From phylogenetic to functional originality: guide through indices and new developments

Authors: Sandrine Pavoine^{a,*}, Michael B. Bonsall^{b,c}, Amaël Dupaix^{a,d}, Ute Jacob^{e,f}, Carlo Ricotta^g

^a *Centre d'Ecologie et des Sciences de la Conservation (CESCO UMR7204), Sorbonne Universités, MNHN, CNRS, UPMC, CP51, 55-61 rue Buffon, 75005, Paris, France*

^b *Mathematical Ecology Research Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK*

^c *St Peter's College, New Inn Hall Street, Oxford. OX1 2DL, UK*

^d *Département de Biologie, ENS Lyon, 46 allée d'Italie, 69007 Lyon, France*

^e *German Centre for Integrative Biodiversity Research (iDiv), Halle- Jena-Leipzig, Leipzig, German*

^f *J.F. Blumenbach Institute of Zoology and Anthropology, University of Goettingen, Berliner Strasse 28, 37073 Goettingen, Germany*

^g *Department of Environmental Biology, University of Rome 'La Sapienza', Piazzale Aldo Moro 5, 00185 Rome, Italy*

*Corresponding author. Sandrine Pavoine; Email address: sandrine.pavoine@mnhn.fr

Contact details for the corresponding author: Muséum national d'Histoire naturelle, UMR 7204 CNRS-UPMC, CESCO, 61 rue Buffon, 75005 Paris, France; tel.: +33140793928; fax: +33140793835.

Running head: Measuring functional originalities

Number of words in abstract: 329

ABSTRACT

In biodiversity studies a species is often classified as original when it has few closely related species, a definition that reflects its phylogenetic originality. More recently, studies have focussed on biological or functional traits that reflect the role(s) that species play within communities and ecosystems. This has led many studies to an alternative evaluation of species' originality: its functional originality. Most indices of species' originality were developed to treat the hierarchical structure of a (phylogenetic) tree. The change in perspective from measures of phylogenetic originality to measures of functional originality thus raises methodological issues particularly around the need to develop indices explicitly appropriate for evaluating functional trait-based originality. We compare indices of species' originality including a new index which we develop to evaluate 1) whether phylogenetic originality could serve as a proxy for functional originality in conservation and ecological studies; 2) whether the transformation of functional data into functional trees modifies the way species are ranked according to their originality measures compared to approaches that directly rely on pairwise functional dissimilarities among species; and more generally, 3) whether different indices provide different views on how original species are from each other, hence reflecting different ecological and evolutionary processes that generated patterns of originality. Using simulations and a real case study, we show that: 1) the strong effects of the choice of a clustering approach can affect reported levels of dissimilarities among species; 2) the tree-based approaches could better reflect the trait-generating processes under constant (Brownian) rates of evolution; and 3) phylogenetic originality measures can depart from functional originality measures when species have large amount of independent evolution. Overall, phylogenies may be used at large scales but cannot replace functional approaches designed for depicting community assembly. Indeed, traits involved in ecological processes may have various histories and thus moderate phylogenetic signals. Our comparative study

51 provides approaches and perspectives on the analysis of originality across biological scales of
52 organization from individuals, through populations, up to the originalities of communities and
53 regions.

54

55 *Keywords:*

56 Biodiversity; carnivores; conservation priorities; entropy; simulations; species distinctiveness

1. Introduction

Atkinson (1989) recommended that "given two threatened taxa, one a species not closely related to other living species and the other [a] widespread and common species, it seems reasonable to give priority to the taxonomically distinct form". May (1990) and Vane-Wright et al. (1991) therefore developed equations to measure how taxonomically distinct a species is compared to a reference set of species. A species was then defined as distinct if it is not closely related to other living species, a concept also known as evolutionary isolation (Jensen et al. 2016). Following Faith (1992), Pavoine et al. (2005) extended the concept of the isolation of a species on a phylogenetic tree to that of originality. They defined originality as the potential rarity of the species' features, where a feature means a particular state of a character. They also considered 'strict uniqueness' as the number of features possessed by this species yet not those shared with the others. Recently, there have been more studies directed on the functional attributes of species: a finite number of physiological, anatomical, behavioural or life-history traits reflecting the roles that species play within communities and ecosystems (e.g. Petchey et al., 2007; Mouillot et al., 2008; Magnuson-Ford et al., 2009; Schmera et al., 2009a; Thompson et al., 2010; Buisson et al. 2013; Mouillot et al., 2013; Godet et al., 2015; Rosatti et al., 2015).

This diversity of approaches led different authors to use the terms distinctiveness, originality and uniqueness in different meanings. Sometimes two expressions have been used to designate the same concept and sometimes a single word was used to mean two different things. Also in the literature there is confusion between the concepts and the methods used to associate quantitative measures to these concepts. Notably originality was used by Pavoine et al. (2005) to design a concept. It was then used again by Buisson et al. (2013) to designate a

measure: the distance, in a functional space, between a species position and the centroid of the space. Generalizing Buisson et al. (2013) framework, Redding et al. (2014) used the expression "originality" to designate the average phylogenetic (patristic) distance to all other species. The fact of being taxonomically distinct was introduced by Atkinson as a concept but "evolutionary distinctiveness" is often used to name an index also known as the "Fair Proportion" measure (Isaac et al., 2007; Jensen et al., 2016; see also Table 1). The concept associated with "evolutionary distinctiveness" was instead often referred to as evolutionary isolation (Redding et al. 2014).

Hereafter we use originality as the core, unifying concept and strict uniqueness as a special case. As a proposal for a unified semantic framework, we define the originality of a given species in a set of species as the rarity of its biological characteristics. Originality can emanate from any characteristics of the species. Notably, it can integrate the evolutionary history (phylogenetic originality) or the functional traits (functional originality) of species. This definition generalizes the definition Pavoine et al. (2005) initially proposed. We consider originality synonymous to the following expressions: distinctiveness (e.g. Atkinson, 1989); isolation (e.g. Redding et al., 2014); degree of uniqueness (e.g. Brooks et al., 2015; Ricotta et al., 2016). We consider originality antonymous to the concept of redundancy (e.g. Buisson et al., 2013; Ricotta et al., 2016). We define strict uniqueness as the minimum difference with any other species in a set. We consider strict uniqueness as a special case of originality. From a biodiversity perspective, strict uniqueness is the amount of diversity that is solely supported by the focal species (driven by unshared characteristics of the species). In contrast, originality is the full contribution of the species to the biodiversity of the set (Pavoine et al. 2005).

Both phylogenetically original and endangered taxa have recently been the focus of conservation actions (Isaac et al., 2007). Depending on the shape of the phylogenetic tree (imbalance and 'tippiness', Heard and Mooers, 2000), the loss of entire species-poor clades

that contain original species could indeed lead to dramatic loss in taxonomic/phylogenetic diversity (Purvis et al., 2000). In contrast, as far as we are aware, very few conservation actions have focused on functionally original and endangered species. Yet, Mouillot et al. (2008), for example, found that protecting the most functionally original species protects high functional fish diversity in the Bonifacio Strait Natural Reserve. In food webs, intermediate species (herbivores) that tend to be more trophically original (they share no or few prey and predators with other species) might be more prone to secondary extinctions. Furthermore, their loss might have great effects on trophic diversity due to their relative originality (Petchey et al., 2008). The concept of species originality has also been studied in ecology and associated with key ecological processes, such as community assembly, ecosystem functioning, and species extinction. Original species could be more likely to invade or colonize, and in addition may have less impact on resident species (Strauss et al., 2006; Strayer et al., 2006). Species original in their functional traits might make a large contribution to ecosystem functions and services, such as gross photosynthetic rate (Petchey et al., 2004). Unique functions of original species in their ecosystems reinforce the importance of originality indices for conservation biology. Developing and comparing measures of originality is thus critical for their efficient use in conservation.

As highlighted above, species originality has been primarily measured from phylogenetic trees. Consequently the methods that were first developed to measure species originality from a phylogeny are now being adapted and applied to the analysis of functional traits. This translation raises new issues on the measurement of functional originality. Indices of phylogenetic originality rely on the tree structure of the phylogeny. Adapting these indices to functional originality thus requires the definition of functional trees (or dendrograms) with a risk of distorting the information provided by functional traits. This is exemplified with the clustering approach used to define the functional tree (Mouchet et al., 2008; Petchey et al.,

2009). Among the indices of phylogenetic originality, the quadratic entropy(QE)-based index developed by Pavoine et al. (2005) was defined for (ultrametric) phylogenetic trees, where the distance from tips to root is constant, which also is a property of functional trees obtained by clustering methods. Here, we extend this QE-based approach to any (phylogenetic or functional) dissimilarity matrix among species. We compare these originality indices related to QE to a range of existing indices introduced in the literature in their ability to discriminate species in terms of their functional originality (Table 1; May, 1990; Eiswerth and Haney, 1992; Redding, 2003; Ricotta, 2004; Redding and Mooers, 2006; see also Redding et al., 2014 for a review). We selected originality indices amongst the most used in the literature. We use numerical simulations and a case study to evaluate the strengths and differences across the range of originality indices. In particular, we evaluate

1) whether measures of phylogenetic originality could serve as a proxy for measures of functional originality in conservation and ecological studies;

2) whether the transformation of functional data into functional trees modifies the way species are ranked according to their measured originalities compared to approaches that directly rely on pairwise functional dissimilarities among species;

3) more generally, whether different indices provide different views on how original species are from each other, hence reflecting the range of ecological and evolutionary processes that generated patterns of originality.

We discuss our results in light of recent developments in the assessment and measurement of a multidimensional view of biodiversity (Pavoine and Bonsall 2011), with the aim to identify (and preserve) the ecological and historical processes that drive biodiversity dynamics.

Table 1

Originality indices discussed in this paper. All indices are measures of originality; but only *PE* and *NN* are measures of strict uniqueness.

Short name	Full name	Dependence on		Reference
		a tree structure	a dissimilarity matrix	
<i>AV</i>	Average distance to other species		X	Eiswerth and Haney, 1992
<i>ES</i>	Equal-Split (branches in a tree are split equally among descending clades)	X		Redding and Mooers, 2006
<i>FP</i>	Fair Proportion (branches in a tree are split fairly among descending species)	X		Redding, 2003
<i>M</i>	May's topological index (number of branches emerging from internal nodes in the path between a species and the root of a tree)	X		May, 1990
<i>NN</i>	Distance to the nearest neighbour		X	This paper
<i>PE</i>	Pendant Edge (terminal branch of a tree)	X		Redding et al., 2014
<i>Qb</i>	Species' proportions that maximize the quadratic entropy diversity index	X	X	Pavoine et al., 2005
<i>Rb</i>	Species' proportions that maximize the <i>R</i> diversity index		X	This paper
<i>tb-AV</i>	<i>AV</i> index applied on tree-based distances among species	X	X	This paper
<i>tb-Rb</i>	<i>Rb</i> index applied on tree-based distances among species	X	X	This paper

2. Methods

2.1. A variety of originality indices

2.1.1. *Phylogenetic originality indices*

Vane-Wright et al. (1991) were probably the first to define a cladistic (taxonomic) measure of originality. Their measure was defined as inversely proportional to the number of internal nodes between the focal species (tip) and the root of the cladistic tree. An improvement to this originality measure (discussed in an earlier study by May, 1990) has considered the effects of unresolved nodes by counting not simply the number of nodes between tip and root but rather the number of branches descending all such nodes. This improved index of species' originality can be applied to any phylogenetic tree. Hereafter we will refer to it as index *M*.

More recently, other measures have been suggested that consider branch lengths on the phylogenetic tree (Table 1; Fig. 1):

i) The pendant edge (*PE*, e.g. Redding et al., 2014) index is defined as the length of the branch that connects a species to the rest of the tree.

ii) The fair proportion index (*FP*, Redding, 2003) distributes the phylogenetic diversity (sum of branch lengths) contained within a tree uniquely among the species at the tips. This is achieved by dividing the shared evolutionary history represented by a branch equally among its *daughter species* at the tips.

iii) The Equal-Split index (*ES*, Redding and Mooers, 2006) also distributes the phylogenetic diversity contained within the tree uniquely among the species at the tips. However, it achieves this by dividing the shared evolutionary history represented by a branch equally among its *daughter branches*.

Insert Fig. 1. Single-column fitting image

2.1.2. *Dissimilarity-based originality indices.*

Considering a set of N species, Eiswerth and Haney (1992) and Ricotta (2004) suggested an alternative measure that allowed genetic and taxonomic distances to be included directly in measures of originality. It consists of computing pair-wise genetic or taxonomic distances among species and then obtaining the average distance between a focal species and all others ($N-1$ species) in the set. We refer to this measure as the average distance index (AV). It is clear that this index has broader applications: it can be applied to any distances among species be they genetic, taxonomic, phylogenetic, or functional. Index AV is related to Schmera *et al.* (2009b) functional value of a species (FV): $FV=[(N-1)*AV]/N$. The conclusions we obtain below for AV also applies to FV because these two indices similarly order species from the least to the most original one. Below, we compare AV to the shortest distance between the focal species and all others in the set (hereafter referred to as the 'nearest neighbour' index, NN).

2.1.3 The special case of the QE-based index framework

To define the QE-based index (hereafter referred to simply as Qb), Pavoine et al. (2005) considered a matrix $\mathbf{D}=(d_{ij})_{1 \leq i \leq N, 1 \leq j \leq N}$ of dissimilarities among N species. The dissimilarity d_{ij} between any two species i and j was calculated on an ultrametric, phylogenetic tree as the sum of branch lengths between each of them and their most recent ancestor. These dissimilarities are ultrametric as they satisfy the following property: $d_{ij} \leq \max(d_{ik}, d_{jk})$ for all i, j, k .

Let $\mathbf{p}=(p_1, ..., p_i, ..., p_N)$ be a vector of species' relative abundance ($\sum_{i=1}^N p_i = 1$) in an hypothetical assemblage. The average dissimilarity among individuals from this assemblage is an established measure of diversity (Rao, 1982; Pavoine et al., 2009):

$$Q(\mathbf{p}, \mathbf{D}) = \sum_i^N p_i \sum_j^N p_j d_{ij} .$$

The index Qb of species' originality is the value of p_i a species i should have to maximize Q (Table 2). The value of p_i that leads to the maximum possible value of Q reflect the originality of species i . Indeed, to a certain extent, the more abundant original species are compared to redundant species, the more diversity there is in a set of species. However, this index Qb is critically dependent on the use of dissimilarities derived from an ultrametric tree (Table 2).

Table 2
Details on how to calculate the originality indices Qb and Rb .

Diversity index	Originality index
Q	<p>Consider that $\mathbf{D}$ is a matrix of patristic distances between species derived from a phylogenetic or functional tree (Fig. 2). The value of p_i that lead to the maximum possible value of Q applied to $\mathbf{D}$ is equal to</p> $Qb_i = \sum_j^N \delta_{ij} / \sum_i^N \sum_j^N \delta_{ij}$ <p>where δ_{ij} is the value at the ith row and jth column of the inverse of $\mathbf{D}$. Pavoine et al. (2005) restricted Qb to ultrametric dissimilarities in $\mathbf{D}$. Indeed without this property, the vector $\mathbf{p}$ that maximizes Q may not be unique and could take null values (zeros) for several species.</p>
R	<p>Consider a given matrix $\mathbf{D}=(d_{ij})_{1 \leq i \leq N, 1 \leq j \leq N}$ of dissimilarities among N species with the only conditions that $d_{ij}=d_{ji}$ for any i,j, $d_{ij}>0$ for any $i \neq j$, and $d_{ii}=0$ for any i. The maximum of R over $\mathbf{p}$ is the first eigenvalue of $\mathbf{D}$. Its unique maximizing vector is the squared first eigenvector of $\mathbf{D}$ (proofs in Appendix A). This maximizing vector is our index Rb.</p>

2.2. The new index

Originality can thus be defined as the amount of abundance a species should have in a theoretical community to provide the maximally, theoretically possible biodiversity to this community. Such a reasoning however means that the measurement of species originality depends on how biodiversity itself is measured. In their study, Pavoine et al. (2005) used Q as the reference biodiversity index. A particularity of Q is that it preferentially weights common species over rare species. However, recent studies have focused on profiles of diversity indices where rare versus common species are weighted with more or less importance (Jost, 2006; Pavoine et al., 2009; Chao et al., 2010; Leinster and Cobbold, 2012). Changing the relative importance given to rare versus common species in functional or phylogenetic diversity measurements has an impact on the amount of abundance a species should have to provide the maximally, theoretically possible biodiversity to a community.

Most of these studies rely on the Hill numbers (Hill, 1973):

$${}^qD(\mathbf{p}) = \left[\sum_i^N (p_i)^q \right]^{\frac{1}{1-q}}, \quad q > 0, q \neq 1$$

where q modifies the relative importance given to rare versus common species: the sensitivity of qD to rare species decreases with q (Patil and Taillie, 1982). The limiting case ($q \rightarrow 1$) serves as a reference where species are weighted directly by their relative abundances (p_i 's) without favoring either rare or common species (Jost, 2006). In the case where the d_{ij} 's are defined in the range $[0,1]$, $1/(1-Q)$ is a generalization of 2D that includes (functional or phylogenetic) distances among species (Ricotta and Szeidl, 2009). The fact that the quadratic diversity (Q) gives high weight to abundant species compared to rare species can be viewed as a weakness when rare species are considered as key drivers of biodiversity of conservation

interest. Developing this, we propose an index to contrast quadratic diversity (Q) by weighing rarity over commonness (see Appendix A):

$$R(\mathbf{p}, \mathbf{D}) = \sum_{i=1}^N \sum_{j=1}^N \sqrt{p_i} \sqrt{p_j} d_{ij}$$

In the particular case where the d_{ij} 's are defined in the range $[0,1]$, $R-1$ is a generalization of Hill number $^{0.5}D$ that includes functional or phylogenetic dissimilarities among species (Appendix A).

Our new index, R , complements Q , both by the similarity of their formulas and by their relatedness to the Hill numbers. Both these indices (Q & R) are anchored within the broader literature on entropy, species diversity, and functional and phylogenetic diversity (see details in Appendix A). For example, when species have even relative abundances, i.e. $p_i=1/N$ for all i , then R equals $MFAD$, an index of diversity introduced by Schmera et al. (2009a):

$$MFAD(\mathbf{D}) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N d_{ij} .$$

As for index Q , the values of p_i that lead to the maximum possible value of R also reflect species originalities. Increasing the abundance of original species compared to redundant species to a certain extent increases the diversity of the set of species (see Appendix A for a simple example). Using the vector of p_i 's that maximizes R , instead of Q , should however provide an alternative view on species' originalities (Table 2). We refer to this vector as Rb . Compared to Qb , an advantage of Rb is that it can be applied to any dissimilarities (even those that are not ultrametric). However, whether Rb provides novel interpretation to species' originality will be analyzed below.

2.3. Case studies

We use both simulation studies and an empirical dataset (on European carnivores) to investigate the measurement of functional originality. For each case (simulation or empirical dataset), we computed all the indices introduced in Table 1 as specified in Fig. 2. We calculated functional dissimilarities among species from raw data using the Euclidean distance metric with quantitative traits in the first case study; and Gower distance (Gower, 1971) with a mix of quantitative and nominal traits in the second case study. We obtained functional trees from functional distances using UPGMA and Ward methods (R function hclust, parameter "average" and "ward.D2", respectively; R Core Team 2016).

Insert Fig. 2. Single-column fitting image

2.3.1. Simulations

For the simulation case study, our objective was to analyze functional originality when traits evolved under a Brownian motion model. Indeed, the Brownian model of trait evolution assumes constant rates of trait changes through time. Under this model, trait-based distances between species are expected to be strongly correlated with phylogenetic distances. We thus chose the Brownian model to evaluate whether, in this extreme, simplified scenario, measures of phylogenetic originalities effectively acted as proxies for measures of functional originalities. If phylogenetic and functional originalities are different even when traits are simulated under a Brownian model of evolution, then the use of phylogenetic originality measures as proxies for functional originality measures would be poor. We also wanted to evaluate whether the correlations between phylogenetic originality measures and functional originality measures depended on methodological choices such as the originality index

chosen, the method used to obtain functional dendrograms, the number of traits considered. Our a priori hypothesis was that, using *Rb*, *AV* and *NN*, functional originality measures would better reflect phylogenetic originality measures. Indeed, these indices are calculated using raw functional data with minimal methodological assumptions and thus minimal data distortion. In contrast, functional dendrograms may distort information obtained from raw traits. Originality indices based on functional dendrograms may thus depart more from the original signal driven by phylogenies.

We simulated four general phylogenetic trees: (i) a pure birth model (BIRTH model, with birth rate of 0.1) leading to relatively well-balanced trees (function “sim.bd.taxa” in the R package TreeSim; Stadler, 2015); (ii) trees with speciation events close to the root (ROOT model); (iii) trees with speciation events near the tips (TIPS model) (using package geiger in R - function “deltaTree” with $\delta = 10$ and 0.1, respectively, Harmon et al. 2008); and finally (iv) asymmetric nonultrametric trees (NU model, where the distance from tips to root is not constant). In the asymmetric trees the topology was generated by splitting randomly the edges (function “rtree” in *ape* package in R, Paradis et al., 2004) and branch lengths were simulated using a log-normal distribution (with $\text{LogN}(0,1)$). Nonultrametric trees represent unequal evolutionary rates in different parts of the phylogeny. They led to asymmetric distributions of species' originalities with many redundant species and a few original ones. We simulated 1000 trees per model with $2^7=128$ tips. One, 25 or 50 traits were simulated for each phylogenetic tree, according to a Brownian model with parameters $\sigma=1$ and $\theta=0$ (function “rTraitCont” in *ape* package).

Our analysis proceeds by first analyzing the effect of the clustering approach (either UPGMA or Ward) on functional originalities using Spearman correlations for each originality index. Next we analyzed the Spearman correlations between indices and detailed the correlations between *Rb* and the other indices calculated with functional data (using one, 25 or

50 traits) for each different trait evolution modes (BIRTH, ROOT, TIPS, or NU models of trait evolution) under different functional (UPGMA vs Ward) dendrograms. For tree-based metrics, correlations with *Rb* were analyzed with centred Principal Component Analysis using the number of traits, the model of trait evolution and the method for tree construction as supplementary variables (using package 'ade4' in R; Dray and Dufour, 2007). Then we analyzed the Spearman correlations between phylogenetic originality measures and functional originality measures for each originality index and simulation case. We used Spearman correlations in all cases as the high number of simulations restricted checking the shape of the relationships between phylogenetic and functional originality measures for each scenario, each originality index and between originality indices.

2.3.2. Functional and phylogenetic originalities of carnivore species in Europe

We investigated the phylogenetic and the functional originalities of European carnivores. We based our study on Temple and Terry (2007) who identified 38 European Carnivora species. Among these species, four were introduced into Europe after 1500 A.D. (*Herpestes javanicus*, *Neovison vison*, *Nyctereutes procyonoides* and *Procyon lotor*), and seven were defined as marginal in Europe (*Cystophora cristata*, *Erignathus barbatus*, *Martes zibellina*, *Odobenus rosmarus*, *Pagophilus groenlandicus*, *Felis chaus* and *Mustela sibirica*). The qualification of marginal occurrence was attributed to these latter species as less than 1% of their population or of their range lies in Europe (Temple and Terry, 2007). We took phylogenetic data from the Carnivora phylogeny established by Nyakatura and Bininda-Emonds (2012). Trait data were obtained from the PanTHERIA database (Jones et al., 2009). Among the 53 traits present in the database, we eliminated those for which values were missing for more than a

quarter of the species. To avoid circularity in our reasoning, we also discarded geographic traits. Indeed, we made a difference in originality between marginal and non-marginal species, marginality being defined by the geographic distribution of species. Among remaining variables, we removed missing values by using alternative data (MacDonald, 2009; Myers et al., 2015). As the trait database was still incomplete (due to missing values) we could not analyze all traits individually. We thus combined traits for this illustration, in order to measure the global functional originality of each species. We provide however in Appendix B a guide on how to deal with individual traits, particularly when some species have identical values for those traits. We discuss also in Appendix B on potential impact of trait selection.

Before combining the traits we removed redundancy by excluding six of the traits with high correlations with body mass (Appendix B). We ended with 9 traits with only 6 missing values: activity cycle (3 attributes: nocturnal only, diurnal only, others), adult body mass, age at eye opening, diet breadth, habitat breadth, inter-birth interval, terrestriality (2 attributes: fossorial and/or ground dwelling only versus above ground dwelling), trophic level (3 attributes: herbivore, omnivore, or carnivore) and weaning age. Body mass, as a 3-dimensional quantitative trait, was log-transformed to avoid extreme originalities. Indeed, we expected extreme originalities to occur because of species with high body mass such as *Odobenus rosmarus* (walrus). The complete set of functional traits retained for this study is given in Appendix C.

We evaluated phylogenetic signal in traits using Mantel correlations for the combined traits (Hardy and Pavoine, 2012; Pavoine and Ricotta, 2013); for individual quantitative traits we used Blomberg et al. (2003) K^* and Pavoine and Ricotta (2013) K_w while for individual nominal traits we used Maddison and Slatkin (1991). We first analyzed the effect of the clustering approach (either UPGMA or Ward) using Spearman correlations for each originality index. Next, we analyzed the Spearman correlations among indices calculated with

phylogenetic and functional data. Then, we analyzed the Spearman correlations between phylogenetic originality measures and functional originality measures for each originality index.

We complemented these analyses with some more specific questions related to the conservation of Carnivora species. We distinguished introduced from native species. Among natives, we distinguished species whose occurrence is marginal in Europe. We first assessed which species were more original using the median of originality values for each index and group of species. For non-marginal native species only, we used an index of extinction risk defined by the International Union for Conservation of Nature (IUCN) European Red List where species are ranked as follows: 1=Least Concern, 2=Near Threatened, 3=Vulnerable, 4=Endangered and 5=Critically Endangered (IUCN 2015). We calculated a Spearman correlation between the (functional or phylogenetic) originality and the extinction risk index.

3. Results

3.1. Simulations

When only one trait was used the correlations between measured functional and phylogenetic originalities were always low (Fig. 3). When the number of traits increased, correlations increased. They were especially high with *PE* and related indices (*ES* and *FP*, Redding et al., 2014) when the pure birth model was used to simulate phylogenies (with or without the speciation events moved towards tips). When the speciation events were concentrated close to the root only the three indices *PE*, *ES* and *FP* led to correlations between measured functional and phylogenetic originality. The impact of the method used for tree construction (i.e. Ward versus UPGMA) on originality depended on the model used to

simulate the phylogeny on which traits evolved. Correlations varied from ≈ 0.30 with *M*, *Qb*, *tb-AV*, and *tb-Rb* when 50 traits were simulated with the ROOT model to ≈ 1.00 with *PE*, *ES* and *FP* (details in Appendix D). In general, correlations were higher with UPGMA than Ward trees. When the simulated phylogenetic trees were not ultrametric (NU model), *AV* and *Rb* led to the highest correlations between functional and phylogenetic originalities. UPGMA and Ward transformations on functional distances thus decreased the connections between the information within functional traits and the model used to simulate them.

Insert Fig. 3. 1.5-column fitting image

On average, the highest correlations (median > 0.90) among originality indices were between *AV* and *Rb*, between *tb-AV* and *tb-Rb*, between *Qb* and *M*, and between *PE* and *NN* (Table 3). High correlations (> 0.70) were also obtained between *Qb*, *ES* and *FP*, between *ES*, *FP* and *PE*, and between *AV*, *Rb*, *tb-AV* and *tb-Rb*. Correlations between our new index *Rb* and the other originality indices were moderate to high, with close-to-1 correlations with *AV* (Fig. 4). The link between the square root of *Rb* and *AV* was close to linear (Fig. 4a). The lowest correlations with *Rb* were obtained with *PE*, *NN*, and *ES* (Table 3 and Fig. 4b; see also Appendix E for more details on these correlations).

Correlations between *Rb* and tree-based metrics tended to be higher when a single trait was used and when traits were simulated using the *NU* model (Fig. 4c). The increase in correlation with only one trait was especially high with *PE*, and related indices (*FP* and *ES*) (Fig. 4c, PCA Axis 1). Correlations of *Rb* with *PE*, *ES* and *FP* were higher with Ward trees and when speciation events were moved close to the root (compared to the pure birth model),

whereas those with the *tb-AV*, *tb-Rb*, *Qb* and *M* were higher with UPGMA and when the speciation events occurred close to tips (Fig. 4c, PCA Axis 2).

Table 3

Median correlations among indices (correlations calculated with phylogeny, traits, and the two clustering approaches) below the diagonal for the Simulation case study and above the diagonal for the Carnivora case study.

	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV	Rb	NN
M		0.60	0.88	0.74	0.92	0.63	0.63	0.60	0.62	0.65
PE	0.36		0.83	0.84	0.77	0.56	0.56	0.60	0.62	0.96
ES	0.69	0.78		0.96	0.95	0.75	0.75	0.75	0.75	0.85
FP	0.61	0.76	0.89		0.96	0.84	0.84	0.84	0.84	0.81
Qb	0.93	0.52	0.79	0.78		0.83	0.84	0.76	0.78	0.80
tb-AV	0.69	0.23	0.42	0.55	0.77		1.00	0.93	0.90	0.55
tb-Rb	0.70	0.24	0.43	0.56	0.78	1.00		0.93	0.90	0.55
AV	0.55	0.24	0.42	0.54	0.62	0.72	0.73		0.99	0.59
Rb	0.54	0.24	0.42	0.54	0.62	0.70	0.71	1.00		0.62
NN	0.36	0.99	0.77	0.78	0.53	0.25	0.25	0.25	0.25	

Insert Fig. 4. 1.5-column fitting image

3.2. Functional and phylogenetic originalities of carnivore species in Europe

Compared to the simulation case study, the phylogenetic signal in functional traits was not known a priori. Although phylogenetic signal was significant in all traits except activity cycle (see details in Appendix B), we found low Mantel correlation between phylogenetic and functional distances when using all combined traits ($r=0.08$, $P=0.071$). As expected from the low Mantel correlation, correlations between the functional and the phylogenetic originalities were also low (Table 4a). The correlation between the originality values obtained from the two methods of tree construction (Ward and UPGMA) were high (Table 4b), which suggests that the chosen clustering method had only a small impact on the originality estimates in this case study. The highest correlations among originality indices were between *FP*, *Qb*, *AV*, *Rb*, *tb-AV* and *tb-Rb* (Table 3). More precisely, our new index *Rb* was correlated with all indices in our case study but particularly strongly with *AV* (Table 3).

Introduced species generally had higher phylogenetic and functional originalities than native species (Fig. 5). Among native species whether the marginally European or the other species were the least phylogenetically and functionally original depended on the index used. There was no evidence for a correlation between the phylogenetic originality of native species and their extinction risk (Table 4c). Correlations with functional originality were higher but still moderate (Table 4d). The highest correlations were obtained with *PE* and *NN*, both characterizing strict functional uniqueness. In this case study, using the clustering approaches decreased the correlation between strict uniqueness and extinction risk and its significance ($P=0.068$ using *PE* but $P=0.039$ with *NN*). All other correlations were not significant ($P>0.1$).

Table 4

Spearman correlations obtained with the Carnivora data set: a) between originality measures obtained using phylogenetic data and those obtained using functional data; b) between functional originality measures obtained with Ward algorithm indices and those obtained with UPGMA; c) between phylogenetic originality measures and IUCN status; d) between functional originality measures and IUCN status. When functional originality is measured, we indicated for each row of the table whether we used raw dissimilarities (noted 'Raw'), Ward trees, or UPGMA trees (see Fig. 2). The notations "(PE)NN", "(tb-)AV" and "(tb-)Rb" mean that we applied indices PE, tb-AV and tb-Rb to trees (phylogenies, Ward and UPGMA functional trees), and indices NN, AV and Rb to raw dissimilarities (see Fig. 2).

	M	(PE)NN	ES	FP	Qb	(tb-)AV	(tb-)Rb
a) cor(phylogenetic originality, functional originality)							
Raw	-	0.19	-	-	-	0.11	0.06
Ward	0.06	0.14	0.30	0.26	0.20	0.21	0.21
UPGMA	-0.19	0.15	0.30	0.24	0.20	0.19	0.19
b) cor(Ward functional originality, UPGMA functional originality)							
Ward-UPGMA	0.72	1.00	0.92	0.99	0.94	0.96	0.96
c) cor(phylogenetic originality, extinction risk)							
Phylogeny	0.01	0.14	0.14	0.08	0.05	-0.12	-0.13
d) cor(functional originality, extinction risk)							
Raw		0.40				0.26	0.29
Ward	0.24	0.36	0.32	0.26	0.27	0.12	0.13
UPGMA	0.18	0.36	0.24	0.27	0.22	0.14	0.14

457 *Insert Fig. 5. Single column fitting image*

458
459 **4. Discussion**

460
461 *4.1. Is phylogenetic originality a proxy for trait-based originality?*

462
463 The concept of species originality for conservation has been effectively applied to define
464 conservation priority most successfully through the Evolutionarily Distinct and Globally
465 Endangered (EDGE) program (<http://www.edgeofexistence.org/conservation/>; Isaac et al.,
466 2007). Within the EDGE program, Isaac et al. (2007) proposed focusing on those species that
467 are both phylogenetically original and threatened with extinction. To date, mammals,
468 amphibians, birds and reef coral species have been assessed through this method (Isaac et al.,
469 2007, 2012; Collen et al., 2011; Huang, 2012; Jetz et al., 2014). The approach is known to
470 depend only on a single originality index (*FP*) and different priority schemes would be
471 expected when using other indices (Redding et al., 2014). The use of phylogenies to measure
472 species originality was justified by the assumption that phylogenetic originality reflects how
473 many character states (including observed and unobserved characters) species share (Faith,
474 2002; Pavoine et al., 2005; Isaac et al., 2007; Cadotte et al., 2010).

475 Using a phylogeny to predict diversity has been particularly important for establishing
476 broad-scale prioritizing schemes in conservation biology (e.g. Barker, 2002; Isaac et al.,
477 2007). In more fundamental ecological studies, for example, Strauss et al. (2006) showed that
478 invaders are more likely to be successful if they are phylogenetically distinct from the natives
479 in the assemblage. According to Strauss et al. (2006), not specifying traits might be useful, in

this context, because of the large diversity of mechanisms that underpin invasion success. However, in community or evolutionary ecology, we might be interested in an identified, finite set of traits that confer functions to species, in relation, for instance, with biotic interactions or abiotic filters (Hooper et al., 2005; Violle et al., 2007). Variation in these focal traits may be more or less predicted by the phylogenetic distances between species (Losos, 2008; Faith, 2015b) depending on the traits.

In our simulations of Brownian trait evolution, correlations between phylogenetic and trait-based strict uniqueness (as measured by *NN* and *PE*) were generally high. However, correlations obtained with other originality indices varied more according to the type of phylogenetic tree simulated and according to the clustering approach applied to traits. Our simulations suggest that a topological index (*M*) should be avoided when the aim is to use phylogenetic originality as a proxy for functional originality. However, it should be noted that the applicability of alternative indices depends on the quality of branch length estimations.

While we might expect that strict phylogenetic uniqueness can act as a proxy for strict functional uniqueness, more generally phylogenetic originalities may rarely act as proxies for trait-based originalities in real case studies. We obtained moderate to high correlations between phylogenetic and trait-based originalities only when several traits were simulated. This latter point supports previous arguments that phylogeny-based approaches in ecology are integrative: they are applicable in a macroevolutionary context to model many traits or features (Webb et al., 2002). They are less likely to be applicable to specific species' functions associated with community assembly (e.g. Gerholds et al., 2015). We indeed found low correlations between phylogenetic and functional originalities in the real case study where traits revealed low (although significant) phylogenetic signals. Functionally but not phylogenetically strictly unique Carnivora species tended to be the most threatened in Europe. These species were notably the nearly threatened *Lutra lutra* (European otter), the vulnerable

Ursus maritimus (polar bear) and *Gulo gulo* (wolverine), and the critically endangered *Mustela lutreola* (European mink) and *Monachus monachus* (Mediterranean monk seal).

4.2. How to choose an index of functional originality?

In some respects the problem of choosing an index of functional originality is a simple extension of choosing an index of phylogenetic originality. However, indices of functional originality are more complex in that traits do not have the intrinsic tree structure on which most current originality indices rely. Our case study on European carnivores highlights that the choice of an originality index can alter the conclusions of a study such as whether original, native species are threatened with extinctions. We found low correlations between part of the indices in the simulation and Carnivora case studies. There is thus no one single way to define and codify originality. This highlights that the choice of an originality index is critically dependent on the research question under scrutiny.

According to Redding et al. (2014), many originality indices depend on two basic measures: the strict uniqueness (as measured by PE or NN) and the average distance to all other species in a set (as measured by tb-AV or AV). As Faith (2008) highlighted, strict uniqueness might be too restrictive as a measure of originality: the originality of the focal species also depends on how many species share its evolutionary history (and/or its functional trait states). For instance, Strauss et al. (2006) found that “the presence or absence of multiple, closely related species and not just a single most closely related species may more effectively limit the success of an invader once it has become established”. That said, the originality indices most correlated with species' extinction risks in European carnivores were those

measuring strict functional uniqueness. Strict uniqueness might thus be an important criterion to consider when defining priorities of conservation. However, the generality of this result requires further work controlling for the traits selected to measure functional originalities, the taxa considered and the spatial scale.

Redding et al. (2014) thus considered the average distance to all other species in a set (AV or $tb-AV$) as the second basic component of species' originalities. Although derived from the same theory as Qb , our case studies clearly highlight the high correlations between the new index introduced in this paper (Rb) and AV . The advantage of AV and the new index Rb over alternative, tree-based indices is that they can both be applied to any dissimilarity matrix without the need to transform them into ultrametric distances, or into trees. An advantage of AV over Rb is the simplicity of its formula. A drawback of AV however, identified by Pavoine et al. (2005), is that it provides very close values for all species. Rb being linearly related to the square of AV better discriminates among the species: it better distinguishes species with high originality from species with close-to-zero originalities.

The other indices discussed in this paper combine information on strict uniqueness and the average distance to all other species in a set (Redding et al., 2014). When applied to functional dissimilarity, all these indices rely, however, on the use of clustering approaches. The question of choosing an originality index for functional data might then be simplified into whether strict uniqueness or average distance is most important for the question at hand, and whether transforming data by a clustering approach is reasonable (i.e., does the transformation allow information in traits to be retained?).

Our hypothesis was that Rb , AV and NN calculated directly on trait dissimilarities should better reflect phylogenetic originalities than indices that require functional trees. This hypothesis failed when traits were simulated from ultrametric trees but was confirmed when

traits were simulated from non-ultrametric trees. In the Carnivora case study, correlations between *Rb* and *tb-Rb* (and similarly between *AV* and *tb-AV*) were high when analyzing all traits together; however these correlations may vary with the traits considered (Appendix B). Only with ultrametric phylogenies and Brownian evolution, a tree-based approach might better reflect the process(es) that generated traits. However, these assumptions of constant rates of evolution for all species need to be thoroughly verified.

The largest differences between the two clustering approaches (Ward and UPGMA) appeared when species had large amount of independent (compared to shared) evolutionary histories (ROOT model). More generally, the UPGMA approach led to higher correlations between phylogenetic and functional originalities whichever the index of originality was used. When a tree-based approach is adopted, a UPGMA functional tree could thus be the most appropriate. However, further research is needed specifically comparing a larger range of clustering approaches (Mouchet et al., 2008).

4.3. Enlarging the perspectives: measure more than species originality

Other criteria that may influence the choice of an index is the possibility to use other attributes (than phylogeny or traits) to define species originalities. For example, Faith (2008, 2015a), Rosauer et al. (2009), Cadotte et al. (2010) modified the *FP* index by weighting species by probabilities of extinction, range size and abundance, respectively. Whether these modified measures of phylogenetic originalities can be applied to functional trees is questionable. Indeed they were developed for rooted phylogenetic trees with branch lengths representing shared evolutionary histories. Further research is needed on how to include

additional information such as threat, endemism and/or abundance into indices of functional originality (See Ricotta et al. 2016 first developments in this direction).

More generally, as dealing with functional originality led us to treat data that are not necessarily connected to a hierarchy (in contrast to phylogenetic trees), the conclusions drawn in this paper have the potential to be extended to assess different types of originality. Indeed the concept of originality can be derived at any scale: using dissimilarities between individuals (measuring data per individual), between populations, between species, or between sites (or assemblages, plots, regions, etc.). The concept of originality could then be adapted for establishing conservation priorities across multiple scales. For example, dissimilarities among communities, even when phylogenetic information is used (e.g. Ives and Helmus, 2010; Chiu et al., 2014; Pavoine, 2016), are rarely derived from trees. At the plot and regional scales, our methodology can also be extended to measure the environmental originality of plots within regions, and of regions. This can be done simply by replacing biological with environmental data when calculating dissimilarities between plots and between regions. Our study thus opens the way to new directions of research where the biological originalities of areas will be compared to their environmental originalities.

5. Conclusions

Originality indices provide critically important but different interpretations to measuring biodiversity. The use of these sorts of indices requires more integrative approaches in both space and/or time. Here we have highlighted methodological similarities between functional and phylogenetic originality. We analyzed alternatives for measuring species originalities that do not depend on the clustering approach. We demonstrated the importance of methodological choices in determining species' originalities. These choices are likely to

impact both the probability of observing significantly original species and the chance to understand the local ecological processes driving species originality. Our new index (Rb) strongly correlates with the average distance to all other species in a set, and its framework is closely related to indices based on quadratic entropy (QE). In contrast to Rb , the QE -based index (Qb) is influenced both by the effects of average distance and strict uniqueness. Both these indices (Rb and Qb) are derived from the concepts of diversity (in terms of effective number of species) and entropy. They highlight direct links between originality and the contribution each species has to both diversity and entropy. Applications to field observational and experimental studies, as well extending theoretical models, are still necessary to evaluate the future impact of species originality as a dimension of community structure in conservation ecology.

Acknowledgements

We thank Owen Petchey for useful comments on a previous version of the paper. U.J. is supported by the German Centre for Integrative Biodiversity Research (iDiv), funded by the German Research Foundation (FZT 118) and funded by the Ministry for Science and Culture of Lower Saxony (BEFmate).

Appendixes A-E. Supplementary data

References

620 Atkinson, J., 1989. Introduced animals and extinctions, in: Western, D., Pearl, M. (Eds.)
 621 Conservation for the twenty-first century. Oxford University Press, Oxford UK, pp.
 622 54-75.

623 Barker, G.M., 2002. Phylogenetic diversity: a quantitative framework for measurement of
 624 priority and achievement in biodiversity conservation. *Biological Journal of the*
 625 *Linnean Society* 76, 165–194.

626 Blomberg, S.P., Garland, T., Ives, A.R., 2003. Testing for phylogenetic signal in comparative
 627 data: behavioral traits are more labile. *Evolution* 57, 717–745.

628 Brooks, T.M., Cuttelod, A., Faith, D.P., Garcia-Moreno, J., Langhammer, P., Pérez-Espona,
 629 S., 2015. Why and how might genetic and phylogenetic diversity be reflected in the
 630 identification of key biodiversity areas? *Philosophical Transactions: Biological*
 631 *Sciences* 370, 20140019.

632 Buisson, L., Grenouillet, G., Villéger, S., Canal, J., Laffaille, P., 2013. Toward a loss of
 633 functional diversity in stream fish assemblages under climate change. *Global Change*
 634 *Biology* 19, 387–400.

635 Cadotte, M.W., Davies, J.T., Regetz, J., Kembel, S.W., Cleland, E., Oakley, T.H., 2010.
 636 Phylogenetic diversity metrics for ecological communities: integrating species
 637 richness, abundance and evolutionary history. *Ecology Letters* 13, 96–105.

638 Chao, A., Chiu, C.H., Jost, L., 2010. Phylogenetic diversity measures based on Hill numbers.
 639 *Philosophical Transactions of the Royal Society of London B: Biological Sciences*,
 640 365, 3599–3609.

641 Chiu, C.H., Jost, L., Chao, A., 2014. Phylogenetic beta diversity, similarity, and
 642 differentiation measures based on Hill numbers. *Ecological monographs* 84, 21–44.

643 Collen, B., Turvey, S.T., Waterman, C., Meredith, H.M.R., Kuhn, T.S., Baillie, J.E.M., Isaac,
644 N.J.B., 2011. Investing in evolutionary history: implementing a phylogenetic approach
645 for mammal conservation. *Philosophical Transactions of the Royal Society, B:*
646 *Biological Sciences* 366, 2611–2622.

647 Dray, S., Dufour, A.B., 2007. The ade4 package: implementing the duality diagram for
648 ecologists. *Journal of Statistical Software* 22, 1–20.

649 Eiswerth, M.E., Haney, J.C., 1992. Allocating conservation expenditures: accounting for
650 inter-species genetic distinctiveness. *Ecological Economics* 5, 235–249.

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

653 Faith, D.P., 2002. Quantifying biodiversity: a phylogenetic perspective. *Conservation*
654 *Biology*, 16, 248–252.

655 Faith, D.P., 2008. Threatened species and the potential loss of phylogenetic diversity:
656 conservation scenarios based on estimated extinction probabilities and phylogenetic
657 risk analysis. *Conservation Biology* 22, 1461–1470.

658 Faith, D.P., 2015a. Phylogenetic diversity, functional trait diversity and extinction: avoiding
659 tipping points and worst-case losses. *Philosophical Transactions of the Royal Society*
660 *of London B: Biological Sciences* 370, 20140011.

661 Faith, D.P., 2015b. The unimodal relationship between species' functional traits and habitat
662 gradients provides a family of indices supporting the conservation of functional trait
663 diversity. *Plant Ecology* 216, 725–740.

664 Gerhold, P., Cahill, J. F., Winter, M., Bartish, I. V., Prinzing, A., 2015. Phylogenetic patterns
665 are not proxies of community assembly mechanisms (they are far better). *Functional*
666 *Ecology* 29, 600–614.

667 Godet, L., Gaüzere, P., Jiguet, F., Devictor, V., 2015. Dissociating several forms of
668 commonness in birds sheds new light on biotic homogenization. *Global Ecology and*
669 *Biogeography* 24, 416–426.

670 Gower, J.C., 1971. A general coefficient of similarity and some of its properties. *Biometrics*,
671 857–871.

672 Hardy, O. J., Pavoine, S., 2012. Assessing phylogenetic signal with measurement error: a
673 comparison of Mantel tests, Blomberg et al.'s K and phylogenetic distograms.
674 *Evolution* 66, 2614–2621.

675 Harmon L.J., Weir, J.T., Brock, C.D., Glor, R.E., Challenger, W., 2008. geiger: investigating
676 evolutionary radiations. *Bioinformatics* 24, 129–131.

677 Heard, S.B., Mooers, A.O., 2000. Phylogenetically patterned speciation rates and extinction
678 risks change the loss of evolutionary history during extinctions. *Proceedings of the*
679 *Royal Society of London Series B - Biological Sciences* 267, 613–620.

680 Hill, M.O., 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology*
681 54, 427–432.

682 Hooper, D.U., Chapin III, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H.,
683 Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J.,
684 Vandermeer, J., Wardle, D.A., 2005. Effects of biodiversity on ecosystem functioning:
685 a consensus of current knowledge. *Ecological monographs* 75, 3–35.

686 Huang, D., 2012. Threatened reef corals of the world. PLoS ONE 7, e34459.

687 Isaac, N.J., Redding, D.W., Meredith, H.M., Safi, K., 2012. Phylogenetically-informed
688 priorities for amphibian conservation. PLoS ONE 7, e43912.

689 Isaac, N.J., Turvey, S.T., Collen, B., Waterman, C., Baillie, J.E., 2007. Mammals on the
690 EDGE: conservation priorities based on threat and phylogeny. PLoS ONE 2, e296.

691 IUCN, 2015. The IUCN Red List of Threatened Species. Version 2015-3.
692 <<http://www.iucnredlist.org>>. Downloaded on May 2015.

693 Ives, A.R., Helmus, M.R., 2010. Phylogenetic metrics of community similarity. The
694 American Naturalist 176, E128–E142.

695 Jensen, E.L., Mooers, A.O., Caccione, A., Russello, M.A., 2016. I-HEDGE: determining the
696 optimum complementarity sets of taxa for conservation using evolutionary isolation.
697 PeerJ 4, e2350.

698 Jetz, W., Thomas, G.H., Joy, J.B., Redding, D.W., Hartmann, K., Mooers, A.O., 2014. Global
699 distribution and conservation of evolutionary distinctness in birds. Current Biology 24,
700 919–930.

701 Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O'Dell, J., Orme, C.D.L., Safi, K., Sechrest,
702 W., Boakes, E.H., Carbone, C., Connolly, C., Cutts, M.J., Foster, J.K., Grenyer, R.,
703 Habib, M., Plaster, C.A., Price, S.A., Rigby, E.A., Rist, J., Teacher, A., Bininda-
704 Emonds, O.R.P., Gittleman, J.L., Mace, G.M., Purvis, A., 2009. PanTHERIA: a
705 species-level database of life history, ecology, and geography of extant and recently
706 extinct mammals. Ecology 90, 2648–2648.

707 Jost, L., 2006. Entropy and diversity. Oikos 113, 363–375.

708 Leinster, T., Cobbold, C.A., 2012. Measuring diversity: the importance of species similarity.
 709 Ecology 93, 477–489.

710 Losos, J.B., 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship
 711 between phylogenetic relatedness and ecological similarity among species. Ecology
 712 Letters 11, 995–1007.

713 MacDonald, D.W. (ed.), 2009. Princeton encyclopedia of Mammals, Princeton University
 714 Press, Princeton, NJ.

715 Maddison, W.P., Slatkin, M., 1991. Null models for the number of evolutionary steps in a
 716 character on a phylogenetic tree. Evolution 45, 1184–1197.

717 Magnuson-Ford, K., Ingram, T., Redding, D.W., Mooers, A.Ø., 2009. Rockfish (Sebastes)
 718 that are evolutionarily isolated are also large, morphologically distinctive and
 719 vulnerable to overfishing. Biological conservation 142, 1787–1796.

720 May, R.M., 1990. Taxonomy as destiny. Nature 347, 129–130.

721 Mouchet, M., Guilhaumon, F., Villéger, S., Mason, N.W., Tomasini, J.A., Mouillot, D., 2008.
 722 Towards a consensus for calculating dendrogram-based functional diversity indices.
 723 Oikos 117, 794–800.

724 Mouillot, D., Bellwood, D.R., Baraloto, C., Chave, J., Galzin, R., Harmelin-Vivien, M.,
 725 Kulbicki, M., Lavergne, S., Lavorel, S., Mouquet, N., Paine, C.E.T., Renaud, J,
 726 Thuiller, W., 2013. Rare species support vulnerable functions in high-diversity
 727 ecosystems. PloS Biology 11, e1001569.

728 Mouillot, D., Culioli, J.M., Pelletier, D., Tomasini, J., 2008. Do we protect biological
729 originality in protected area? A new index and an application to the Bonifacio Strait
730 Natural Reserve. *Biological Conservation* 141, 1569–1580.

731 Myers, P., Espinosa, R., Parr, C.S., Jones, T., Hammond, G.S., Dewey, T.A., 2015. The
732 Animal Diversity Web (online). Accessed at <http://animaldiversity.org>, June 2015.

733 Nyakatura, K., Bininda-Emonds, O., 2012. Updating the evolutionary history of Carnivora
734 (Mammalia): a new species-level supertree complete with divergence time estimates.
735 *BMC Biology* 10, 12.

736 Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in
737 R language. *Bioinformatics* 20, 289–290.

738 Patil, G.P., Taillie, C., 1982. Diversity as a concept and its measurement. *Journal of the*
739 *American Statistical Association* 77, 548–561.

740 Pavoine, S., 2016. A guide through a family of phylogenetic dissimilarity measures among
741 sites. *Oikos* 125, 1719–1732.

742 Pavoine, S., Bonsall, M.B., 2011. Measuring biodiversity to explain community assembly: a
743 unified approach. *Biological Reviews* 86, 792–812.

744 Pavoine, S., Love, M.S., Bonsall, M.B., 2009. Hierarchical partitioning of evolutionary and
745 ecological patterns in the organization of phylogenetically-structured species
746 assemblages: application to rockfish (genus: *Sebastes*) in the Southern California
747 Bight. *Ecology letters* 12, 898–908.

748 Pavoine, S., Ollier, S., Dufour, A.B., 2005. Is the originality of a species measurable? *Ecology*
749 *Letters* 8, 579–586.

750 Pavoine, S., Ricotta, C., 2013. Testing for phylogenetic signal in biological traits: the ubiquity
751 of cross-product statistics. *Evolution* 67, 828–840.

752 Petchey, O.L., Eklöf, A., Borrvall, C., Ebenman, B., 2008. Trophically unique species are
753 vulnerable to cascading extinction. *The American Naturalist* 171, 568–579.

754 Petchey, O.L., Evans, K.L., Fishburn, I.S., Gaston, K.J., 2007. Low functional diversity and
755 no redundancy in British avian assemblages. *Journal of Animal Ecology* 76, 977–985.

756 Petchey, O.L., Hector, A., Gaston, K.J., 2004. How do different measures of functional
757 diversity perform? *Ecology* 85, 847–857.

758 Petchey, O.L., O’Gorman, E.J., Flynn, D.F., 2009. A functional guide to functional diversity
759 measures, in: Naeem, S., Bunker, D.E., Hector, A., Loreau, M., Perrings, C. (Eds.),
760 Biodiversity, ecosystem functioning, and human wellbeing: an ecological and
761 economic perspective. Oxford University Press, Oxford, pp. 49–60.

762 Purvis, A., Agapow, P.M., Gittleman, J.L., Mace, G.M., 2000. Non-random extinction and the
763 loss of evolutionary history. *Science* 288, 328–330.

764 R Core Team, 2016. R: A language and environment for statistical computing, R Foundation
765 for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

766 Rao, C.R., 1982. Diversity and dissimilarity coefficients: a unified approach. *Theoretical*
767 *Population Biology* 21, 24–43.

768 Redding, D.W., 2003. Incorporating genetic distinctness and reserve occupancy into a
769 conservation prioritisation approach. Master thesis: University of East Anglia, Norwich.

770 Redding, D.W., Mazel, F., Mooers, A.Ø., 2014. Measuring evolutionary isolation for
771 conservation. *PloS ONE* 9, e113490.

772 Redding, D.W., Mooers, A.Ø., 2006. Incorporating evolutionary measures into conservation
773 prioritization. *Conservation Biology* 20, 1670–1678.

774 Ricotta, C., 2004. A parametric diversity measure combining the relative abundances and
775 taxonomic distinctiveness of species. *Diversity and Distributions* 10, 143–146.

776 Ricotta, C., de Bello, F., Moretti, M., Caccianiga, M., Cerabolini, B.E.L., Pavoine, S., 2016.
777 Measuring the functional redundancy of biological communities: a quantitative guide.
778 *Methods in Ecology and Evolution* 7, 1386–1395.

779 Ricotta, C., Szeidl, L., 2009. Diversity partitioning of Rao's quadratic entropy. *Theoretical*
780 *Population Biology* 76, 299–302.

781 Rosatti, N.B., Silva, D.M., Batalha, M.A., 2015. Loss of phylogenetic and functional
782 originalities of woody cerrado species in simulated extinction scenarios. *Austral*
783 *Ecology* 40, 267–274.

784 Rosauer, D.F., Laffan, S.W., Crisp, M.D., Donnellan, S.C., Cook, L.G., 2009. Phylogenetic
785 endemism: a new approach for identifying geographical concentrations of
786 evolutionary history. *Molecular Ecology* 18, 4061–4072.

787 Schmera, D., Erős, T., Podani, J., 2009a. A measure for assessing functional diversity in
788 ecological communities. *Aquatic Ecology* 43, 157–167.

789 Schmera, D., Podani, J., Erős, T., 2009b. Measuring the contribution of community members
790 to functional diversity. *Oikos* 118, 961–971.

791 Stadler, T., 2015. TreeSim: Simulating Phylogenetic Trees. R package version 2.2.
792 <https://CRAN.R-project.org/package=TreeSim>.

793 Strauss, S.Y., Webb, C.O., Salamin, N., 2006. Exotic taxa less related to native species are
794 more invasive. *Proceedings of the National Academy of Sciences* 103, 5841–5845.

795 Strayer, D.L., Eviner, V.T., Jeschke, J.M., Pace, M.L., 2006. Understanding the long-term
796 effects of species invasions. *Trends in Ecology & Evolution* 21, 645–651.

797 Temple, H.J., Terry, A., 2007. The status and distribution of European mammals, Office for
798 Official Publications of the European Communities, Luxembourg. Available:
799 www.iucn-redlist.org.

800 Thompson, K., Petchey, O.L., Askew, A.P., Dunnett, N.P., Beckerman, A.P., Willis, A.J.,
801 2010. Little evidence for limiting similarity in a long-term study of a roadside plant
802 community. *Journal of Ecology* 98, 480–487.

803 Vane-Wright, R.I., Humphries, C.J., Williams, P.H., 1991. What to protect? Systematics and
804 the agony of choice. *Biological Conservation* 55, 235–254.

805 Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., Garnier, E., 2007.
806 Let the concept of trait be functional! *Oikos* 116, 882–892.

807 Webb, C.O., Ackerly, D.D., McPeck, M.A., Donoghue, M.J., 2002. Phylogenies and
808 community ecology. *Annual Review of Ecology and Systematics* 33, 475–505.

809

810

811

812

813

814

815

816

817

818

819

Fig. 1. Tree-based indices used in this paper: a) the theoretical tree used in all panels with indication of branches and branch lengths, tips (species from A to F), and root; b) Pendant Edge (PE) index; c) Fair Proportion (FP) index; d) Equal-Split (ES) index. In each panel, “Total” indicates the originality values attributed by each index to each species. Letters and numbers above and below branches on the tree indicate how much each branch contributes to the originality of each species. For example, in b) only terminal branches contribute to the PE measure of originality.

Fig. 2. Link between data type (functional or phylogenetic data) and the originality indices applied to the data. Patristic dissimilarities designate the distance between two leaves (species) measured along the tree: here, half the sum of branch lengths on the shortest path that connects the two species on the tree. We computed *Rb*, *AV* and *NN* directly on the defined functional dissimilarities among species. We also computed them indirectly from the established functional trees using patristic distances. In this latter case, the indices *Rb* and *AV* were named *tb-Rb* and *tb-AV*, respectively, where '*tb*' stands for 'tree-based' to explicitly specify that their calculation depended on a functional tree. *NN* computed indirectly on a tree, noted *tb-NN*, corresponded to *PE* (the pendant edge).

Fig. 3. Box plots of correlations between functional and phylogenetic originalities for each originality index, simulation type [clustering approach (UPGMA, Ward, or raw distances = no clustering), number of traits (1 trait, 25 traits, or 50 traits), and phylogenetic tree model (BIRTH, ROOT, TIPS, or NU)].

Fig. 4. Relationships between Rb and the other originality indices used in the simulations. a) Typical links obtained between AV and Rb (this example was simulated with BIRTH model and 25 traits). b) Box plots of the Spearman correlations between Rb and the other originality indices in the simulation case study. c) Principal component analysis of the Spearman correlations between Rb and tree-based originality indices (see also Appendix E for a summary table of the correlations). In c) the coordinates of indices and simulation types are given on axes 1 (57% of variation; horizontal axis) and 2 (26%; vertical axis); supplementary variables (underlined in panel c) were added on the map at the centre of their associated simulation types: the clustering approaches (U=UPGMA, W=Ward), the number of traits (1, 25, or 50), and the phylogenetic tree model (BIRTH, ROOT, TIPS, or NU)].

Fig. 5. Carnivore case study: box plots of the relative species originalities grouped into introduced species (noted 'I' on the left of each panel), native species with marginal occurrence in Europe (noted 'M' at the middle of each panel) and other native species (noted 'N' on the right of each panel). Originality was calculated using a) phylogenetic data, b) raw functional dissimilarities, c) functional data and Ward clustering, d) functional data and UPGMA clustering. Codes for the originality indices used are indicated at the top of the figure. The notations "(PE)NN", "(tb-)AV" and "(tb-)Rb" mean that we applied indices PE, tb-AV and tb-Rb to trees (phylogenies, Ward and UPGMA functional trees), and indices NN, AV and Rb to raw dissimilarities (see Fig. 2).