Ecological Indicators 82 (2017) 196–205

Contents lists available at ScienceDirect

Ecological Indicators

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

Original Articles

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

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, Germany
[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

ARTICLE INFO

Keywords:
Biodiversity
Carnivores
Conservation priorities
Entropy
Simulations
Species distinctiveness

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

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

---

[*] Corresponding author at: Muséum national d'Histoire naturelle, UMR 7204 CNRS-UPMC, CESCO, 61 rue Buffon, 75005 Paris, France.
E-mail address: sandrine.pavoine@mnhn.fr (S. Pavoine).

http://dx.doi.org/10.1016/j.ecolind.2017.06.056
Received 22 February 2017; Received in revised form 27 June 2017; Accepted 27 June 2017

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

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		Refs.
		a tree structure	a dissimilarity matrix	
AV	Average distance to other species		X	Eiswerth and Haney (1992)
ES	Equal-Split (branches in a tree are split equally among descending clades)	X		Redding and Mooers (2006)
FP	Fair Proportion (branches in a tree are split fairly among descending species)	X		Redding (2003)
M	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)
NN	Distance to the nearest neighbour		X	This paper
PE	Pendant Edge (terminal branch of a tree)	X		Redding et al. (2014)
Qb	Species' proportions that maximize the quadratic entropy diversity index	X	X	Pavoine et al. (2005)
Rb	Species' proportions that maximize the *R* diversity index		X	This paper
tb-AV	*AV* index applied on tree-based distances among species	X	X	This paper
tb-Rb	*Rb* index applied on tree-based distances among species	X	X	This paper

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.

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*.

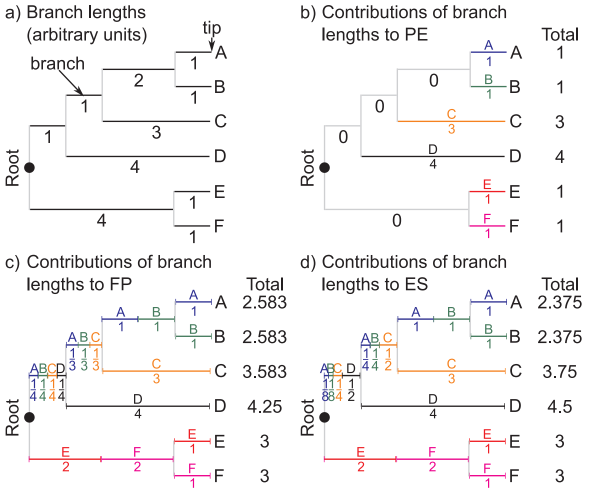

Fig. 1. Tree-based indices used in this paper: (a) the theoretical tree used in all panels with indication of branches and branch lengths, tips (species from A to F), and root; (b) Pendant Edge (PE) index; (c) Fair Proportion (FP) index; (d) Equal-Split (ES) index. In each panel, "Total" indicates the 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.

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*): $FV = [(N\text{-}1)^*AV]/N$. 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 $\mathbf{D} = (d_{ij})_{1 \le i \le N, 1 \le j \le N}$ 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: $d_{ij} \le \max(d_{ik}, d_{jk})$ for all *i,j,k*.

Let $\mathbf{p} = (p_1, \ldots, p_i, \ldots, p_N)$ 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):

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

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* reflects 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).

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):

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

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

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

where q modifies the relative importance given to rare versus common species: the sensitivity of ^{q}D to rare species decreases with q (Patil and Taillie, 1982). The limiting case ($q \to 1$) serves as a reference where species are weighted directly by their relative abundances (p_is) 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 ^{2}D 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 in the Supplementary material):

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

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 in the Supplementary material).

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 in the Supplementary material). 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 in the Supplementary material for a simple example). Using the vector of p_is 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).

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 non-ultrametric 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 $LogN(0,1)$). 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

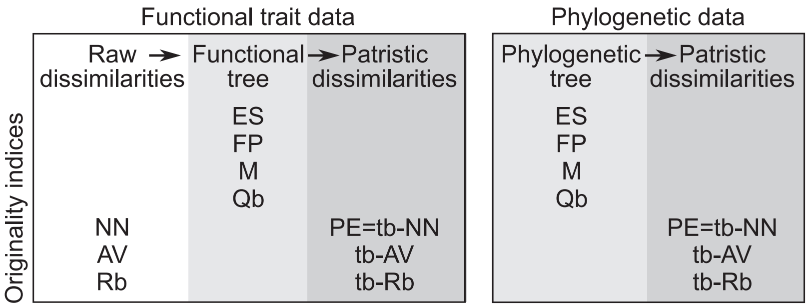

Fig. 2. Link between data type (functional or phylogenetic data) and the originality indices applied to the data. Patristic dissimilarities designate the distance between two leaves (species) measured along the tree: here, half the sum of branch lengths on the shortest path that connects the two species 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).

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 CE. (*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-Edmonds (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 (in the Supplementary material) 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 (in the Supplementary material) 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 in the Supplementary material). 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 in the Supplementary material.

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) *Kw* 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 the Supplementary material). 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.

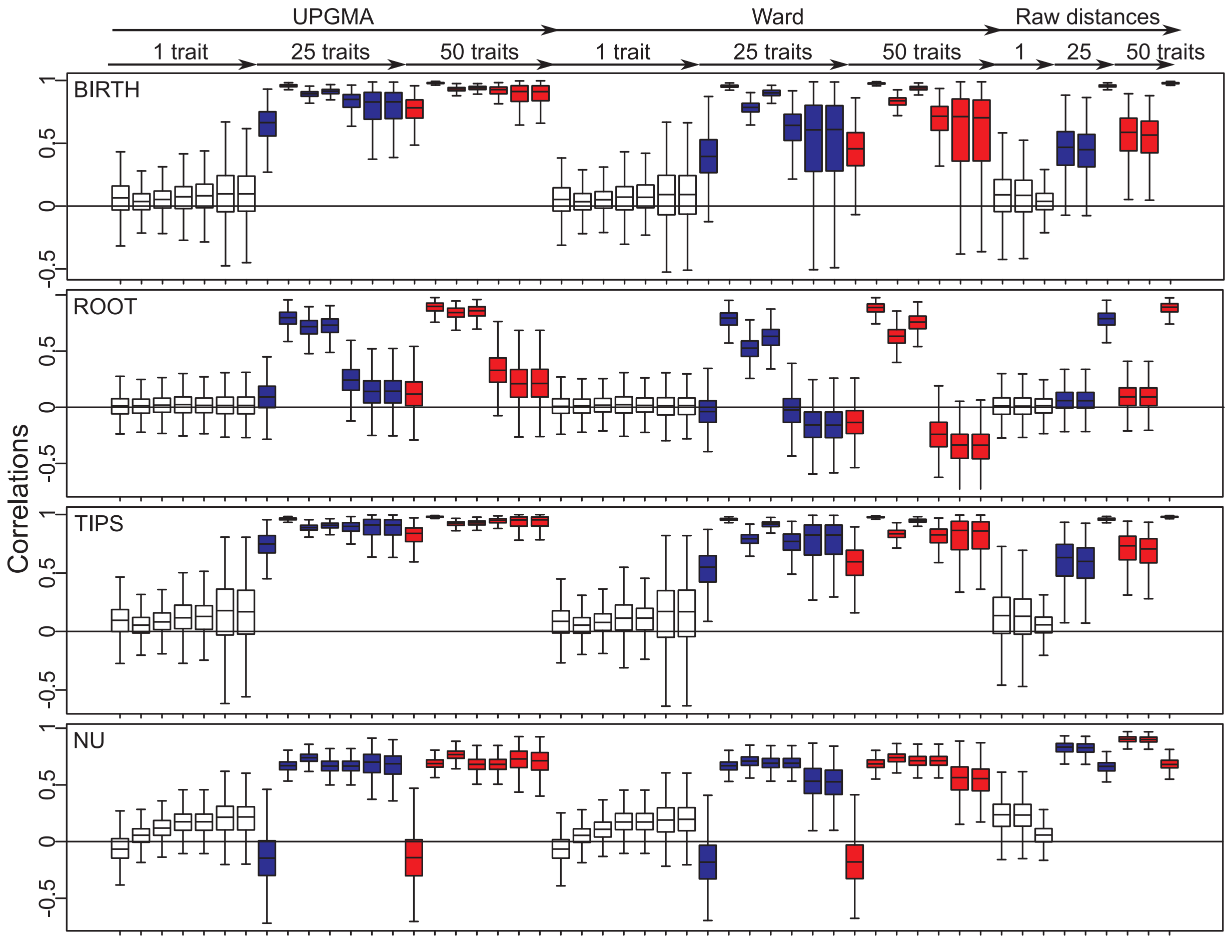

Fig. 3. Box plots of correlations between functional and phylogenetic originalities for each originality index, simulation type [clustering approach (UPGMA, Ward, or raw distances = no clustering), number of traits (1 trait, 25 traits, or 50 traits), and phylogenetic tree model (BIRTH, ROOT, TIPS, or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

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	

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 in the Supplementary material 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).

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 in the Supplementary material), 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).

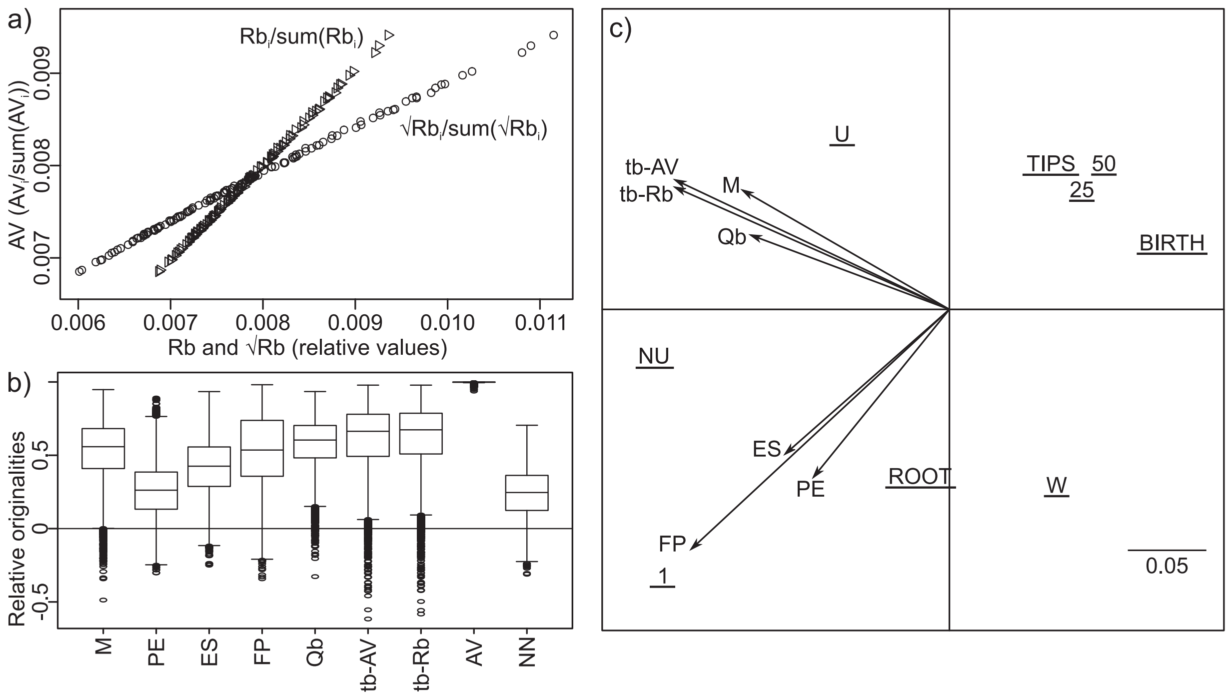

Fig. 4. Relationships between *Rb* and the other originality indices used in the simulations. (a) Typical links obtained between *AV* and *Rb* (this example was simulated with BIRTH model and 25 traits). (b) Box plots of the Spearman correlations between *Rb* and the other originality indices in the simulation case study. (c) Principal component analysis of the Spearman correlations between *Rb* and tree-based originality indices (see also Appendix E in the Supplementary material 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)].

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

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).

4. Discussion

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

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

Using a phylogeny to predict diversity has been particularly important for establishing broad-scale prioritizing schemes in conservation biology (e.g. Barker, 2002; Isaac et al., 2007). In more fundamental ecological studies, for example, Strauss et al. (2006) showed that invaders are more likely to be successful if they are phylogenetically distinct from the natives 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,

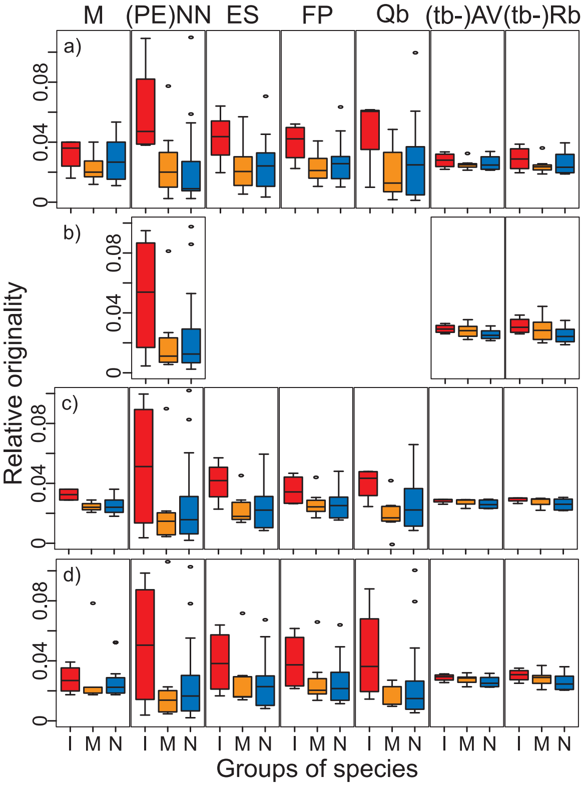

Fig. 5. Carnivore case study: box plots of the relative species originalities grouped into introduced species (noted 'I' on the left of each panel), native species with marginal occurrence in Europe (noted 'M' at the middle of each panel) and other native species (noted 'N' on the right of each panel). 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).

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. Gerhold 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 in the Supplementary material). 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

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolind.2017.06.056.

References

Atkinson, J., 1989. Introduced animals and extinctions. In: Western, D., Pearl, M. (Eds.), Conservation for the Twenty-First Century. Oxford University Press, Oxford UK, pp. 54–75.

Barker, G.M., 2002. Phylogenetic diversity: a quantitative framework for measurement of priority and achievement in biodiversity conservation. Biol. J. Linn. Soc. 76, 165–194.

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

Brooks, T.M., Cuttelod, A., Faith, D.P., Garcia-Moreno, J., Langhammer, P., Pérez-Espona, S., 2015. Why and how might genetic and phylogenetic diversity be reflected in the identification of key biodiversity areas? Philos. Trans.: Biol. Sci. 370, 20140019.

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

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

Chao, A., Chiu, C.H., Jost, L., 2010. Phylogenetic diversity measures based on Hill numbers. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 365, 3599–3609.

Chiu, C.H., Jost, L., Chao, A., 2014. Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. Ecol. Monogr. 84, 21–44.

Collen, B., Turvey, S.T., Waterman, C., Meredith, H.M.R., Kuhn, T.S., Baillie, J.E.M., Isaac, N.J.B., 2011. Investing in evolutionary history: implementing a phylogenetic approach for mammal conservation. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 366, 2611–2622.

Dray, S., Dufour, A.B., 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. 22, 1–20.

Eiswerth, M.E., Haney, J.C., 1992. Allocating conservation expenditures: accounting for inter-species genetic distinctiveness. Ecol. Econ. 5, 235–249.

Faith, D.P., 1992. Conservation evaluation and phylogenetic diversity. Biol. Conserv. 61, 1–10.

Faith, D.P., 2002. Quantifying biodiversity: a phylogenetic perspective. Conserv. Biol. 16, 248–252.

Faith, D.P., 2008. Threatened species and the potential loss of phylogenetic diversity: conservation scenarios based on estimated extinction probabilities and phylogenetic risk analysis. Conserv. Biol. 22, 1461–1470.

Faith, D.P., 2015a. Phylogenetic diversity, functional trait diversity and extinction: avoiding tipping points and worst-case losses. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 370, 20140011.

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

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

Godet, L., Gaüzere, P., Jiguet, F., Devictor, V., 2015. Dissociating several forms of commonness in birds sheds new light on biotic homogenization. Global Ecol. Biogeogr. 24, 416–426.

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

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

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

Heard, S.B., Mooers, A.O., 2000. Phylogenetically patterned speciation rates and extinction risks change the loss of evolutionary history during extinctions. Proc. R. Soc. Lond. Ser. B: Biol. Sci. 267, 613–620.

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

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

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

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

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

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

Ives, A.R., Helmus, M.R., 2010. Phylogenetic metrics of community similarity. Am. Nat. 176, E128–E142.

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

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

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

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

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

Losos, J.B., 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecol. Lett. 11, 995–1007.

MacDonald, D.W. (Ed.), 2009. Princeton Encyclopedia of Mammals. Princeton University Press, Princeton, NJ.

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

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

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

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

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

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

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

Nyakatura, K., Bininda-Edmonds, O., 2012. Updating the evolutionary history of carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. BMC Biol. 10, 12.

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

Patil, G.P., Taillie, C., 1982. Diversity as a concept and its measurement. J. Am. Stat. Assoc. 77, 548–561.

Pavoine, S., Bonsall, M.B., 2011. Measuring biodiversity to explain community assembly: a unified approach. Biol. Rev. 86, 792–812.

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

Pavoine, S., Ollier, S., Dufour, A.B., 2005. Is the originality of a species measurable? Ecol. Lett. 8, 579–586.

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

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

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

Petchey, O.L., Evans, K.L., Fishburn, I.S., Gaston, K.J., 2007. Low functional diversity and no redundancy in British avian assemblages. J. Anim. Ecol. 76, 977–985.

Petchey, O.L., Eklöf, A., Borrvall, C., Ebenman, B., 2008. Trophically unique species are vulnerable to cascading extinction. Am. Nat. 171, 568–579.

Petchey, O.L., O'Gorman, E.J., Flynn, D.F., 2009. A functional guide to functional diversity measures. In: Naeem, S., Bunker, D.E., Hector, A., Loreau, M., Perrings, C. (Eds.), Biodiversity, Ecosystem Functioning, and Human Wellbeing: an Ecological and Economic Perspective. Oxford University Press, Oxford, pp. 49–60.

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

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria URL. http://www.R-project.org/.

Rao, C.R., 1982. Diversity and dissimilarity coefficients: a unified approach. Theor. Popul. Biol. 21, 24–43.

Redding, D.W., Mooers, A.Ø., 2006. Incorporating evolutionary measures into conservation prioritization. Conserv. Biol. 20, 1670–1678.

Redding, D.W., Mazel, F., Mooers, A.Ø., 2014. Measuring evolutionary isolation for conservation. PLoS One 9, e113490.

Redding, D.W., 2003. Incorporating Genetic Distinctness and Reserve Occupancy into a Conservation Priorisation Approach. Master Thesis. University of East Anglia, Norwich.

Ricotta, C., Szeidl, L., 2009. Diversity partitioning of Rao's quadratic entropy. Theor. Popul. Biol. 76, 299–302.

Ricotta, C., de Bello, F., Moretti, M., Caccianiga, M., Cerabolini, B.E.L., Pavoine, S., 2016. Measuring the functional redundancy of biological communities: a quantitative guide. Methods Ecol. Evol. 7, 1386–1395.

Ricotta, C., 2004. A parametric diversity measure combining the relative abundances and taxonomic distinctiveness of species. Divers. Distrib. 10, 143–146.

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

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

Schmera, D., Erős, T., Podani, J., 2009a. A measure for assessing functional diversity in ecological communities. Aquat. Ecol. 43, 157–167.

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

Stadler, T., 2015. TreeSim: Simulating Phylogenetic Trees. R Package Version 2.2. https://CRAN.R-project.org/package=TreeSim.

Strauss, S.Y., Webb, C.O., Salamin, N., 2006. Exotic taxa less related to native species are more invasive. Proc. Natl. Acad. Sci. 103, 5841–5845.

Strayer, D.L., Eviner, V.T., Jeschke, J.M., Pace, M.L., 2006. Understanding the long-term effects of species invasions. Trends Ecol. Evol. 21, 645–651.

Temple, H.J., Terry, A., 2007. The Status and Distribution of European Mammals. Office for Official Publications of the European Communities, Luxembourg Available: www.iucn redlist.org.

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

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

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

Webb, C.O., Ackerly, D.D., McPeek, M.A., Donoghue, M.J., 2002. Phylogenies and community ecology. Ann. Rev. Ecol. Syst. 33, 475–505.

Appendix A. Proof and definitions

In our study, we compared indices that either used a matrix of dissimilarity among species directly or functional data transformed into a functional tree. We did not consider methods that use ordination approaches to position species into Euclidean space (Magnusson et al., 2009; Faith, 2015). Further studies on the need to use such approaches and how they compare to tree-based or distance-based approaches are still very much required. Connections exist among these alternative approaches: *Qb* can also be derived from an ordination method (Pavoine et al., 2005).

Definitions

For any real ($n \times n$) symmetric matrix $\mathbf{A}$, there exists a unitary matrix $\mathbf{U}$ with n rows (i.e. $\mathbf{U}^t\mathbf{U}=\mathbf{I}_n$, the $n \times n$ identity matrix) such that $\mathbf{U}^t\mathbf{A}\mathbf{U}=\mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is a diagonal matrix of the eigenvalues of $\mathbf{A}$. Eigenvalues are real and eigenvectors, columns of $\mathbf{U}$, can be taken to be real (e.g. Seber 2008).

Proof that the first eigenvector of a dissimilarity matrix can be taken to be positive

Let $\mathbf{D}=(d_{ij})_{1\leq i\leq N, 1\leq j\leq N}$ be a matrix of pairwise dissimilarities with $d_{ij}=d_{ji}$ for any i,j, $d_{ij}>0$ for any $i\neq j$, and $d_{ii}=0$ for any i.

It is shown below that the first eigenvector of $\mathbf{D}$ has values of same sign, which can be taken to be positive.

According to Noutsos (2006, definition 2.2), a matrix $\mathbf{A} \in \mathbb{R}^{N,N}$ possesses the strong Perron-Frobenius property (in reference to the Perron-Frobenius theorem) if its dominant eigenvalue λ_1 is positive, the only one in the circle $|\lambda_1|$ ($\lambda_1 > |\lambda_i|$, $i = 2, 3, \ldots, N$) and the corresponding eigenvector $x^{(1)}$ is positive (which means that it has only positive values).

According to Noutsos (2006, definition 2.3), a matrix $\mathbf{A} \in \mathbb{R}^{N,N}$ is said to be *eventually positive* if there exists a positive integer k_0 such that $A^k > 0$ ($=A^k$ only has positive values) for all $k \geq k_0$. The demonstration that $\mathbf{D}$ is eventually positive stems from the fact that $d_{ij}>0$ for any $i\neq j$, and $d_{ii}=0$ for any i. With the latter conditions for the values of $\mathbf{D}$, $\mathbf{D}^k > 0$ for all $k \geq 2$.

According to Noutsos (2006, theorem 2.1), given that $\mathbf{D}$ is eventually positive and symmetric ($d_{ij}=d_{ji}$ for any i,j), then $\mathbf{D}$ possesses the strong Perron-Frobenius property.

□

Proof that the unique maximizing vector of R is the squared, first eigenvector of the matrix of dissimilarities

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

Consider $u_i=\sqrt{p_i}$ and $\boldsymbol{u}=(u_1,...,u_N)$.

$$R(\mathbf{u},\mathbf{D})=\sum_{i=1}^{N}\sum_{j=1}^{N}u_i u_j d_{ij}$$

It is well known that the maximum of **R**, given that $\sum_{i=1}^{N}u_i^2=1$ is the first eigenvalue of **D** that is reached when **u** is the first eigenvector of **D**. The fact that this eigenvector has positive values has been proved above so that the maximum of $\max_{\mathbf{p}}\left\{R(\mathbf{p},\mathbf{D})\right\}$ is the first eigenvalue of **D** that is reached for the squared, first eigenvector of **D**.

□

Details on Rao's Q

As above, we begin by considering a matrix $\mathbf{D}=(d_{ij})_{1\leq i\leq N,1\leq j\leq N}$ of dissimilarities among N species with N the number of species such that the dissimilarities are symmetric: $d_{ij}=d_{ji}$ for any i,j, $d_{ij}>0$ for any $i\neq j$, and $d_{ii}=0$ for each i. We also consider $\mathbf{p}=(p_1, ..., p_i, ..., p_N)$ a vector of species' relative abundance ($\sum_{i=1}^{N}p_i=1$). The average dissimilarity among individuals from an assemblage is then an established measure of diversity, based on Rao's (1982) quadratic diversity or entropy (QE):

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

If $d_{ij}=1$ for all $i\neq j$, then the measure of diversity (Q) reduces to

$$T_2(\mathbf{p})=1-\sum_{i}^{N}p_i^2\ ,$$

which is the Gini-Simpson index and is a special case (when q=2) of a more generalized measure of entropy (the Tsallis entropy: T_q):

$$T_q(\mathbf{p})=\frac{1-\sum_{i}^{N}p_i^q}{q-1},\ q>0,\ q\neq 1$$

When $q\to 1$, $T_q(\mathbf{p})\underset{q\to 1}{\to}-\sum_{i=1}^{N}p_i\ln(p_i)$ and this is the Shannon entropy.

The use of quadratic entropy (Q) is then a generalization of the Tsallis entropy when q=2 that includes distances among species and can be used to measure functional and phylogenetic diversity. In the particular case where the d_{ij}'s are defined in the range [0,1], then $Q \leq T_2$. The maximum of T_2 is obtained when species' abundances are even (p_k=1/N for all i): $\max(T_2) = (N-1)/N$.

The Tsallis entropy is a function of Hill numbers (Hill 1973):

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

Indeed,

$$T_q(\mathbf{p}) = \frac{\left[{}^{q}D(\mathbf{p})\right]^{1-q}-1}{1-q},\ q>0,\ q\neq 1$$

$$\lim_{q\to 1}\left[T_q(\mathbf{p})\right] = \ln\left(\lim_{q\to 1}\left[{}^{q}D(\mathbf{p})\right]\right)$$

$1/\left[1-T_2(\mathbf{p})\right]$ thus is Hill number of order 2. It is an effective number of species: it is equal to the number of equally abundant species needed to obtain the same value of diversity. If the the d_{ij}'s are defined in the range [0,1], $1/\left[1-Q(\mathbf{p})\right]$ can then be interpreted as an effective number of evenly abundant and maximally dissimilar species (Ricotta & Szeidl 2009). Using $1/\left[1-Q(\mathbf{p})\right]$, abundant species are given high weight compared to rare species.

Pavoine et al. (2005) introduced the vector of species' abundances that maximizes Q, varying **p** and fixing **D**, as an index of species' originality (named QE-based index). However, this vector can contain null values (zeros) and several distinct vectors can simultaneously maximize Q for a given matrix **D**. To avoid this, Pavoine et al. (2005) proposed the restricted use of their index of species' originality to distances that have ultrametric properties (i.e. $d_{ij} \leq \max(d_{ik}, d_{jk})$ for all i,j,k) including distances obtained from trees with constant tip-to-root distance (distances = half the sum of branch lengths on the shortest path that connects two tips in a tree).

Details on the new index of biodiversity, *R*

In the main text, we decreased the importance given to the most abundant species in Q as follows:

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

The rest of the reasoning below follows exactly the same steps as for the QE-based index. For d_{ij} = 1 for all $i \neq j$, R reduces to

$$S_2(\mathbf{p}) = \left(\sum_i^N \sqrt{p_i}\right)^2 - 1,$$

S_2 is an established statistical metric (Behara & Chawla 1974) and physical quantity (Yamano 2014) as a special case of the generalized γ-entropy with γ=2:

$$S_\gamma(\mathbf{p}) = \frac{1 - \left[\sum_i^N (p_i)^{1/\gamma}\right]^\gamma}{1 - 2^{\gamma-1}}, \ \gamma > 0, \ \gamma \neq 1$$

When $\gamma \to 1$, $S_\gamma(\mathbf{p}) \underset{\gamma\to 1}{\to} -\sum_{i=1}^N p_i \log_2(p_i)$. R is thus a generalization of the γ-entropy when γ=2 that includes distances among species and thus eventually measures functional and phylogenetic diversity. In the particular case where the d_{ij}'s are defined in the range [0,1], then $R \leq S_2$. The maximum of S_2 is obtained when species' abundances are even (p_k=1/N for all i): $\max(S_2) = N - 1$.

The generalized γ-entropy has been underexplored in ecology when analyzing biodiversity. As for Tsallis entropy, the generalized γ-entropy is a function of Hill numbers (Hill 1973). Indeed, considering q=1/γ,

$$S_\gamma(\mathbf{p}) = \frac{1 - \left[\sum_i^N (p_i)^q\right]^{1/q}}{1 - 2^{(1-q)/q}} = \frac{1 - \left[{}^qD(\mathbf{p})\right]^{(1-q)/q}}{1 - 2^{(1-q)/q}}, \ q = 1/\gamma, \ \gamma \neq 1$$

$$\lim_{\gamma\to 1}\left[S_\gamma(\mathbf{p})\right] = \ln\left(\lim_{q\to 1}\left[{}^qD(\mathbf{p})\right]\right) / \ln(2)$$

As a particular case, $S_2(\mathbf{p}) = {}^{0.5}D(\mathbf{p}) - 1$. $S_2 + 1$ thus is Hill number of order 0.5. If the d_{ij}'s are defined in the range [0,1], R+1 can thus be interpreted as an equivalent number of evenly abundant and maximally dissimilar species. In R and R+1, rare species are given significant weight compared to Q and 1/[1-Q].

If the dissimilarities vary in the range [0,1] and if both the distances among species and the abundances of species vary, the maximum of R is N-1. In contrast, if the distances among species are fixed (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) and only the abundances of the species vary, the maximum of R is the first eigenvalue of matrix **D** and the unique maximizing vector is its first eigenvector and has positive values (proof above). This maximizing vector was used by Champely and Chessel (2002) to rescale Q between 0 and 1.

The contribution of species to biodiversity

We provide in Fig. A.1 an illustration of the fact that increasing the abundance of original species compared to redundant species to a certain extent increases the diversity of the set of species.

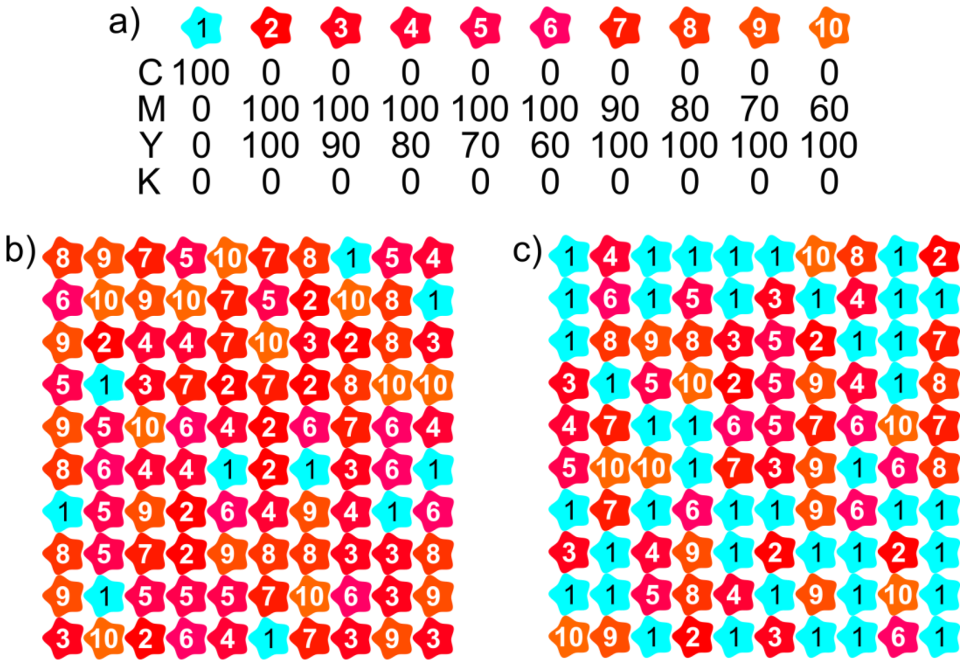

Fig. A.1. Theoretical example illustrating the increase in diversity when increasing the frequency of the most original species. Here species are symbolized by flowers and sets of species by fields of flowers. Each species numbered from 1 to 10 is associated with a color. We selected the CMYK color model used in color printing to define the color of the 10 flowers. The CMYK model is based on three elementary colors (C=cyan, M=magenta, Y=yellow), and on the black color (K). The percentage of each color used to define each flower is indicated in a). Species#1 is by far the most original, having only the cyan color, while others were created by mixing magenta and yellow. In b) each species is represented by 10 flowers. In c) the species with the most original color (species#1) is represented by 42 flowers, species#2 to #4 and species#7 and #8 by 6 flowers each and species#5, #6, #9, #10 by 7 flowers each. To obtain the number of flowers per species in c) we calculating Manhattan distance between the color profiles of species (given in a) and applied *Rb* to this inter-specific distance matrix.

References

Behara, M., Chawla, J.S., 1974. Generalized gamma-entropy. Selecta statistica canadiana 2, 15−38.

Champely, S., Chessel, D. 2002. Measuring biological diversity using Euclidean metrics. Environmental and Ecological Statistics, 9, 167−177.

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

Hill, M.O., 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54, 427−432.

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

Noutsos, D. 2006. On Perron–Frobenius property of matrices having some negative entries. Linear Algebra and its Applications 412, 132–153.

Pavoine, S., Ollier, S., Pontier, D., 2005. Measuring diversity from dissimilarities with Rao's quadratic entropy: are any dissimilarities suitable? Theoretical population biology 67, 231–239.

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

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

Seber, G.A.F. 2008. A matrix handbook for statisticians. Wiley, Hoboken, New Jersey.

Yamano, T., 2014. Thermodynamic meaning of the index of γ-entropy. Results in Physics 4, 50–51.

Appendix B. Details on the second case study

1. Selection of the traits

We discarded 6 traits, although they were available in Pantheria, because of their high correlation with body mass (see Figures B.1 to B.6): codes for those traits in Pantheria are "X13.1_AdultHeadBodyLen_mm", "X9.1_GestationLen_d", "X15.1_LitterSize", "X17.1_MaxLongevity_m", "X5.3_NeonateBodyMass_g", "X23.1_SexualMaturityAge_d".

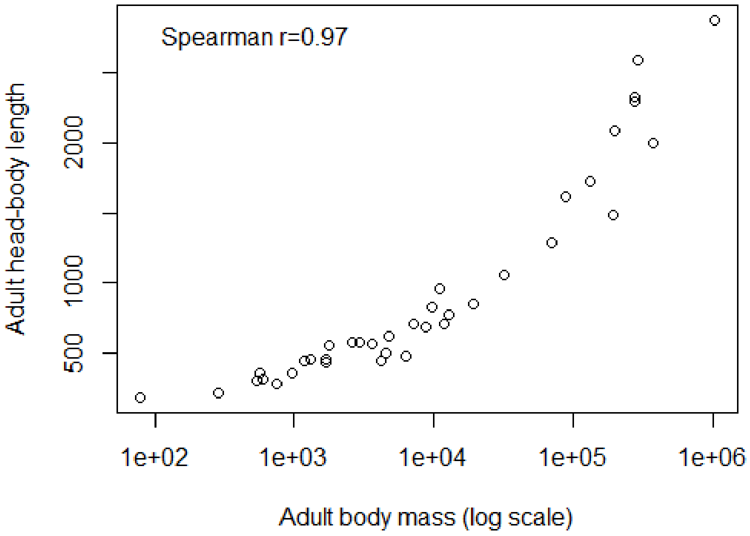

Fig. B.1 As expected body mass was highly correlated with body length

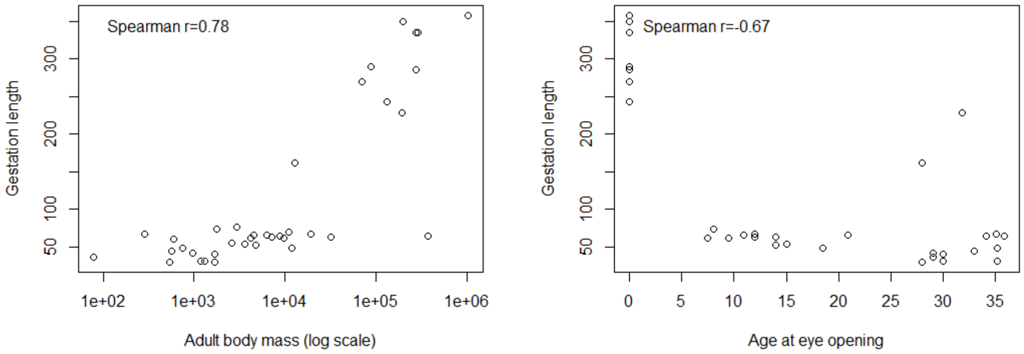

Fig. B.2 With few exceptions, high gestation length was associated with large body mass and eye opening at birth.

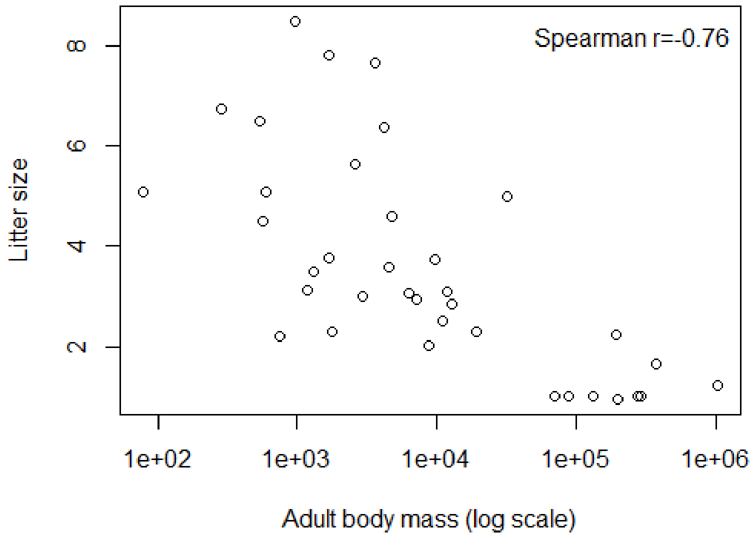

Fig. B.3 Litter size decreased with body mass

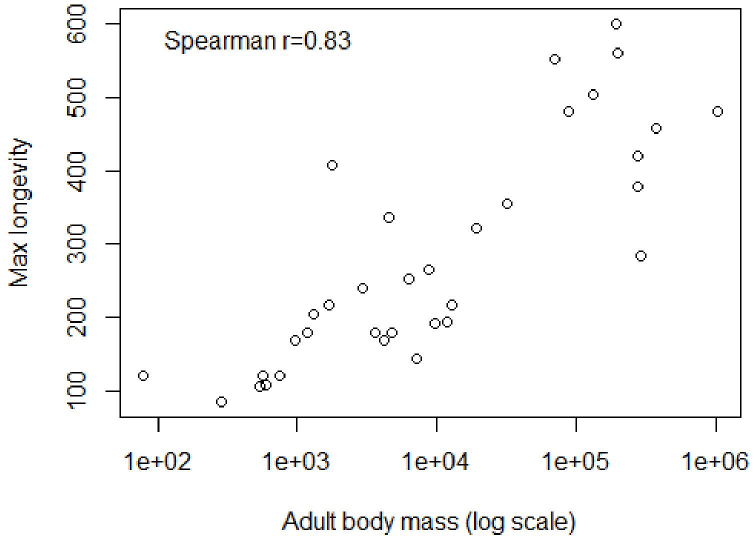

Fig. B.4 Maximum longevity increased with body mass

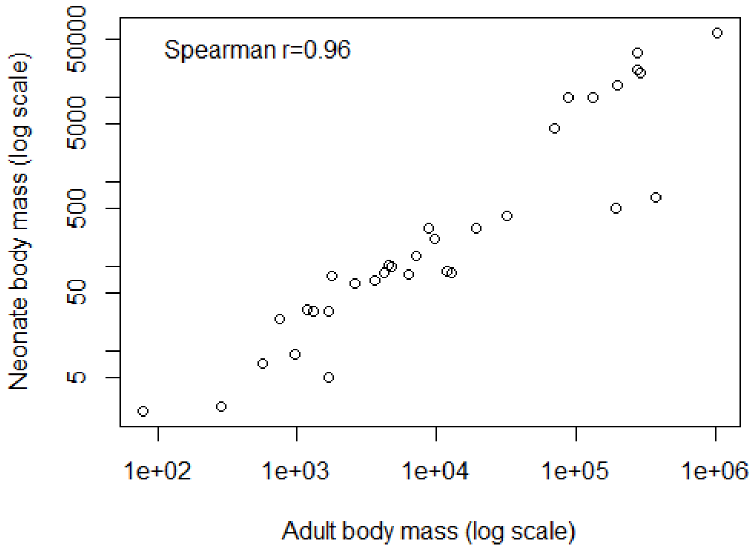

Fig. B.5 Adult and neonate body mass were highly correlated

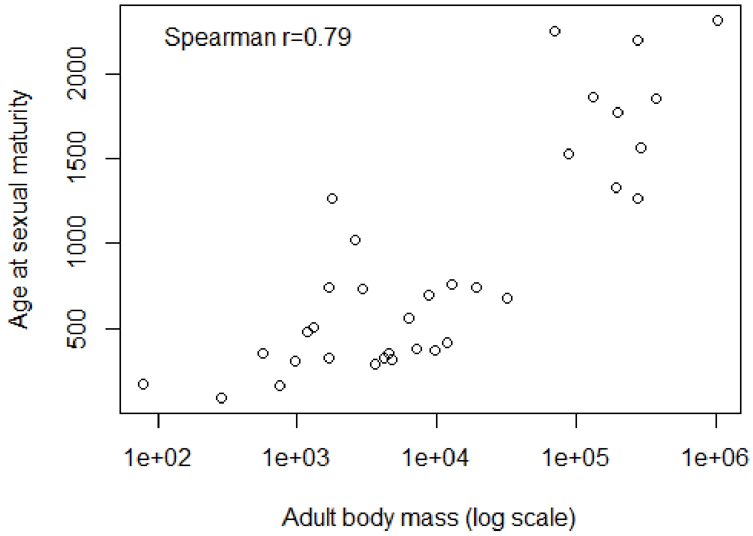

Fig. B.6 Age at sexual maturity, as maximum longevity, increased with body mass

2. Tests for phylogenetic signal

The phylogenetic tree used in this case study is displayed in Figure B.7.

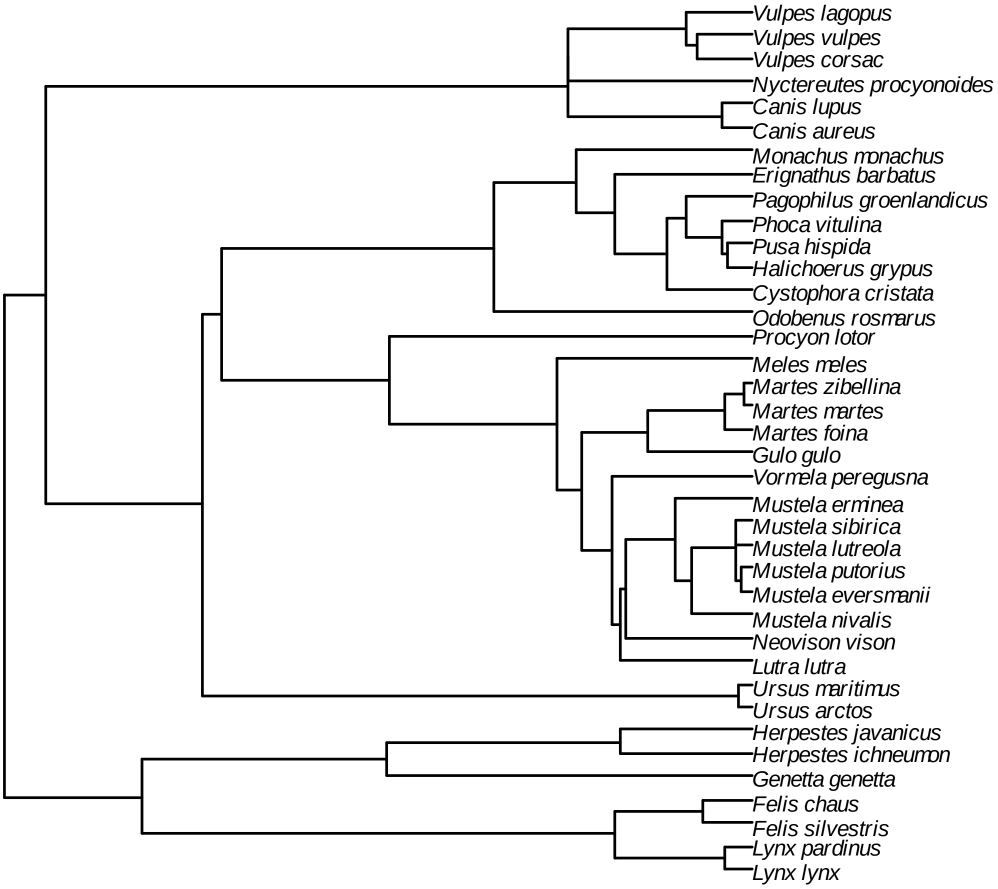

Fig. B.7. Phylogenetic tree for the second case study (extracted from Nyakatura and Bininda-Edmonds, 2012)

We used permutation tests for phylogenetic signal and 1000 permutations of data in each test.

Phylogenetic signal in trophic level was significant: Maddison and Slatkin test, Pvalue = 0.015, 2 levels, observed number of evolutionary transitions = 8, median of this number obtained by randomization = 13 (min. 7, max 16) [Maddison and Slatkin test was applied to all traits with unordered transitions among all levels (even rates)]. Details are given in Fig. B.8.

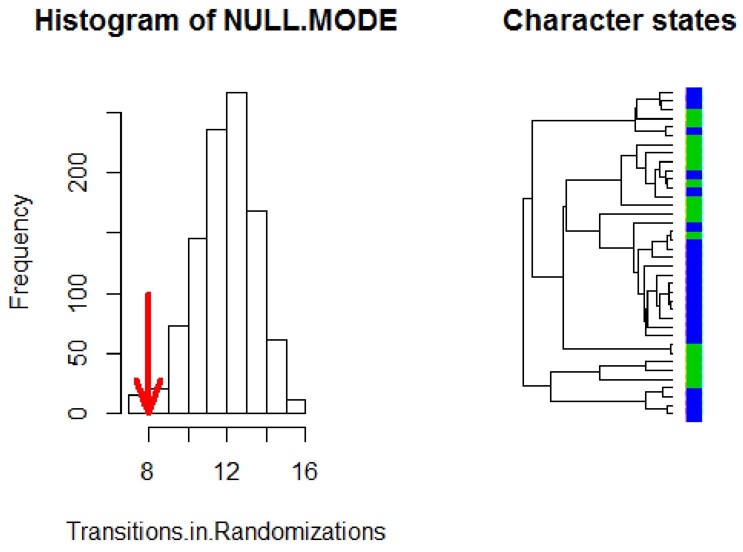

Fig. B.8 Details on Maddison and Slatkin (1991) test applied to trophic level providing on the left the histogram of null number of transitions and the observed number of transitions (red arrow), and on the right the phylogenetic tree with each color on the tips showing different trait states [Blue = carnivore; Green = omnivore].

Phylogenetic signal in terrestriality was not tested as all species were classified fossorial and/or ground dwelling only in Pantheria except Procyon_lotor which was classified as above ground dwelling.

Phylogenetic signal in activity cycle was not significant: Maddison and Slatkin test, Pvalue = 0.379, 3 levels, observed number of evolutionary transitions = 12, median of this number obtained by randomization = 13 (min. 8, max 14). Details are given in Fig. B.9.

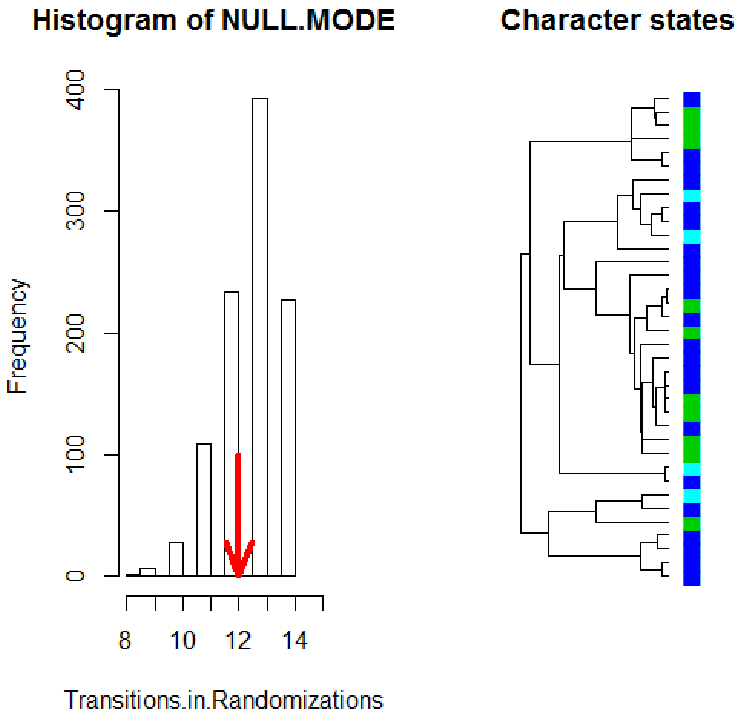

Fig. B.9 Details on Maddison and Slatkin (1991) test applied to activity cycle [on 36 species due to missing data]: on the left the histogram of null number of transitions and the observed number of transitions (red arrow), and on the right the phylogenetic tree with each color on the tips showing different trait states [Green = nocturnal only; Dark blue = nocturnal/crepuscular, cathemeral, crepuscular or diurnal/crepuscular; Light blue = diurnal only].

Phylogenetic signal was significant in body mass (K*=0.98, P-value=0.001; Kw=0.73, P-value=0.001), age at eye opening (K*=1.12, P-value=0.001; Kw=0.87, P-value=0.001), interbirth interval (K*=0.55, P-value=0.002; Kw=0.78, P-value=0.001), weaning age (K*=0.62, P-value=0.001; Kw=0.92, P-value=0.02).

Diet breadth and habitat breadth are discrete quantitative traits. We tested phylogenetic signal in these traits using Maddison and Slatkin (1991) test but considering different rates of transitions between levels (rates are equal only between similar differences in breadth for example from 1 to 2 and from 2 to 3, but the rate of transitions from 1, 2, or 3 to 6 is assumed to be different from that between 1 and 2). Habitat breadth is the number of habitat layers used by each species from 1 to 3. Diet breadth is the number of dietary categories eaten by each species from 1 to 6. Phylogenetic signal was significant in habitat breadth (P-value = 0.000, 3 levels, observed number of evolutionary transitions = 6, median of this number obtained by randomization = 12 [min. 7, max 13]) and diet breadth (P-value = 0.009, 4 levels, observed number of evolutionary transitions = 28, median of this number obtained by randomization = 40 [min. 27, max 49]); details in Figs B.10 and B.11.

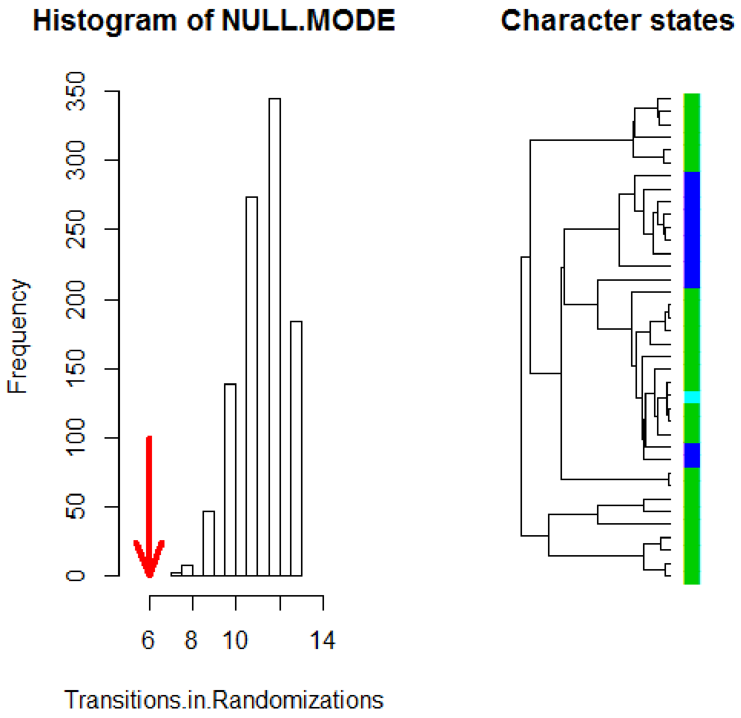

Fig. B.10 Details on Maddison and Slatkin (1991) test applied to habitat breadth: on the left the histogram of null number of transitions and the observed number of transitions (red arrow), and on the right the phylogenetic tree with each color on the tips showing different trait states [Green = 1 habitat layer is used by the species; Dark blue = 2; Light blue = 3].

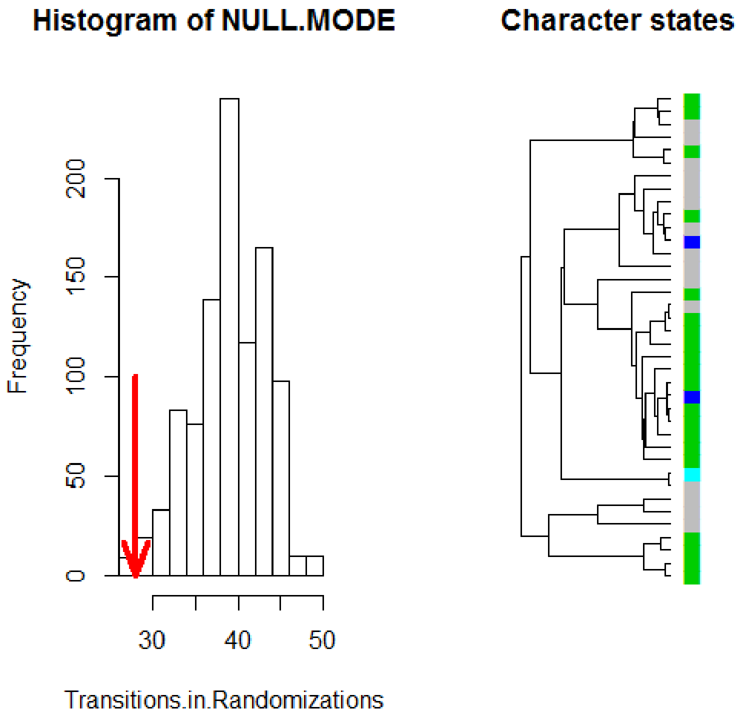

Fig. B.11 Details on Maddison and Slatkin (1991) test applied to diet breadth: on the left the histogram of null number of transitions and the observed number of transitions (red arrow), and on the right the phylogenetic tree with each color on the tips showing different trait states [Green = 1 diet category is used by the species; Dark blue = 2; Light blue = 3; Grey = 6].

3. Dealing with identical sets of species.

Translating the problem of measuring originality from phylogeny to a small set of traits raises a critical new question because, according to a limited set of discrete traits (such as nominal or binary traits), two or more species might be identical (i.e, these species form an apparently *identical set* of species). Two species thus belong to the same identical set if they have identical values for all traits considered. This raises the question of whether these apparently identical species are considered as not distinctive at all (originality measure = 0). According to PE and NN these species will be indeed non distinctive. However, with the other indices they can be attributes positive originalities. First, a species within a speciose identical set can be considered as less original than a species within an identical set of a few species. Second, sets of identical species are not equally distinct; some might be very different from others in trait space. In this case, species within a relatively distinct set will be considered more functionally distinct than species within less distinct identical sets, all else being equal. Overall, we can state that the originality of a species increases with the originality of the identical set of species to which it belongs, and decreases with the number of species within the identical set.

With the QE-based index (QEb), our solution is to partition the originality of an identical set equally among the species within it. An identical set of S species and originality I will contain species each with a originality value of I/S. As the quadratic entropy is unchanged by splitting a set of identical species (Shimatani 2001), the vector of originality defined by this process also maximizes the quadratic entropy (which is the criterion used by Pavoine *et al.* 2005 to define QEb; Pavoine and Bonsall 2009).

With the Fair-proportion (FP) and Equal-Split (ES) indices, the tree is modified by adding artificial branches of length 0 at the tips. Each tip represents an identical set and there are as many artificial branches at that tip as there are species within the identical set. Note that with this rule, ES shares the following property with QEb: if an identical set contains S species and has originality equal to I, then the originality of each species in the set is I/S.

With AV, the formula is unaltered as the relative originality of a species decreases with the presence of several zero distances with other species. Consider that a species is the sole member of its identical set and that its originality value is D/n, where D is the sum of all distances between this focal species and the other n species. Now we add s species identical to the focal species. The total distance D remains unchanged, but the number of distances increases by s: the originality of the species is reduced to $D/(n+s)$.

With Rb, as with AV, the formula was unaltered (although for a given i, the d_{ij} were allowed to be zero for one or several j but not all) and the observed correlations between AV and Rb were maintained with sets of identical species in our Carnivora case study as shown below.

4. Analyses per trait or group of traits

Correlations between phylogenetic and functional originalities were low for all traits except those related to the reproduction (Table B.1). They were even negative for some traits like those related to the habitat and for example for body size when indices Rb and AV are used to measure originality. The effect of the use of a clustering approach before applying Rb and AV depends on the traits used. For example using a clustering approach increased correlations between functional and phylogenetic originalities as measured by Rb and AV with all combined traits and with body mass. However those correlations were decreased with reproduction and the Ward clustering approach.

In general, the choice of the clustering approach (Ward versus UPGMA) had low effect on the vectors of functional originalities but this also depended on the traits considered and on the originality index selected (Table B.2). For example, the lowest correlations between the originality vectors obtained after the Ward versus the UPGMA approach were observed with body size, particularly when using tb-AV and tb-Rb (Table B.2).

Regarding the link between our new originality index Rb and the other originality indices, this also depended on the traits used, except for AV (Table B.3). Correlations between Rb and AV were close to unity whatever the traits used. Correlations between Rb and uniqueness (NN and PE) varied from 0 with body mass to 0.6 with all combined traits. Correlations between Rb and tb-Rb were higher with the UPGMA approach implying that, at least in this example, the UPGMA may less distort the functional distances among species than the Ward approach.

Introduced species had higher originality according to the phylogeny, all traits together, and behaviour traits (Table B.4). Results for other traits were more mitigated and depended on the originality index chosen. For body size, introduced species were (in terms of median value) more original than other species according to PE, whereas the most original species were the marginal ones according to the other indices. For diet traits, the most original species were either introduced species alone (with M and ES) or both marginal and introduced species (according to the other indices, except uniqueness, i.e. PE and NN). For habitat traits, the most original species were the introduced ones (with M and PE) or the marginal ones (with FP, tb-AV, tb-Rb, AV and Rb). For reproduction traits, the most original species were the introduced ones (with tree-based indices and NN) or the marginal ones (with AV and Rb, with five out of seven marginal species being more original than the most original introduced species with AV and Rb and only one with tb-AV and tb-Rb). The native species were thus the least phylogenetically and functionally original. However, whether the introduced or the marginal species were the most original depended on the index and traits used. Although both used to measure uniqueness, PE and NN disagreed on which of the introduced or marginal species were the most original according to body mass (consensus was obtained with the other traits).

No correlation was observed between the phylogenetic originality of native species and their extinction risk (IUCN Red List status, Table B.5). But there was a weak evidence for a positive relation between the extinction risk and the functional originality, more precisely the strict functional uniqueness, measured by PE for all traits except behaviour traits and body size and according to most originality indices for habitat and reproduction traits. The highest correlations were obtained with PE and NN, both characterizing the strict uniqueness, in habitat traits and all combined traits (correlations between 0.36 and 0.40, Table B.5).

Table B.1 Spearman correlations between functional and phylogenetic originalities in the Carnivora case study

	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
All traits - Ward	0.06	0.14	0.30	0.26	0.20	0.21	0.21	0.11	0.06	0.19
All traits - UPGMA	-0.19	0.15	0.30	0.24	0.20	0.19	0.19			
Behaviour - Ward	-0.08	0.26	-0.02	0.16	0.16	0.02	0.03	0.04	0.07	0.26
Behaviour - UPGMA	-0.08	0.26	-0.02	0.16	0.16	0.03	-0.03			
Body Mass - Ward	0.09	0.11	0.10	0.16	0.11	0.02	0.01	-0.38	-0.40	0.06
Body Mass - UPGMA	-0.05	0.09	0.08	0.09	0.02	-0.22	-0.21			
diet - Ward	-0.31	-0.21	-0.23	0.30	0.27	0.34	0.33	0.34	0.34	-0.21
diet - UPGMA	-0.31	-0.21	-0.26	0.30	0.27	0.34	0.37			
Habitat - Ward	-0.37	-0.18	-0.30	-0.11	-0.20	-0.22	-0.22	-0.22	-0.22	-0.18
Habitat - UPGMA	-0.37	-0.18	-0.29	-0.11	-0.20	-0.22	-0.22			
Reproduction - Ward	0.32	0.44	0.49	0.45	0.43	0.04	0.03	0.13	0.12	0.34
Reproduction - UPGMA	0.24	0.45	0.47	0.51	0.36	0.11	0.11			

* These indices are independent on the clustering approach; for phylogenetic originalities only, AV=tb-AV, Rb=tb-Rb, NN=PE

Table B.2 Correlations between functional originalities obtained with Ward versus UPGMA clustering in the Carnivora case study

	M	PE	ES	FP	Qb	tb-AV	tb-Rb
All traits	0.72	1.00	0.92	0.99	0.94	0.96	0.96
Behaviour	1.00	1.00	1.00	1.00	1.00	0.97	1.00
Body Mass	0.77	1.00	0.97	0.86	0.88	0.58	0.62
Diet	1.00	1.00	1.00	1.00	1.00	1.00	0.97
Habitat	1.00	1.00	1.00	1.00	1.00	1.00	0.99
Reproduction	0.96	1.00	0.99	0.98	0.97	0.92	0.91

Table B.3 Spearman correlations between Rb and the other originality indices in the Carnivora case study using groups of traits

	M	PE	ES	FP	Qb	tb-AV	tb-Rb*	AV*	NN*
Ward									
All traits	0.62	0.62	0.69	0.85	0.77	0.89	0.89	0.99	0.62
Behaviour	0.56	0.30	0.59	0.95	0.95	0.92	0.95	0.99	0.30
Body mass	0.16	0.06	0.17	0.46	0.22	0.49	0.53	0.99	0.04
Diet	0.29	0.31	0.33	0.93	0.93	1.00	0.99	1.00	0.31
Habitat	0.45	0.34	0.48	1.00	1.00	1.00	1.00	1.00	0.34
Reproduction	0.66	0.56	0.63	0.77	0.65	0.72	0.71	1.00	0.48
UPGMA									
All traits	0.34	0.62	0.80	0.83	0.77	0.90	0.90		
Behaviour	0.59	0.30	0.58	0.95	0.95	0.95	0.95		
Body mass	0.63	0.06	0.30	0.60	0.55	0.85	0.85		

Diet	0.29	0.31	0.33	0.93	0.93	1.00	0.98		
Habitat	0.51	0.34	0.51	1.00	1.00	1.00	0.99		
Reproduction	0.65	0.59	0.62	0.73	0.71	0.83	0.83		

* AV and NN are independent of the clustering method (Ward versus UPGMA); with phylogenetic data, tb-AV=AV, tb-Rb=Rb and PE=NN.

Table B.4 Median of species relative originalities (expressed as %) per group of species: Introduced species (Int), Marginal species (Mar), Native species (Nat) (Carnivora Case Study)

Phylogeny										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
Int	3.60	4.72	4.39	4.22	6.05	2.80	2.88			
Mar	2.00	2.00	2.05	2.11	1.27	2.49	2.36			
Nat	2.67	0.90	2.42	2.57	2.49	2.48	2.33			
All traits - Ward										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
Int	3.25	5.12	4.18	3.43	4.34	2.92	3.01	2.92	3.04	5.39
Mar	2.41	1.47	1.79	2.43	1.70	2.89	2.95	2.81	2.83	1.11
Nat	2.41	1.57	2.21	2.52	2.22	2.58	2.60	2.50	2.43	1.25
All traits - UPGMA										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb			
Int	2.69	5.04	3.82	3.73	3.63	2.94	3.08			
Mar	2.24	1.38	1.61	2.04	1.11	2.85	2.89			
Nat	2.24	1.66	2.28	2.15	1.49	2.51	2.45			
Behaviour - Ward										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
Int	4.16	0	5.09	4.58	4.75	3.59	3.60	3.89	4.18	0
Mar	2.05	0	0.18	1.38	0.99	1.93	1.83	1.99	1.79	0
Nat	1.68	0	0.05	1.38	0.99	1.93	1.83	1.99	1.79	0
Behaviour - UPGMA										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb			
Int	3.69	0	4.49	3.99	3.66	3.81	4.02			
Mar	1.85	0	0.18	1.12	0.90	1.87	1.69			
Nat	1.68	0	0.05	1.12	0.90	1.87	1.69			
Body mass - Ward										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
Int	2.39	1.15	1.75	2.33	1.85	2.53	2.52	2.19	1.95	1.43
Mar	2.87	0.98	2.66	2.51	2.70	2.80	2.85	3.11	3.03	1.47
Nat	2.39	0.82	1.85	2.29	1.95	2.61	2.67	2.16	1.89	1.06
Body mass - UPGMA										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb			
Int	2.07	1.03	1.72	1.83	1.10	2.25	2.09			
Mar	2.21	0.94	1.91	2.19	1.32	3.45	3.65			
Nat	2.21	0.78	1.60	1.84	1.31	2.28	2.12			
Diet - Ward										

	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
Int	2.39	0	0.48	2.86	1.67	3.10	3.09	3.12	3.12	0
Mar	1.89	0	0.10	2.86	1.67	3.10	3.09	3.12	3.12	0
Nat	1.71	0	0.05	2.08	1.25	2.26	2.25	2.25	2.24	0
Diet - UPGMA										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb			
Int	2.39	0	0.46	2.49	1.74	3.08	3.03			
Mar	1.89	0	0.11	2.49	1.74	3.08	3.03			
Nat	1.71	0	0.04	1.79	1.19	2.24	2.19			
Habitat - Ward										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
Int	2.46	0	0.39	2.69	2.10	3.04	3.00	2.87	2.54	0
Mar	1.64	0	0.03	3.71	2.33	4.17	4.12	3.87	3.31	0
Nat	1.79	0	0.04	1.67	1.88	1.90	1.89	1.86	1.77	0
Habitat - UPGMA										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb			
Int	2.31	0	0.36	1.90	1.77	2.89	2.59			
Mar	1.63	0	0.03	2.49	2.49	3.97	3.53			
Nat	1.63	0	0.03	1.30	1.05	1.82	1.66			
Reproduction - Ward										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
Int	2.71	4.06	3.05	2.44	2.82	2.59	2.50	2.06	1.66	1.61
Mar	2.11	1.10	1.33	1.75	1.27	2.59	2.50	2.75	2.64	0.30
Nat	2.46	0.63	1.81	1.80	1.48	2.59	2.50	2.36	2.11	0.55
Reproduction - UPGMA										
	M	PE	ES	FP	Qb	tb-AV	tb-Rb			
Int	2.25	3.01	2.91	2.36	1.42	2.30	2.00			
Mar	1.87	0.76	1.27	1.36	0.59	2.28	1.97			
Nat	1.87	0.60	1.27	1.47	0.52	2.28	1.96			

* These indices are independent on the clustering approach; for phylogenetic originalities only, AV=tb-AV, Rb=tb-Rb, NN=PE

Table B.5 Spearman correlation between originality and extinction risks (Carnivora case study)

	M	PE	ES	FP	Qb	tb-AV	tb-Rb	AV*	Rb*	NN*
Phylogeny	0.01	0.14	0.14	0.08	0.05	-0.12	-0.13			
All traits - Ward	0.24	0.36	0.32	0.26	0.27	0.12	0.13	0.26	0.29	0.40
All traits - UPGMA	0.18	0.36	0.24	0.27	0.22	0.14	0.14			
Behaviour - Ward	0.05	0.00	0.05	0.09	0.09	0.09	0.09	0.11	0.09	0.00
Behaviour - UPGMA	0.04	0.00	0.04	0.09	0.09	0.09	0.09			
Body Size - Ward	-0.38	-0.12	-0.31	-0.02	-0.34	0.24	0.21	0.15	0.13	-0.15
Body Size - UPGMA	-0.23	-0.10	-0.27	-0.19	-0.23	-0.19	-0.19			
Diet - Ward	0.06	0.29	0.05	0.18	0.18	0.11	0.11	0.11	0.11	0.29
Diet - UPGMA	0.06	0.29	0.05	0.18	0.18	0.11	0.07			

Habitat - Ward	0.19	0.39	0.20	0.33	0.33	0.33	0.33	0.33	0.33	0.39
Habitat - UPGMA	0.19	0.39	0.20	0.33	0.33	0.33	0.33			
Reproduction - Ward	0.19	0.35	0.26	0.32	0.25	0.25	0.26	0.23	0.23	0.15
Reproduction - UPGMA	0.20	0.38	0.25	0.34	0.27	0.22	0.22			

* These indices are independent on the clustering approach; for phylogenetic originalities only, AV=tb-AV, Rb=tb-Rb, NN=PE

Appendix C - Carnivora species traits (legends for traits can be found at http://esapubs.org/Archive/ecol/E090/184/metadata.htm)

	X1.1 Activity Cycle	X5.1 Adult Body Mass (g)	X2.1_Age at Eye Opening (d)	X6.1 Diet Breadth	X12.1 Habitat Breadth	X14.1 Interbirth Interval (d)	X12.2 Terrestriality	X6.2 Trophic Level	X25.1 Weaning Age (d)
Canis_aureus	2	9658.7	7.5	6	1	365	1	2	61.3
Canis_lupus	2	31756.51	14.01	1	1	365	1	3	44.82
Erignathus_barbatus	3	279999.99	0	6	2	547.5	1	2	20.94
Genetta_genetta	1	1756.17	8	6	1	182.5	1	2	71.8
Cystophora_cristata	3	278896.81	0	6	2	365	1	2	9.5
Felis_chaus	2	7157.99	11.94	1	1	130.01	1	3	95.49
Gulo_gulo	1	12792.49	27.94	1	1	821.25	1	3	83.64
Halichoerus_grypus	NA	197570.01	0	2	2	365	1	3	18.37
Herpestes_ichneumon	2	2980.02	49	6	1	365	1	2	58.09
Herpestes_javanicus	3	750	18.5	6	1	152.08	1	2	43.59
Felis_silvestris	2	4573.08	10.95	1	1	148.25	1	3	76.01
Monachus_monachus	2	294881.33	0	6	2	730	1	2	42.75
Meles_meles	2	11884.03	35.16	1	1	365	1	3	91.3
Lutra_lutra	1	8868.69	35.87	1	2	365	1	3	100.8
Lynx_lynx	2	19300	11.94	1	1	365	1	3	81.85
Lynx_pardinus	2	11050	NA	1	1	365	1	3	70
Martes_foina	2	1675	NA	1	1	365	1	3	74.11
Martes_martes	1	1299.99	29.96	1	1	365	1	3	54.19
Martes_zibellina	2	1173.91	35.16	6	1	NA	1	2	52.23
Mustela_erminea	2	284.5	35.1	1	1	365	1	3	60.93
Mustela_eversmanii	1	1684.21	30	1	1	365	1	3	48.8
Mustela_lutreola	2	566.44	33	2	3	365	1	3	58.71
Mustela_nivalis	2	78.45	29.08	1	1	273.75	1	3	36.74
Mustela_putorius	1	975.55	29.08	1	1	252.5	1	3	54.39
Mustela_sibirica	2	530.85	27.94	1	1	365	1	3	58.2
Neovison_vison	1	928.5	24.5	1	2	365	1	3	45
Nyctereutes_procyonoides	1	4214.99	9.5	6	1	365	1	2	49.69

Odobenus_rosmarus	2	1042996.25	0	6	2	821.25	1	2	591.36
Pagophilus_groenlandicus	NA	132250	0	6	2	365	1	2	11.45
Phoca_vitulina	2	87316.66	0	1	2	365	1	3	30.74
Procyon_lotor	2	6373.72	20.91	6	2	365	2	2	104.16
Pusa_hispida	2	70963.6	0	6	2	365	1	2	45.23
Ursus_arctos	2	196287.5	31.82	6	1	912.5	1	2	182.5
Ursus_maritimus	3	371703.81	34.12	3	1	831.67	1	2	205.17
Vormela_peregusna	2	594.13	39	1	1	365	1	3	52
Vulpes_corsac	1	2615.33	NA	6	1	365	1	2	54.56
Vulpes_lagopus	2	3584.37	15.03	1	1	365	1	3	49.5
Vulpes_vulpes	1	4820.36	14.01	1	1	365	1	3	50.71

Appendix D Median of the correlations between the originality vectors calculated with UPGMA trees and those calculated with Ward trees in the simulation case study

BIRTH

	M	PE	ES	FP	Qb	tb-AV	tb-Rb
1 trait	0.81	1.00	0.91	0.93	0.91	0.86	0.87
25 traits	0.63	1.00	0.86	0.95	0.78	0.74	0.75
50 traits	0.60	1.00	0.86	0.95	0.76	0.74	0.74

ROOT

	M	PE	ES	FP	Qb	tb-AV	tb-Rb
1 trait	0.81	1.00	0.91	0.93	0.91	0.83	0.84
25 traits	0.50	0.99	0.85	0.93	0.60	0.38	0.39
50 traits	0.34	0.99	0.83	0.93	0.38	0.26	0.26

TIPS

	M	PE	ES	FP	Qb	tb-AV	tb-Rb
1 trait	0.82	1.00	0.91	0.93	0.91	0.87	0.87
25 traits	0.72	1.00	0.86	0.93	0.86	0.86	0.86
50 traits	0.70	1.00	0.85	0.93	0.86	0.87	0.87

N-U

	M	PE	ES	FP	Qb	tb-AV	tb-Rb
1 trait	0.81	1.00	0.91	0.93	0.92	0.85	0.86
25 traits	0.77	0.99	0.87	0.92	0.86	0.77	0.78
50 traits	0.78	0.99	0.86	0.92	0.86	0.78	0.78

Appendix E. Median of the correlations between Rb and the other indices in the simulation case study

Tree-based indices							Approaches used:		
M	PE	ES	FP	Q	tr-AV	tr-Rb	UPGMA	NbofTraits	Model
0.49	0.34	0.48	0.73	0.56	0.77	0.78	Y	1	BIRTH
0.54	0.1	0.21	0.23	0.54	0.61	0.61	Y	25	BIRTH
0.54	0.11	0.22	0.25	0.57	0.66	0.66	Y	50	BIRTH
0.44	0.33	0.4	0.8	0.49	0.72	0.74	N	1	BIRTH
0.42	0.1	0.3	0.37	0.51	0.48	0.49	N	25	BIRTH
0.35	0.11	0.29	0.37	0.48	0.51	0.51	N	50	BIRTH
0.51	0.37	0.5	0.76	0.59	0.82	0.82	Y	1	ROOT
0.73	0.31	0.44	0.44	0.71	0.69	0.69	Y	25	ROOT
0.7	0.25	0.36	0.35	0.67	0.67	0.67	Y	50	ROOT
0.44	0.37	0.42	0.83	0.51	0.74	0.76	N	1	ROOT
0.5	0.31	0.56	0.54	0.59	0.34	0.34	N	25	ROOT
0.39	0.26	0.5	0.47	0.47	0.28	0.28	N	50	ROOT
0.46	0.33	0.46	0.69	0.54	0.75	0.76	Y	1	TIPS
0.54	0.08	0.22	0.25	0.57	0.69	0.69	Y	25	TIPS
0.56	0.09	0.23	0.26	0.6	0.76	0.76	Y	50	TIPS
0.44	0.32	0.38	0.77	0.48	0.71	0.73	N	1	TIPS
0.47	0.08	0.31	0.43	0.55	0.63	0.63	N	25	TIPS
0.43	0.09	0.3	0.45	0.55	0.68	0.68	N	50	TIPS
0.74	0.57	0.73	0.84	0.74	0.82	0.83	Y	1	NU
0.69	0.31	0.52	0.54	0.65	0.77	0.77	Y	25	NU
0.7	0.32	0.52	0.55	0.66	0.78	0.79	Y	50	NU
0.67	0.58	0.69	0.92	0.75	0.61	0.68	N	1	NU
0.68	0.31	0.6	0.7	0.74	0.5	0.52	N	25	NU
0.68	0.31	0.61	0.71	0.75	0.5	0.52	N	50	NU

NU = non ultrametric model

Y=UPGMA
N=Ward

Other indices independent of the clustering approaches

		Approaches used:	
AV	NN	NbofTrai	Model
1	0.36	1	BIRTH
1	0.10	25	BIRTH
1	0.11	50	BIRTH
1	0.39	1	ROOT
1	0.32	25	ROOT
1	0.26	50	ROOT
1	0.34	1	TIPS
1	0.09	25	TIPS
1	0.10	50	TIPS
1	0.39	1	NU
1	0.23	25	NU
1	0.23	50	NU