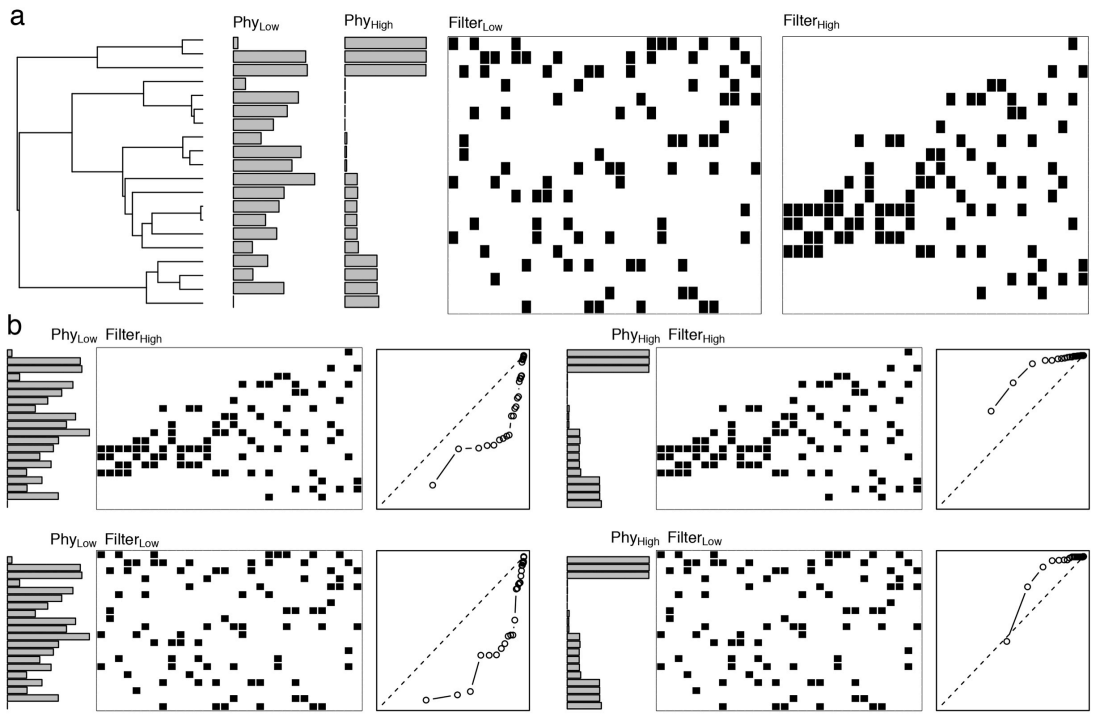

Figure 1. Representation of simulated scenarios of phylogenetic habitat filtering and analysis. (a) The data simulation starts by simulating ultrametric phylogenetic trees, from these trees we simulated traits with different levels of phylogenetic signal, represented in the figure by horizontal bars, similar the trait values have similar the bar lengths. Independent traits are simulated in the metacommunities, based in the same phylogenetic trees. The communities are organized according to the hypothetical environment gradient. In one extreme of the environmental gradient only few species, the ones that are highly phylogenetic related, will be present. In another extreme, all species, independent of their phylogenetic similarity, could be present. Each small square represents the presence of a species in the community. The species traits and the species in the community are matching with the species in the phylogenetic tree at the left. (b) The parameters combinations for the data analyses. Each level of phylogenetic signal and each metacommunity are generated in paired combinations, as well as one observed curve and its statistical significance under the two null models. To simplify the representation of the simulation process, only the extremes scenarios, with examples of corresponding generated data, are represented here; nine parameters combinations compose the complete simulation.

Data analysis

Based on the simulation of three levels of phylogenetic signal and three levels of phylogenetic filters, we generate nine combinations of parameters. For each combination we obtained the community-weighted means (CWM), which express the functional composition of communities. The trait averages at the community level express the trait variation of species along environmental gradient (Garnier et al. 2004; Pillar and Duarte 2010). Furthermore, for each combination of parameters we used phylogenetic distances and the metacommunity for defining the phylogeny-weighted species composition (Pillar and Duarte 2010). This method, used to scale-up the phylogenetic information to the community level, consists in using the phylogenetic distances for defining degrees of belonging between species present in the metacommunity following the fuzzy set theory (Zadeh 1965). Based on their phylogenetic similarity, every species within the metacommunity specifies a fuzzy set to which every species belongs, including itself, with a certain degree of belonging, i.e., each specie simultaneously belong to more than one specie with certain degrees (Pillar et al. 2009; Pillar and Duarte 2010). This degree of belonging will range 0 to 1 for each pair of species, being that the degrees of belonging of a given species across the entire fuzzy sets are standardized to unit total. Using matrix multiplication it is possible to scale-up the information contained in the phylogenetic matrix to the community level (Pillar et al. 2009; Pillar and Duarte 2010). This multiplication results in a matrix that contains species composition after fuzzy weighting based on phylogenetic similarities, and expresses the phylogenetic structure of each community. We performed a Principal Coordinates Analysis in the phylogeny-weighted matrix of species composition using the square root of Bray-Curtis as the resemblance measure (Legendre and Legendre 2012). This procedure generated axis

of phylogenetic variation of phylogenetic gradient, the Principal Coordinates of Phylogenetic Structure (PCPS, Duarte 2011), where each vector describes one phylogenetic gradient of community (Duarte et al. 2012). The PCPS were used as predictors in a linear regression for modeling the trait averages at communities to explain the proportion of the variability of traits in the community, representing the phylogenetic signal at the metacommunity level (Pillar and Duarte 2010; Debastiani and Duarte 2014). We used a sequential approach in order to model trait average as a function of PCPS in the metacommunity, in other words, we first model trait average as being a function of the first PCPS, then, as being a function of first two PCPS, and so on (function *pcps.curve* package PCPS [Debastiani and Duarte 2014]). We plot the determination coefficient of each model and the percentage of accumulation of eigenvalues in the PCPS used in each model for represent the curve of phylogenetic signal at metacommunity level (Debastiani and Duarte 2014; Duarte et al. submitted). Under neutral diffusion of the clades among communities we expect that coefficient of determination and percentage of accumulation of eigenvalues in the PCPS to be linearly related and to generate a diagonal line that increases in 1:1 ratio in the graphic. When the curve is above the line of the ratio between eigenvalues and the percentage of explained variation of the phylogenetic structure of the metacommunity, it explains more than the expected variance of the traits in the community. When the curve of eigenvalue accumulation is below the diagonal line, the explanation of variance of the traits in the community is less than the amount of information contained in the phylogenetic composition matrix (Figure 2). Under strong trait conservatism the relationship between CWM trait and the first PCPS shows a high coefficient of determination. With more labile traits models relating CWM trait and the first PCPS show a low coefficient of determination and with the

increase of numbers of PCPS in the model have increments in coefficient of determination (Duarte et al. submitted).

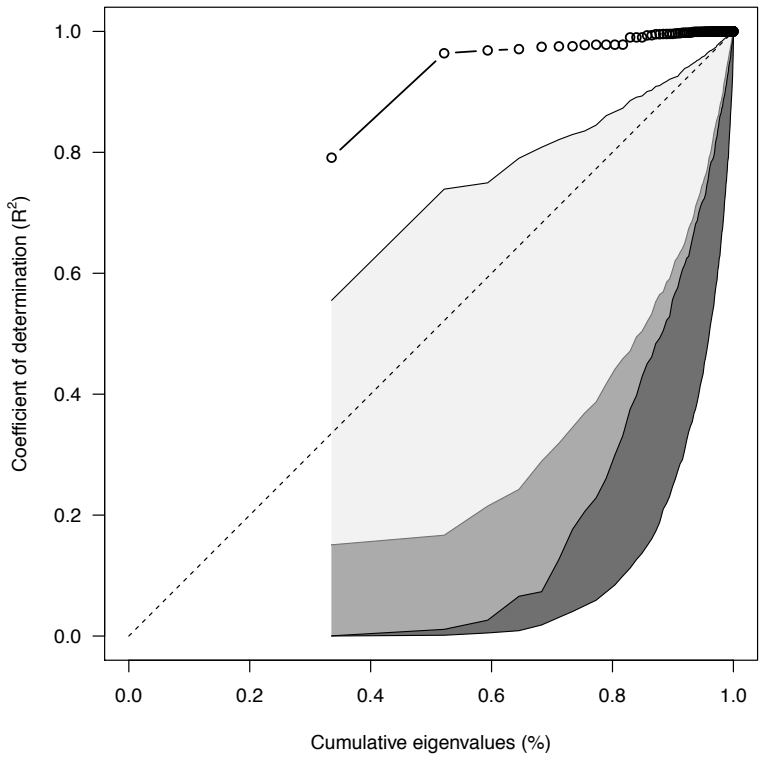

Figure 2. Curve representing phylogenetic signal at the metacommunity level. The diagonal line represents the 1:1 ratio between cumulative eigenvalues of PCPS and coefficient of determination of the CWM regression on PCPS. The dots represent the observed curve for a hypothetical example. The shaded areas show 95% confidence intervals based on null curves with two null models. The dark gray shaded area shows null curves generated using a null model that shuffles the terminal tips across the phylogenetic tree (TS null model) and the light gray shaded area shows null curves generated using a the null model that simulate traits under the Brownian motion (BM null model). The intermediary gray shaded is only the overlap between dark gray and light gray.

The shape of the incremental curve is not enough to define if it is a true pattern of phylogenetic niche conservatism, because the information of species

composition are kept constants in both side of regression, in the response variable and in the predictors, therefore we used two null models to access the significance of observed curves. The first null model is used to control for the influence of species compositional variation across the metacommunity and test if the phylogeny has importance for the structure of community composition (Debastiani and Duarte 2014). The null model keeps the species composition and shuffle terminal tips across the phylogenetic tree to compute a set of null PCPS and generates a set of random PCPS used to recalculate the curves. This null model is called taxa shuffle (TS). If the observed curve falls above the confidence intervals for the distribution of null curve under this null model, it indicates that the phylogeny has an important role for the structure of the metacommunity. The second null model also keeps the species composition and estimate neutral phenotypic diffusion across the metacommunity, given the phylogenetic structure of a metacommunity (Duarte et al. submitted). This null model consists in simulating a trait evolving under Brownian motion, the simplest neutral model for evolution of traits used for continuous characters (Felsenstein 1985), based on the phylogenetic relationship among the species occurring in the metacommunity. The neutral curves are computed using CWM values for traits evolving under the Brownian motion, keeping the PCPS constants. This null model is called Brownian motion (BM). Observed curves above the confidence intervals for the distribution of neutral curve indicate that the observed CWM variation between communities is larger than expected by a phylogenetically neutral trait diffusion process across the set of communities.

For each combination of parameters, we ran 500 simulations and evaluated the average accumulation curve and the proportion of significant curve under each null model (i.e., rejection rate) based in 999 permutations. We recorded the proportion of

observed curve that falls above the confidence intervals for the distribution of each null model to estimate the statistical power and type I error rate (i.e., rate of false positives). For the TS null model the type I error was estimated considering the rejection rates of simulations performed using Grafen's $\rho = 0.0001$. For BM null model the type I error was estimated considering the rejection rates of simulations performed using Grafen's $\rho = 0.0001$ and 1.0. The rejection rate was achieved for a targeted level of $\alpha = 0.05$, considering curves that are above the null curves under both null models. All numerical simulations were conducted in R (R Development Core Team 2014).

Results

Our results showed that the main difference in the results produced using the different parameters was due to the phylogenetic signal in the species pool. Different phylogenetic filters related to clades had similar behaviors of accumulation curves when phylogenetic signal is the same (Figure 3). In the absence of phylogenetic signal in the species pool level, values of the coefficient of determination are lower than expected by the accumulation of eigenvalues in the PCPS, independently of the intensity of phylogenetic habitat filter (Figure 3, column 1). Furthermore, there is little variation between simulations. When traits evolve as expected by Brownian motion evolution of the average of the simulations, it shows that the values of coefficient of determination increases linearly with the eigenvalues retained by PCPS (Figure 3, column 2). When there is phylogenetic signal in the traits of the species, the accumulation curve is higher than expected by phylogenetic eigenvalues (Figure 3,

column 3) and the variation around the average is much higher than what is found when there is no phylogenetic signal in species traits.

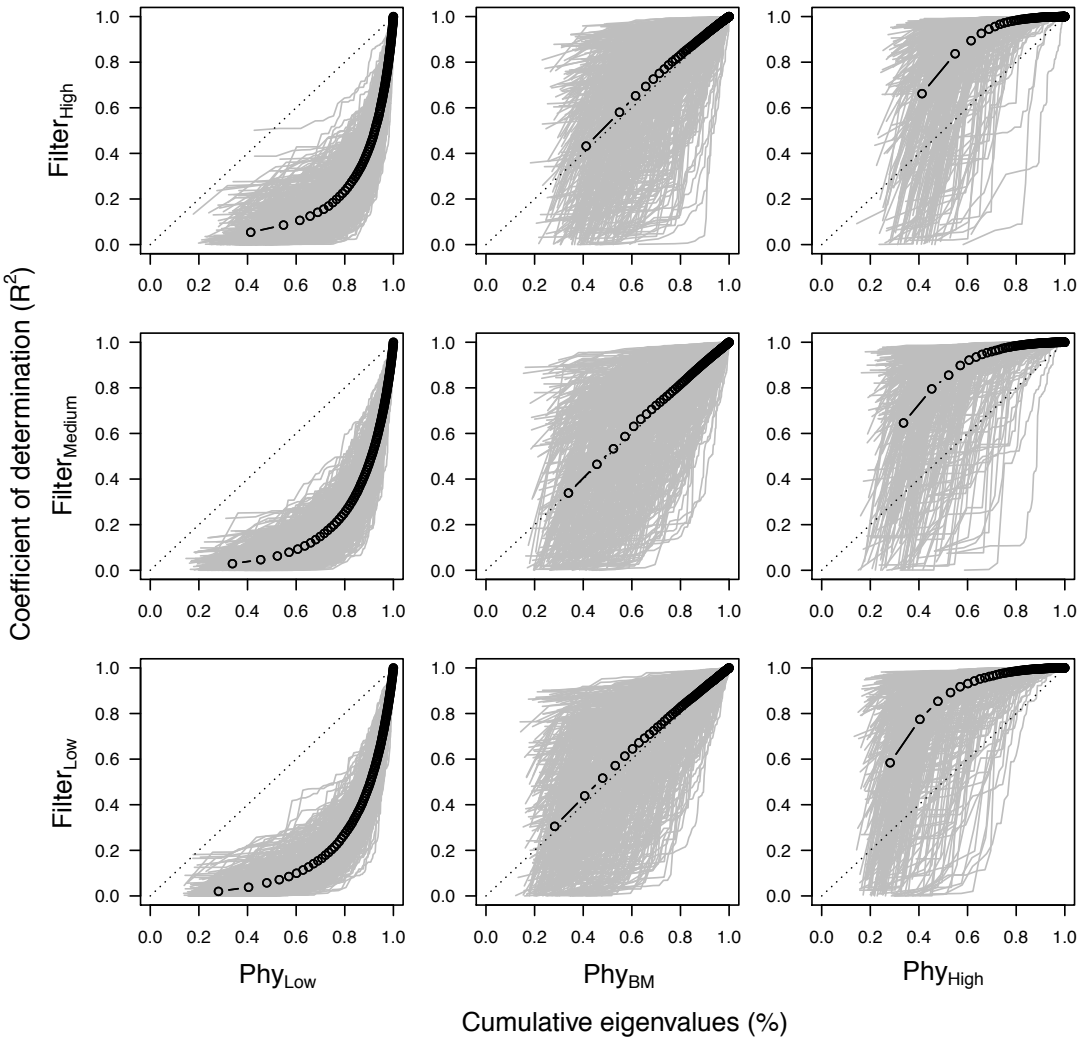

Figure 3. Phylogenetic signal at the metacommunity level for each parameter combination. Each curve shows the accumulation of coefficient of determination derived from regressions between CWM and sequential PCPS. The first column of graphics shows the results without phylogenetic signal (Grafen's ρ = 0.0001), following the second column the Brownian motion (Grafen's ρ = 1.0) and the last one the high phylogenetic signal for the pool of species (Grafen's ρ = 5.0). The first line shows high phylogenetic habitat filter in the metacommunity (PhyHigh, Filter = 0.9),

following for the medium habitat filter (FilterMedium, Filter = 0.5) and by metacommunity without any habitat filter (FilterLow, Filter = 0). Points in black are the mean for 500 simulations; each grey line shows the observed curve for individual simulation.

Phylogenetic habitat filters exerted a minor influence on the generated patterns if habitat filters are high (filter = 0.9) the values of the first PCPS captures around 40% of the variation in the phylogeny-weighted species composition matrix. When habitat filters are missing or have intermediate levels, filter = 0 and filter = 0.5 respectively, the first PCPS tends to accumulate around 30% of the variation in the phylogeny-weighted species composition matrix.

When traits do not show phylogenetic signal in the species pool and the null model that shuffles terminal tips across the phylogenetic tree (TS null model), which test if phylogenetic information have any relationship with trait variation at the community level, the rejection rate ranged from 0.048 to 0.160 for the 10 first PCPSs for all metacommunities, independent of the species co-occurrence in metacommunities (Figure 4, column 1 and Table S1). For the Brownian motion null model (BM null model) the rejection rate was zero for traits that do not show phylogenetic signal in the species pool. For scenarios with traits evolving under Brownian motion, the rejection rate ranged from 0.734 to 0.986 for the TS null model and from 0.034 to 0.082 for the BM null model for 10 firsts PCPSs. With high phylogenetic signal the rejection rate ranged between 0.868 and 1 for the TS null model and from 0.362 to 0.914 for the BM null model for the 10 firsts PCPS (Figure 4, column 2 and Table S1). In all parameter combination the rejection rate was similar

within the same level of phylogenetic signal in the species pool, independent of species distribution in the metacommunity (Figure 4 and Table S1).

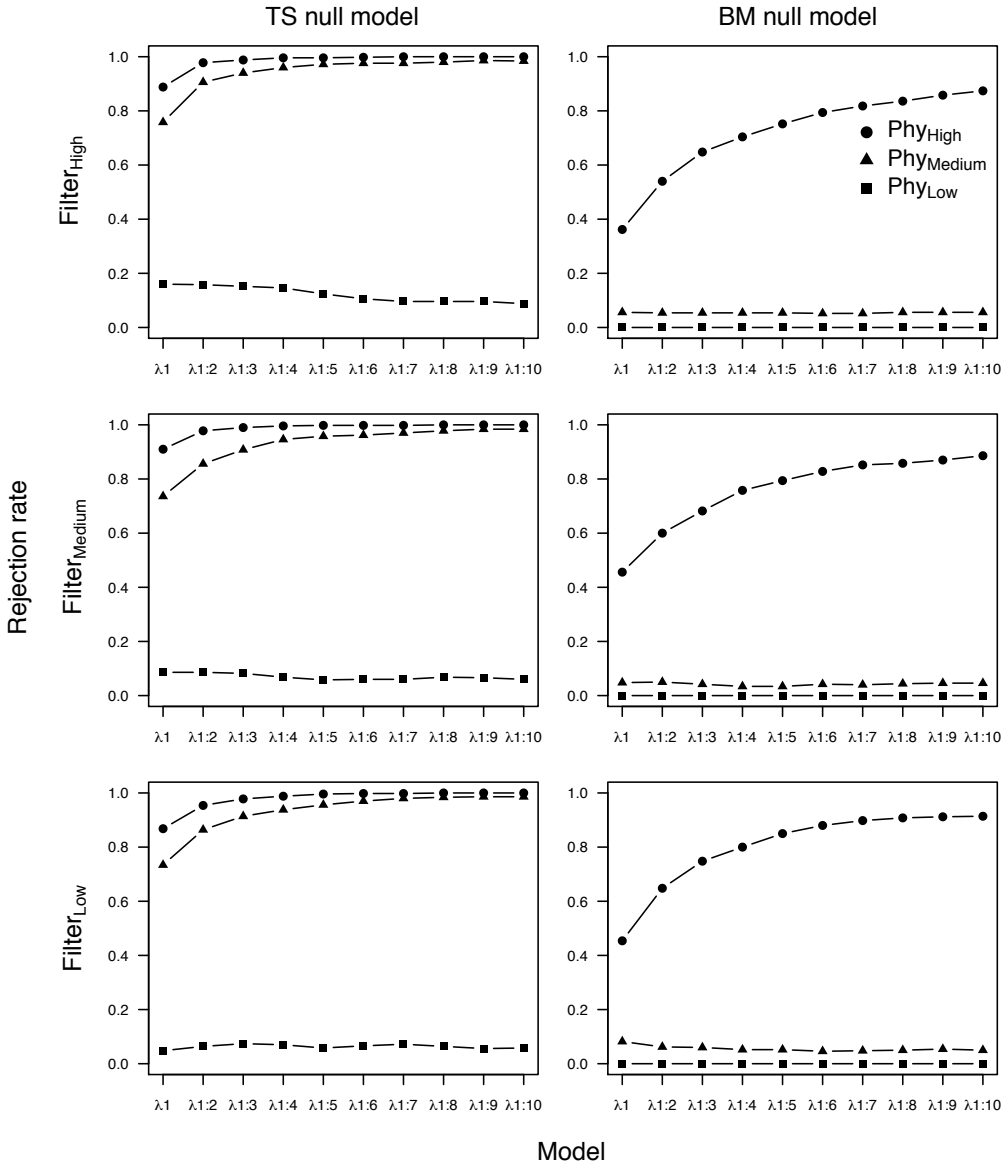

Figure 4. Rejection rate for each null model in all nine parameters combination. Each model (λ 1, λ 1:2, λ 1:3, …) shows the results for each PCPS modeling the CWM following the sequential approach (PCPS1, PCPS1+PCPS2, and so on). The first column of graphics shows the results for the TS null model, which shuffles the terminal tips across the phylogenetic tree, followed by the second column that shows the results for BM null model, which estimates neutral phenotypic diffusion across the metacommunity. The rejection rate is calculated independently for each model, considering the target level of $\alpha = 0.05$, considering the proportion of curves that are

above the null curves under both null models. PhyLow are the results of traits without phylogenetic signal (Grafen's ρ = 0.0001); PhyBM are the results of traits under Brownian motion (Grafen's ρ = 1.0); PhyHigh are the results of traits conserved with high phylogenetic signal for the species pool (Grafen's ρ = 5.0); FilterLow are metacommunity without any habitat filter (Filter = 0); FilterMedium shows results with medium habitat filter (Filter = 0.5); and FilterHigh shows results with high phylogenetic habitat filter in the metacommunity (Filter = 0.9).

Discussion

Our results with simulated data show that when the species pool does not exhibit phylogenetic signal, the trait community-weighted means fluctuate in a random manner and it is not possible to predict trait variation based on the phylogenetic structure of the metacommunity. However, when the simulated species pool presents phylogenetic signal, it is possible to predict part of trait variation in the simulated metacommunity based on its phylogenetic structure, even when species were randomly distributed across the metacommunity. Determining the magnitude of phylogenetic signal in the species pool is a crucial question to access the relationship between phylogenetic structure and trait variation in the community level. When the species are randomly distributed, the correlation between the community-weighted means and any environmental variable is will likely be always low, but if species traits are similar in close related species, it is possible to predict the variation of traits in the community (CWM) based only on the PCPS. Our result expands the result found by Duarte et al. (submitted) that shows that when phylogenetic signal is high the phylogenetic structure of community predicts community weighted trait means across metacommunities, even when species are randomly distributed across communities.

Our simulation used a wide range of phylogenetic signal in the species pool that varied from traits with no phylogenetic signal to highly conserved traits. Our results shows that the phylogenetic signal expressed in traits can be transposed to the metacommunity level when the relationship between traits at the metacommunity level and phylogenetic structure is analyzed. It seems to partially contradict some studies that suggested that the phylogenetic signal in the species pool-level is poorly maintained at the metacommunity level even when traits were phylogenetically conserved (Mason and Pavoine 2013). The accumulation curve in the metacommunity level, developed in this work, follows the same interpretation as the work developed by Diniz-filho et al. (2012), where accumulation curves are used to analyze the signal in the species pool level. The mean of accumulation curve is close to the relation 1:1 between cumulative PCPS and coefficient of determination when the traits evolve by a Brownian motion model. Our results show that in intermediate levels of phylogenetic signal the pattern presented by accumulation curves are similar to what is observed when traits evolve under Brownian motion (Figure S2 and Table S2, Supplementary information S2 and S3 respectively). Under Brownian motion or intermediate levels of phylogenetic signal the curves showed a wide range of explained variation of CWM. However, in highly conserved traits, the proportion of explanation of PCPS in predicting CWM variation is high.

We expected that different intensity of the simulated environmental filter would have different effects in the accumulation curve of eigenvalues, but this did not happen. Weak or high phylogenetic habitat filters showed similar patterns of phylogenetic signal at the metacommunity level. The phylogeny-weighted species composition is formed by the information of species composition plus phylogenetic information. If all species are entirely independent (i.e, star phylogeny), the

phylogeny-weighted species composition is the same as the species composition matrix (Pillar and Duarte 2010). In this case the expected relationship between the cumulative eigenvalues for the PCPS and the coefficient of determination is equal to the ratio 1:1 (Supplementary information S4). This can be explained purely by the accumulation of information in the compositional matrix, as each species is used for both, calculating the CWM and for obtaining the phylogeny-weighted species composition matrix (i.e., composition matrix when all species are entirely independent). When traits are labile, with less phylogenetic signal than expected by the neutral model, the phylogenetic information does not predict anything and only disrupts the relationship between the phylogenetic structure of community and the CWM. When the traits evolve under Brownian motion or more conserved than it, a strong relationship between the phylogenetic distribution of species in the communities and the CWM variation will be observed.

Our null model that shuffles terminal tips across the phylogenetic tree (TS null model) shows appropriate rejection rate (type I error) when traits do not show phylogenetic signal in the species pool (values near to the targeted level of $\alpha = 0.05$). The power is high, superior to 0.734, for traits evolving under Brownian motion or traits with high phylogenetic signal. Likewise, the second null model that estimate neutral phenotypic diffusion across the metacommunity also shows appropriate type I error rates for traits under Brownian motion evolution. The statistical power of BM null model is high for highly conserved traits, but for intermediate levels of phylogenetic signal it is weak, e.g. rejection rate up to 0.174 for Grafen's $\rho = 2.0$ for the CWM modeled by the first PCPS (Table S2, Supplementary information S3). Traits simulated with intermediate levels of phylogenetic signal generate curves with a wide range of variation, rarely observing curves staying outside the confidence

intervals generated by BM null model. This could be explained because trait variation in the metacommunity is summarized using trait average at the community level (CWM). As not all species were involved in the calculation of CWM (each community had only some species from the total species pool), being that the variation of CWM with simulated traits was very wide and consequently the prediction by phylogenetic structure was very wide, as well. Furthermore, trait average at the community level (CWM) that could be non-informative in some situations being that other measures of community trait variation could be more informative. Another potential limitation of our modeling approach is that we considered phylogenetic trees with homogeneous rates of speciation in all lineages, without soft polytomies, and this could influence the results (Mazel et al. 2015).

In a recent contribution, Duarte et al. (submitted) suggested that empirical communities show distinct shapes of accumulation curves that can be generated by the number of communities in the metacommunity, species relative abundances and the topology and distributions of branch lengths of the phylogenetic tree (Duarte et al. submitted). Diniz-Filho et al. (2015) demonstrate that when measuring phylogenetic signal at the species pool level, non-stationarity in trait evolution, with pronounced morphological or physiological differentiation in a relatively short period of time may cause abrupt changes in the shape of the curves. However, we suggest here that at the metacommunity level, these abrupt changes may also be a consequence of important historical factors and species distribution patterns that are not uniform among clades, i.e., species in one clade are much more abundant than species in other clades.

Since the accumulation curves of eigenvalues at metacommunity level is not enough to assess phylogenetic niche conservatism, the use of null models is extremely

important. The null models used here only test if phylogenies have an important role for the structure of the metacommunity and if CWM variation among communities is larger than expected by phylogenetically neutral trait diffusion, always keeping the species co-occurrence constant. Null models could be used to discriminate other patterns in species distribution along the shape of curves (Gotelli 2000).

We showed that strong relationship between phylogeny and traits can be found even when traits are not related to the environment. If phylogenetic niche conservatism must be related to specific environmental variable, additional tests are required (Pillar and Duarte 2010; Ulrich et al. 2012), using traits and environmental variable that are chosen appropriately. Despite its importance, phylogenetic niche conservatism is a complicated concept (Pyron et al. 2015), as there are several elements that must be combined in order to have a cohesive framework to evaluate the interplay between evolutionary and ecological processes. The analytical approach established in this work attempt to formalize operationally the concept of phylogenetic signal at metacommunity level incorporating species co-occurrences and approximate to concept of phylogenetic niche conservatism. The challenge consists in understanding the interplay between both local processes, such as competition and habitat filtering, and long term processes, such as dispersal, trait evolution and speciation and extinction dynamics (Pennell and Harmon 2013).

References

Blomberg SP, Garland T, Ives AR (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717–45

Chase JM, Leibold MA (2003) Ecological niches: linking classical and contemporary approaches. University of Chicago Press, Chicago

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

Crisp MD, Arroyo MTK, Cook LG, et al. (2009) Phylogenetic biome conservatism on a global scale. Nature 458:754–756

Davis CC, Willis CG, Primack RB, Miller-Rushing AJ (2010) The importance of phylogeny to the study of phenological response to global climate change. Philos Trans R Soc Lond B Biol Sci 365:3201–3213

Debastiani VJ, Duarte LDS (2014) PCPS - an R-package for exploring phylogenetic eigenvectors across metacommunities. Front Biogeogr 6:144–148

Debastiani VJ, Muller SC, Oliveira JM, et al. (2015) Recurrent patterns of phylogenetic habitat filtering in woody plant communities across phytogeographically distinct grassland-forest ecotones. Community Ecol 16:1–9

Diniz-Filho JAF, Alves DMCC, Villalobos F, et al. (2015) Phylogenetic eigenvectors and nonstationarity in the evolution of theropod dinosaur skulls. J Evol Biol 28:1410–1416

Duarte LDS (2011) Phylogenetic habitat filtering influences forest nucleation in grasslands. Oikos 120:208–215

Duarte LDS, Prieto P V., Pillar VD (2012) Assessing spatial and environmental drivers of phylogenetic structure in Brazilian Araucaria forests. Ecography 35:952–960

Duarte LDS, Debastiani VJ, Carlucci MB, et al. (submitted). Community-weighted trait means: Should phylogeny stay or should it go? Functional Ecology.

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

Filho JAFD, Rangel TF, Santos T, et al. (2012) Exploring patterns of interspecific variation in quantitative traits using sequential phylogenetic eigenvector regressions. Evolution 66:1079–1090

Freckleton RP, Harvey PH (2006) Detecting non-Brownian trait evolution in adaptive radiations. PLoS Biol 4:e373

Garnier E, Cortez J, Billes G, et al. (2004) Plant functional markers capture ecosystem properties during secondary succession. Ecology 85:2630–2637

Gotelli N (2000) Null model analysis of species co-occurrence patterns. Ecology 81:2606–2621

Grafen A (1989) The phylogenetic regression. Philos Trans R Soc Lond B Biol Sci 326:119–157

Harmon LJ, Weir JT, Brock CD, et al. (2008) GEIGER: Investigating evolutionary radiations. Bioinformatics 24:129–131

Harvey PH, Pagel MD (1991) The comparative method in evolutionary biology. Oxford University Press, Oxford

Holt RD (2009) Bringing the Hutchinsonian niche into the 21st century: ecological

and evolutionary perspectives. Proc Natl Acad Sci U S A 106:19659–19665

Hutchinson GE (1957) Concluding Remarks. Cold Spring Harb Symp. 415–427

Kembel SW, Cowan PD, Helmus MR, et al. (2010) Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464

Legendre P, Legendre L (2012) Numerical ecology. 3nd ed. Elsevier

Leibold MA, Holyoak M, Mouquet N, et al. (2004) The metacommunity concept: a framework for multi-scale community ecology. Ecol Lett 7:601–613

Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecol Lett 11:995–1003

Losos JB (2011) Seeing the forest for the trees: The limitations of phylogenies in comparative biology. Am Nat 177:709–727

Losos JB, Leal M, Glor RE, et al. (2003) Niche lability in the evolution of a Caribbean lizard community. Nature 424:542–545

Mason NWH, Pavoine S (2013) Does trait conservatism guarantee that indicators of phylogenetic community structure will reveal niche-based assembly processes along stress gradients? J Veg Sci 24:820–833

Mazel F, Davies JT, Gallien L, et al. (2015) Influence of tree shape and evolutionary time-scale on phylogenetic diversity metrics. Ecography 38:1–8

McPeek M A (2008) The ecological dynamics of clade diversification and community assembly. Am Nat 172:E270–284

Münkemüller T, Boucher FC, Thuiller W, Lavergne S (2015) Phylogenetic niche conservatism - common pitfalls and ways forward. Funct Ecol 29:627–639

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

Paradis E (2004) APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20:289–290

Pelletier F, Garant D, Hendry a P (2009) Eco-evolutionary dynamics. Philos Trans R Soc Lond B Biol Sci 364:1483–1489

Pennell MW, Harmon LJ (2013) An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. Ann N Y Acad Sci 1289:90–105

Pillar VD, Duarte LDS, Sosinski EE, Joner F (2009) Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. J Veg Sci 20:334–348

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

Prinzing A, Durka W, Klotz S, Brandl R (2001) The niche of higher plants: evidence for phylogenetic conservatism. Proc R Soc London Ser B Biol Sci 268:2383–2389

Pyron RA, Costa GC, Patten MA, Burbrink FT (2015) Phylogenetic niche conservatism and the evolutionary basis of ecological speciation. Biol Rev 90:1248–1262

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

Revell LJ, Harmon LJ, Collar DC (2008) Phylogenetic signal, evolutionary process, and rate. Syst Biol 57:591–601

Roy K, Goldberg EE (2007) Origination, extinction, and dispersal: integrative models for understanding present-day diversity gradients. Am Nat 170 Suppl :S71–85

Schoener TW (2011) The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. Science 331:426–429

Seger GDS, Duarte LDS, Debastiani VJ, et al. (2013) Discriminating the effects of phylogenetic hypothesis, tree resolution and clade age estimates on phylogenetic signal measurements. Plant Biol 15:858–867

Ulrich W, Piwczyński M, Maestre FT, Gotelli NJ (2012) Null model tests for niche conservatism, phylogenetic assortment and habitat filtering. Methods Ecol Evol 3:930–939

Webb CO, Ackerly DD, Mcpeek MA, Donoghue MJ (2002) Phylogenies and community ecology. Annu Rev Ecol Syst 33:475–505

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

Wiens JJ, Graham CH (2005) Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. Annu Rev Ecol Evol Syst 36:519–539

Willis CG, Ruhfel B, Primack RB, et al. (2008) Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. Proc Natl Acad Sci U S A

105:17029–17033

Zadeh LA (1965) Fuzzy sets. Inf Control 338–353

Supplemental Material

Supplementary information S1

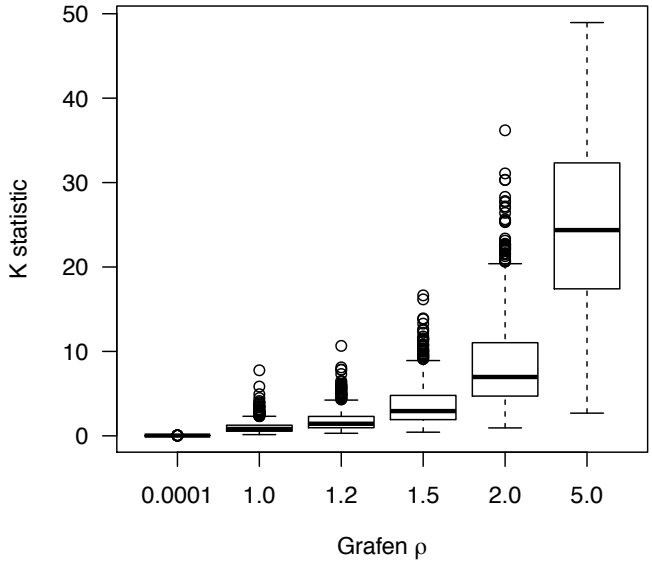

Figure S1. Standard box-whisker-plots showing K statistic under different Grafen's ρ. Simulation were performed with ultrametric phylogenetic trees with 100 species following parameter and procedure described in the main text.

Supplementary information S2

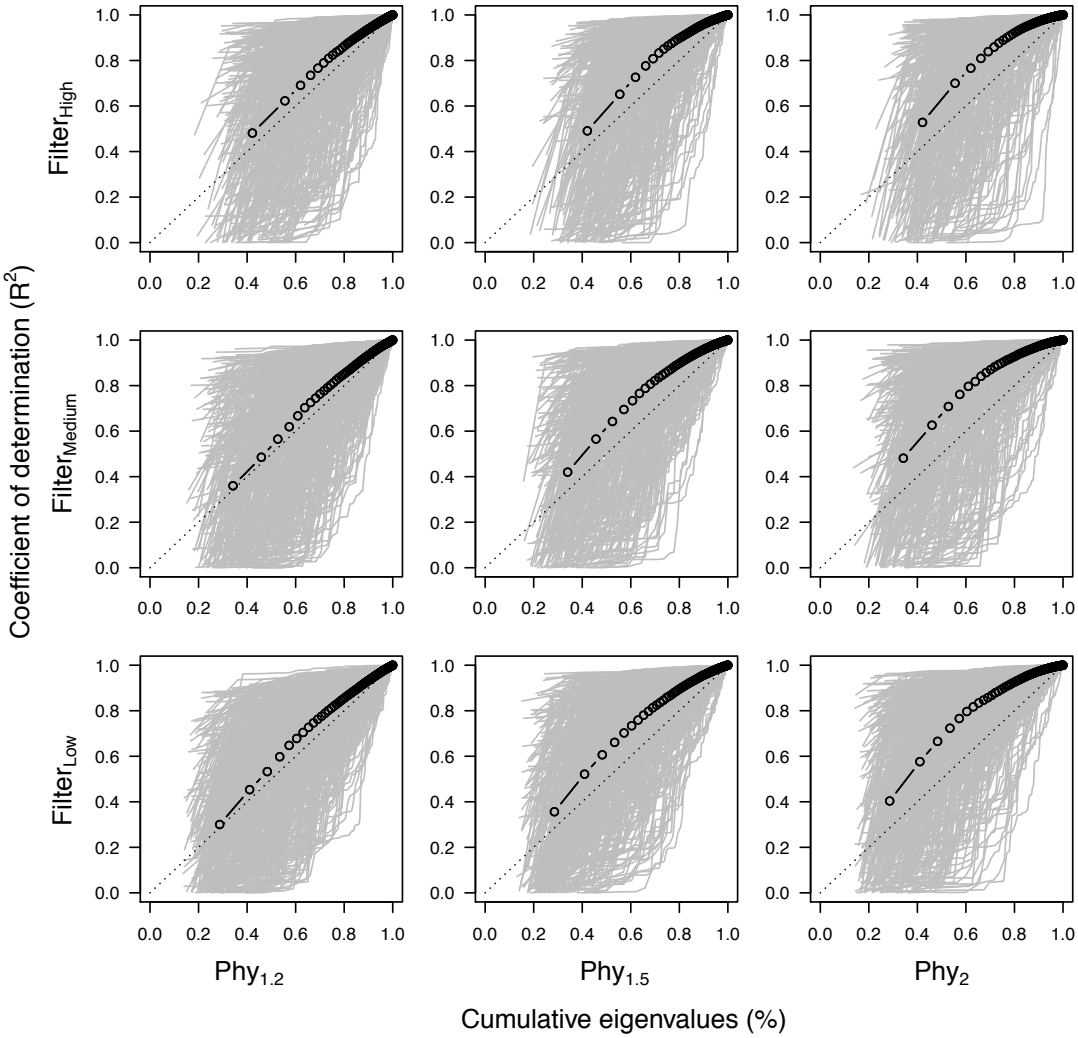

Figure S2. Phylogenetic signal at the metacommunity level, considering intermediate level of phylogenetic signal in the pool of species for each parameter combination of phylogenetic habit filter. Each curve shows the proportion of PCPS accumulation and the coefficient of determination for traits variation in metacommunity level. The first column of graphics shows the results from the phylogenetic signal for the pool of species with Grafen's ρ = 1.2, following the second column with Grafen's ρ = 1.5 and Grafen's ρ = 2.0. The first line shows high phylogenetic habitat filter in the metacommunity (FilterHigh, Filter = 0.9), following for the medium habitat filter (FilterMedium, Filter = 0.5) and by metacommunity without any habitat filter

(FilterLow, Filter = 0). Points in black are the mean for 500 simulations; each grey line shows the observed curve for individual simulation.

Supplementary information S3

Table S1. Rejection rate, for each null model, in all nine parameters combination. Each model (λ 1, λ 1:2, λ 1:3, …) shows the results for each PCPS modeling the CWM following the sequential approach (PCPS1, PCPS1+PCPS2, and so on). TS null model shuffle terminal tips across the phylogenetic tree and BM null model estimate neutral phenotypic diffusion across the metacommunity. The rejection rate is calculated independently for each model, considering the target level of $\alpha = 0.05$, considering the proportion of curves that are above the null curves under both null models. PhyLow are results with traits without phylogenetic signal (Grafen's $\rho = 0.0001$); PhyBM are results with traits under Brownian motion (Grafen's $\rho = 1.0$); PhyHigh are results with traits conserved with high phylogenetic signal for the pool of species (Grafen's $\rho = 5.0$); FilterLow are metacommunity without any habitat filter (Filter = 0); FilterMedium shows results with medium habitat filter (Filter = 0.5); and FilterHigh shows results with high phylogenetic habitat filter in the metacommunity (Filter = 0.9). More details can be found in main text.

Null model	Parameter	λ1	λ1:2	λ1:3	λ1:4	λ1:5	λ1:6	λ1:7	λ1:8	λ1:9	λ1:10
TS null model	PhyLow - FilterLow	0.048	0.064	0.074	0.07	0.058	0.066	0.072	0.064	0.056	0.058
	PhyLow - FilterMedium	0.086	0.086	0.082	0.068	0.058	0.06	0.06	0.068	0.066	0.06
	PhyLow - FilterHigh	0.16	0.158	0.152	0.146	0.124	0.106	0.096	0.096	0.096	0.088
	PhyBM - FilterLow	0.734	0.864	0.914	0.938	0.956	0.97	0.98	0.984	0.986	0.986
	PhyBM - FilterMedium	0.736	0.856	0.908	0.946	0.958	0.962	0.97	0.978	0.984	0.984
	PhyBM - FilterHigh	0.758	0.906	0.94	0.96	0.972	0.976	0.976	0.98	0.986	0.984
	PhyHigh - FilterLow	0.868	0.954	0.978	0.988	0.996	0.998	0.998	1	1	1
	PhyHigh - FilterMedium	0.91	0.978	0.99	0.996	0.998	0.998	0.998	1	1	1
	PhyHigh - FilterHigh	0.888	0.978	0.988	0.996	0.996	0.998	1	1	1	1
BM null model	PhyLow - FilterLow	0	0	0	0	0	0	0	0	0	0

	PhyLow - FilterMedium	0	0	0	0	0	0	0	0	0	0
	PhyLow - FilterHigh	0	0	0	0	0	0	0	0	0	0
	PhyBM - FilterLow	0.082	0.062	0.06	0.052	0.052	0.046	0.048	0.05	0.054	0.05
	PhyBM - FilterMedium	0.048	0.05	0.042	0.034	0.034	0.042	0.04	0.044	0.046	0.046
	PhyBM - FilterHigh	0.056	0.054	0.054	0.054	0.054	0.052	0.052	0.056	0.056	0.056
	PhyHigh - FilterLow	0.454	0.648	0.748	0.8	0.85	0.88	0.898	0.908	0.912	0.914
	PhyHigh - FilterMedium	0.456	0.6	0.682	0.758	0.794	0.828	0.852	0.858	0.87	0.886
	PhyHigh - FilterHigh	0.362	0.54	0.648	0.704	0.752	0.794	0.818	0.836	0.858	0.874

Table S2. Rejection rate with intermediate levels of phylogenetic signal, for each null model. Each model (λ 1, λ 1:2, λ 1:3, …) shows the results for each PCPS modeling the CWM following the sequential approach (PCPS1, PCPS1+PCPS2, and so on). TS null model shuffle terminal tips across the phylogenetic tree and BM null model estimate neutral phenotypic diffusion across the metacommunity. The rejection rate is calculated independently for each model, considering the targed level of $\alpha = 0.05$, considering the proportion of curves that are above the null curves under both null models. Phy_1.2, Phy_1.5, Phy_2.0 with traits conserved with phylogenetic signal for the pool of species (Grafen's $\rho = 1.2$, Grafen's $\rho = 1.5$, Grafen's $\rho = 2.0$, respectively); FilterLow are metacommunity without any habitat filter (Filter = 0); FilterMedium shows results with medium habitat filter (Filter = 0.5); and FilterHigh shows results with high phylogenetic habitat filter in the metacommunity (Filter = 0.9). More details can be found in main text.

Null model	Parameter	λ1	λ1:2	λ1:3	λ1:4	λ1:5	λ1:6	λ1:7	λ1:8	λ1:9	λ1:10
TS null model	Phy_1.2 - FilterLow	0.722	0.88	0.928	0.958	0.98	0.984	0.992	0.992	0.994	0.992
	Phy_1.2 - FilterMedium	0.768	0.87	0.916	0.938	0.964	0.978	0.984	0.986	0.99	0.988
	Phy_1.2 - FilterHigh	0.792	0.91	0.944	0.974	0.978	0.988	0.994	0.998	0.998	0.996
	Phy_1.5 - FilterLow	0.754	0.916	0.954	0.976	0.986	0.986	0.99	0.994	0.996	0.996
	Phy_1.5 - FilterMedium	0.804	0.912	0.952	0.966	0.978	0.992	0.994	1	1	1
	Phy_1.5 - FilterHigh	0.784	0.914	0.96	0.968	0.98	0.986	0.992	0.994	0.994	0.996
	Phy_2.0 - FilterLow	0.808	0.914	0.96	0.982	0.984	0.99	0.99	0.994	0.994	0.994
	Phy_2.0 - FilterMedium	0.836	0.944	0.97	0.982	0.986	0.988	0.994	0.994	0.994	0.994
	Phy_2.0 - FilterHigh	0.812	0.942	0.966	0.978	0.978	0.984	0.99	0.994	0.994	0.994
BM null model	Phy_1.2 - FilterLow	0.064	0.086	0.094	0.09	0.102	0.102	0.106	0.118	0.11	0.116
	Phy_1.2 - FilterMedium	0.068	0.066	0.07	0.07	0.072	0.086	0.084	0.084	0.086	0.086
	Phy_1.2 - FilterHigh	0.07	0.084	0.092	0.102	0.1	0.1	0.106	0.096	0.098	0.104
	Phy_1.5 -	0.124	0.176	0.204	0.196	0.206	0.22	0.226	0.236	0.248	0.262

FilterLow										
Phy_1.5 - FilterMedium	0.118	0.134	0.16	0.174	0.18	0.188	0.208	0.218	0.212	0.224
Phy_1.5 - FilterHigh	0.118	0.13	0.144	0.16	0.182	0.196	0.204	0.202	0.192	0.208
Phy_2.0 - FilterLow	0.174	0.254	0.282	0.314	0.36	0.37	0.41	0.41	0.404	0.416
Phy_2.0 - FilterMedium	0.202	0.236	0.262	0.298	0.324	0.342	0.354	0.388	0.406	0.414
Phy_2.0 - FilterHigh	0.18	0.21	0.234	0.26	0.288	0.302	0.312	0.336	0.334	0.358

Supplementary information S4

Data simulation

Simulation of phylogeny and phylogenetic signal performed following the procedure described in the main text. We simulated ultrametric phylogenetic trees with 100 species and levels of phylogenetic signal using the Grafen's method (see main text). Traits without phylogenetic signal (Grafen's $\rho = 0.0001$), traits under Brownian motion (Grafen's $\rho = 1.0$) and traits with high phylogenetic signal for the pool of species (Grafen's $\rho = 5.0$). We generate metacommunity composed by 100 communities, with 10 species in each local community. We drawn randomly the 10 species from phylogenetic pool without any phylogenetic restriction (FilterLow, filter = 0 in the main text) and characterized only by species incidence.

Data analyses

Based on the simulation of three levels of phylogenetic signal and the random metacommunity we obtained the community-weighted means (CWM) for each parameter combination of phylogenetic signal. We defined the phylogeny-weighted species composition using not the phylogenetic distance between species, but a phylogenetic matrix when all species was entirely different from other species (star phylogenetic tree). We performed a Principal Coordinates Analysis in the phylogeny-weighted matrix of species composition using the square root of Bray-Curtis as the resemblance measure with the aim of calculate the PCPS. The PCPSs without phylogenetic information were used as predictors in a linear regression for modeling

the trait averages at communities to explain the proportion of the variability of traits in the metacommunity (see details in the main text). For each combination of parameters, we ran 100 simulations and evaluated the average accumulation curve.

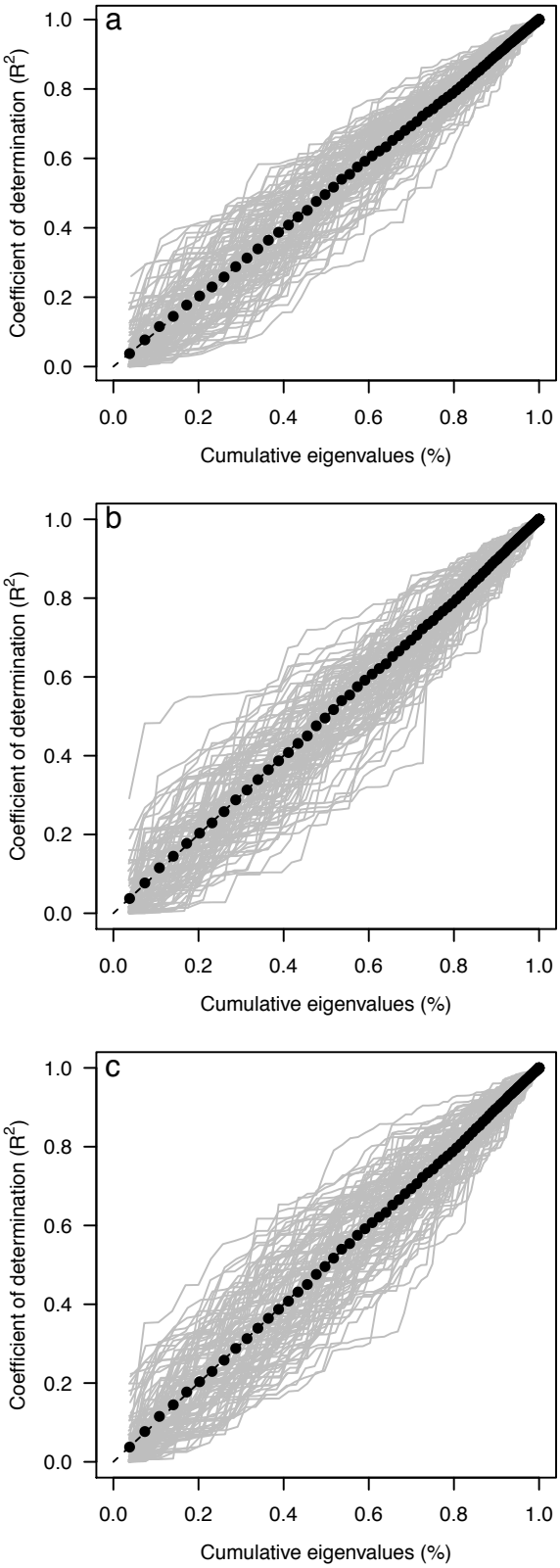

Figure S3. Curve shows the proportion of PCPS accumulation and the coefficient of determination for traits derived from regressions between CWM and sequential PCPS.

PCPS without uses any phylogenetic information. (a) Traits without phylogenetic signal (Grafen's $\rho = 0.0001$), (b) traits under Brownian motion (Grafen's $\rho = 1.0$) and (c) traits with high phylogenetic signal for the pool of species (Grafen's $\rho = 5.0$). Points in black are the mean for 100 simulations; each grew line show the observed curve for individual simulation. Each metacommunity was composed by 100 communities, these being by 10 species randomly distributed without any phylogenetic filter.

CONSIDERAÇÕES FINAIS

O conceito de conservação filogenética de nicho é importante para diversas áreas, bastante difícil de contextualizar e analisar. Vários métodos já estão disponíveis, mas nem todos estão de acordo com os conceitos utilizados em ecologia. O próprio conceito de nicho ecológico não está totalmente definido dentro do termo de conservação filogenética de nicho. Nesta tese buscamos avançar alguns pontos cruciais na tentativa de conectar melhor os processos evolutivos e ecológicos atuantes sobre as comunidades atuais. Tentamos ligar os conceitos de nicho ecológico às abordagens estatísticas, buscando desenvolver ferramentas que possam ser aplicadas aos conjuntos de dados reais, aumentando a compreensão dos sistemas biológicos.

No primeiro capítulo foi desenvolvido uma abordagem estatística para quantificar sinal filogenético usando o teste de Mantel. A novidade apresentada aqui se refere a incorporação de modelos evolutivos específicos para comparar os valores da estatística de Mantel com o esperado por determinados modelos de evolução. Essa novidade é essencial para quantificar sinal filogenético e inferir conservação filogenética de nicho, resolvendo críticas em relação ao uso do teste de Mantel. A análise das propriedades estatísticas do teste demonstrou que a abordagem proposta é bastante robusta apresentado taxas de erro tipo I compatíveis com o esperado. O teste também apresentou um bom poder estatístico, que aumenta consideravelmente com a intensidade do sinal filogenético e tipo de atributo a ser considerado na análise. De maneira geral, o teste mostrou ser um boa alternativa para medir o sinal filogenético em atributos binários e categóricos e para bases de dados com múltiplos atributos.

No segundo capítulo foi apresentado um pacote para explorar os padrões filogenéticos no nível de metacomunidade. O pacote é uma ferramenta flexível para

um dos principais programas utilizados em ecologia, permitindo aplicar os métodos de maneira fácil a vários conjuntos de dados. O pacote facilita a descrição de comunidades em termos filogenéticos e permite relacionar esses padrões filogenéticos às características funcionais e variáveis ambientais das metacomunidades. O pacote já vem sendo usado para responder importantes perguntas tanto em escala local quanto em escalas mais amplas.

No terceiro capítulo foi investigada a relação entre sinal filogenético e alguns padrões de coocorrência de espécies em metacomunidades. Além disso, foram testadas as propriedades estatísticas de um método que incorpora a coocorrência das espécies a fim de quantificar sinal filogenético no nível de metacomunidade. Foram utilizadas as propriedades estatísticas do método em relação a dois modelos nulos, os quais incorporam diferentes aspectos da estrutura das comunidades e da evolução dos atributos. Os resultados demostram que o sinal filogenético no nível do conjunto de espécies local pode ser detectado no nível de metacomunidade, havendo pouca influência dos padrões de coocorrência no padrão de sinal filogenético neste nível. Os resultados da análise das propriedades estatísticas dos modelos nulos testados mostram que ambos modelos apresentam taxas de erro tipo I adequadas. No entanto, a variação da proporção de explicação das médias dos atributos no nível de metacomunidade é bastante ampla quando os atributos seguem o modelo neutro de evolução ou são conservados em um nível intermediário, sendo difícil detectar um sinal filogenético no nível de metacomunidade acima do esperado pelo modelo de difusão neutro dos atributos nas comunidades.

Quantificar sinal filogenético e conservação filogenética de nicho são tarefas complicadas. Nessa tese buscamos relacionar os conceitos já estabelecidos em

ecologia com o desenvolvimento de novos métodos para quantificar o sinal filogenético, incorporando conceitos raramente utilizados nos métodos atuais. Superar as limitações dos métodos atuais permite compreender melhor a interação entre os processos locais, como a coocorrência de espécies e filtragem de habitat, com os processos de longo prazo, tais como a dispersão, evolução e dinâmicas de especiação e extinção. Nossa expectativa é que a incorporação de conceitos importantes nas análises ecológicas propostas por esta tese, possam contribuir de alguma forma para o desenvolvimento dos métodos utilizados para quantificar sinal filogenético e conservação filogenética de nicho.

REFERÊNCIAS BIBLIOGRÁFICAS

Adams, D. 2014. A Generalized K Statistic for Estimating Phylogenetic Signal from Shape and other High-dimentional Multivariate Data. Syst. Biol. 63:685–697.

Blomberg, S. P., T. Garland, e A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution (N. Y). 57:717–745.

Boettiger, C., G. Coop, e P. Ralph. 2012. Is your phylogeny informative? Measuring the power of comparative methods. Evolution (N. Y). 66:2240–2251.

Cavender-bares, J., K. H. Kozak, P. V. A. Fine, e S. W. Kembel. 2009. The merging of community ecology and phylogenetic biology. Ecol. Lett. 12:693–715.

Chase, J. M., e M. A. Leibold. 2003. Ecological Niches. Chicago and London: University of Chicago Press.

Cooper, N., W. Jetz, e R. P. Freckleton. 2010. Phylogenetic comparative approaches for studying niche conservatism. J. Evol. Biol. 23:2529–2539.

Duarte, L. D. S., V. J. Debastiani, M. B. Carlucci, A. V. L. Freitas, L. O. Gonçalves, Y. R. Súarez, W. Vicentin, J. Pereira e J. A. F. Diniz-Filho. (submetido). Community-weighted trait means: Should phylogeny stay or should it go? Functional Ecology.

Gause, G. F. 1934. The Struggle for Existence. Baltimore: Williams & Wilkins.

Grinnell, J. 1917. The niche relationships of the California Thrasher. Auk, 34, 427-433.

Crisp, M. D., M. T. K. Arroyo, L. G. Cook, M. a Gandolfo, G. J. Jordan, M. S. McGlone, P. H. Weston, M. Westoby, P. Wilf, e H. P. Linder. 2009. Phylogenetic biome conservatism on a global scale. Nature 458:754–6. Nature Publishing Group.

Elton, C. 1927. Animal ecology. Sidgwick & Jackson, London.

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

Filho, J. A. F. D., T. F. Rangel, T. Santos, L. M. Bini, J. A. F. Diniz Filho, e J. A. F. Diniz-Filho. 2012. Exploring patterns of interspecific variation in quantitative traits using sequential phylogenetic eigenvector regressions. Evolution (N. Y). 66:1079–1090.

Freckleton, R. P., e P. H. Harvey. 2006. Detecting non-Brownian trait evolution in adaptive radiations. PLoS Biol. 4:e373.

Fritz, S. A., e A. Purvis. 2010. Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. Conserv. Biol. 24:1042–51.

Gerhold, P., J. F. Cahill, M. Winter, I. V Bartish, e A. Prinzing. 2015. Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). Funct. Ecol. 29:600–614.

Harvey, P. H., e M. Pagel. 1991. The Comparative Method in Evolutionary Biology. Oxford: Oxford Uni versity Press.

HilleRisLambers, J., P. B. Adler, W. S. Harpole, J. M. Levine, e M. M. Mayfield. 2012. Rethinking Community Assembly through the Lens of Coexistence

Theory. Annu. Rev. Ecol. Evol. Syst. 43:227–248.

Hubbell, S.P. 2001. The Unified Neutral Theory of Biodi versity and Biogeography. Princeton, NJ: Princeton Uni versity Press.

Hutchinson, G. E. 1957. Concluding Remarks. Cold Spring Harb Symp.

Losos, J. B. 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecol. Lett. 11:995–1003.

Losos, J. B. 2011. Seeing the forest for the trees: the limitations of phylogenies in comparative biology. Am. Nat. 177:709–27.

Losos, J. B., M. Leal, R. E. Glor, K. De Queiroz, P. E. Hertz, L. Rodríguez Schettino, A. C. Lara, T. R. Jackman, e A. Larson. 2003. Niche lability in the evolution of a Caribbean lizard community. Nature 424:542–545.

MacArthur, R. H., e R. Levins. 1967. The limiting similarity, convergence and divergence of coexisting species. American Naturalist 101: 377 385.

McPeek, M. A. 2008. The ecological dynamics of clade diversification and community assembly. Am. Nat. 172:E270–84.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–84.

Pelletier, F., D. Garant, e A. P. Hendry. 2009. Eco-evolutionary dynamics. Philos. Trans. R. Soc. B Biol. Sci. 364:1483–1489.

Pillar, V. D., e L. S. Duarte. 2010. A framework for metacommunity analysis of phylogenetic structure. Ecol. Lett. 13:587–596.

Prinzing, A., W. Durka, S. Klotz, e R. Brandl. 2001. The niche of higher plants: evidence for phylogenetic conservatism. Proc. R. Soc. London. Ser. B Biol. Sci. 268:2383–2389. The Royal Society.

Pyron, R. A., G. C. Costa, M. A. Patten, e F. T. Burbrink. 2015. Phylogenetic niche conservatism and the evolutionary basis of ecological speciation. Biol. Rev. 90:1248–1262.

Revell, L. J., L. J. Harmon, e D. C. Collar. 2008. Phylogenetic signal, evolutionary process, and rate. Syst. Biol. 57:591–601.

Rosado, B. H. P., M. de S. L. Figueiredo, E. A. de Mattos, e C. E. V Grelle. 2016. Eltonian shortfall due to the Grinnellian view: functional ecology between the mismatch of niche concepts. Ecography (Cop.). 39:001–008.

Schoener, T. W. 2011. The Newest Synthesis: Understanding Ecological Dynamics. Science (80). 331:426–429.

Ulrich, W., M. Piwczyński, F. T. Maestre, e N. J. Gotelli. 2012. Null model tests for niche conservatism, phylogenetic assortment and habitat filtering. Methods Ecol. Evol. 3:930–939.

Webb, C. O., D. D. Ackerly, M. A. Mcpeek, e M. J. Donoghue. 2002. Phylogenies and community ecology. Annu. Rev. Ecol. Syst. 33:475–505.

Wiens, J. J., D. D. Ackerly, A. P. Allen, B. L. Anacker, L. B. Buckley, H. V Cornell, E. I. Damschen, T. Jonathan Davies, J. Grytnes, S. P. Harrison, B. A. Hawkins, R. D. Holt, C. M. McCain, P. R. Stephens, L. B, e T. Jonathan. 2010. Niche conservatism as an emerging principle in ecology and conservation biology. Ecol. Lett. 13:1310–24. Wiley Online Library.

Wiens, J. J., e C. H. Graham. 2005. Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. Annu. Rev. Ecol. Evol. Syst. 36:519–539.