



**HAL**  
open science

# Measuring inter-specific diversity. Benefits of species originality indices in community ecology and conservation biology

Anna Kondratyeva

► **To cite this version:**

Anna Kondratyeva. Measuring inter-specific diversity. Benefits of species originality indices in community ecology and conservation biology. Biodiversity and Ecology. Muséum National d'Histoire Naturelle Paris, 2019. English. NNT: . tel-04066501

**HAL Id: tel-04066501**

**<https://hal.science/tel-04066501v1>**

Submitted on 12 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

MUSÉUM NATIONAL D'HISTOIRE NATURELLE  
École Doctorale 227  
Sciences de la Nature et de l'Homme

Année 2019



N° attribué par la bibliothèque

|\_|\_|\_|\_|\_|\_|\_|\_|\_|\_|\_|\_|\_|\_|\_|\_|

THÈSE

pour obtenir le grade de

DOCTEUR DU MUSÉUM NATIONAL D'HISTOIRE NATURELLE

Spécialité : Ecologie

Présentée et soutenue publiquement par  
Anna KONDRATYEVA

Le 11 décembre 2019

---

**Measuring inter-specific diversity.  
Benefits of species originality indices in community ecology and  
conservation biology**

---

Sous la direction de : **Sandrine Pavoine et Philippe Grandcolas**

Devant le jury :

<b>Mme Sandrine Pavoine</b>	Maitre de conférences au MNHN	Directrice de Thèse
<b>Mr Philippe Grandcolas</b>	Directeur de recherche au CNRS	Directeur de Thèse
<b>Mme Sophie Nadot</b>	Professeure à l'Université Paris-Sud	Rapportrice
<b>Mr Cyrille Violle</b>	Directeur de recherche au CNRS	Rapporteur
<b>Mme Hélène Morlon</b>	Directrice de recherche au CNRS	Examinatrice
<b>Mr Hervé Daniel</b>	Maitre de conférences à l'Agrocampus Ovest	Examineur



## REMERCIEMENTS

---

Se lancer dans une thèse de doctorat est une grande décision mais quand c'est fait c'est comme une porte qui s'ouvre sur un grand monde de la recherche scientifique qui grouille de personnes curieuses, passionnées par leur travail. Je voudrais garder des liens avec ce monde pour pouvoir transmettre ses savoirs auprès des plus jeunes, dans les écoles en tant qu'intervenante ou professeure – on verra...

Je voudrais remercier toutes les personnes qui ont joué un rôle plus au moins important pour moi pendant ces trois années. Cette aventure ne serait pas possible sans ma directrice et mon directeur de thèse. **Sandrine**, merci de m'avoir donnée cette chance de m'essayer dans la recherche. Je te suis reconnaissante d'avoir fait confiance et de m'avoir accompagnée depuis le début, merci pour tes encouragements, ton calme, ta patience et ta sagesse. **Philippe**, ta justesse des mots et l'incroyable vision globale des choses m'épateront toujours, j'ai appris beaucoup à tes côtés.

J'adresse de chaleureux remerciements aux membres de mon jury – **Sophie Nadot, Cyrille Violle, Hélène Morlon et Hervé Daniel** – pour avoir accepté d'évaluer ce travail.

Je remercie également le **Labex BCDiv** pour le financement de ce projet de thèse.

**Anne-Cécile Haussonne** je te remercie de m'avoir aidé dans toutes ces démarches administratives à se casser la tête, sans toi je serai encore en train d'attendre un ordinateur et serais jamais partie aussi simplement en congrès.

Merci à l'Ecole Doctorale 227 et tout particulièrement à **Nathalie Machon** et **Anaïs Ranera** (même si tu as changé de poste) pour tout ce que vous faites pour la vie des doctorants au Muséum.

Je voudrais remercier vivement l'équipe allemande de Helmholtz Centre for Environmental Research – UFZ, Halle (Saale) pour m'y avoir accueilli pendant deux mois et pour m'avoir aidé avec les données et les analyses. Dankeschön à **Walter Durka** qui connaît les espèces végétales sur les bouts des doigts, à **Ingolf Kühn** pour les précisions sur les analyses spatiales et à **Sonja Knapp** pour les idées qui ont aidé à structurer notre travail en commun.

Un travail efficace n'est pas imaginable pour moi sans un environnement agréable. Je voudrais donc remercier tout.e.s mes collègues auprès qui j'ai eu plaisir de passer ces trois années. Tout d'abord merci aux personnes du bureau coquelicot: **Typhaine R.** pour ton inépuisable bonne humeur, même quand les modèles ne tournent plus et il faut faire une pause smarties ; **Gabrielle M.** prête à aider à tout moment, tu me manquais à la fin pour tenir le club de travaille-tard ; **Alain D.**, pour ton grand sourire réconfortant et ta sérénité d'être; **Margaux A.** pour des conseils déteints ; **Fabien C.** pour ton humour qui manque maintenant dans le bureau, **Kévin B.** le grand



sportif toujours discret, mais aussi **Olivier B.**, **Mona O.**, **Ana-Cristina T.**, **Tanguy L-L.**, **Laura T.** (merci pour le voyage à Nancy, c'était presque ma seule session de terrain), **Eduardo**, le dernier arrivé qui commence tout juste, **Nélida P.**, qui es partie trop vite, (j'étais trop contente de se retrouver à Rennes !).

En dehors du bureau coquelicot,  
**Minh-Xuan T.**, toujours soucieux de bien-être de ces collègues, merci pour tes conseils et ta bienveillance, **Jean-François J.** merci pour ton inépuisable stock de connaissances naturalistes, **François D.** merci pour des conseils en stat et des blagues osées;  
Un mot tout particulier à **Adrienne**. Merci d'avoir jamais perdu l'énergie et l'envie de m'écouter, merci pour ces échanges rassurants mais aussi pour des astuces dans le R;

Même si on était séparés entre deux bâtiments, je passais souvent vous faire un coucou et je voudrais particulièrement remercier **Diane** pour ne jamais se prendre la tête, le bureau **des barbuis** qui a disparu au compte de goutte et merci **Théo** (le dernier survivant du bureau) d'y m'avoir accueilli à la fin; **Olivier B.**, on t'a perdu de vue ces derniers temps, mais merci pour les pauses discussion; **Grégoire L.**, tu m'étonnerais toujours avec tes bêtises et tes histoires incroyables. Il est compliqué de citer encore toutes ces personnes que j'ai côtoyées pendant ces trois années au sein de ce laboratoire et qui ont eu d'importance à un moment ou à un autre grâce à leurs conseils, connaissances et tout simplement leur bienveillance. Une pensée pour **Jean-Baptiste M.**, **Colin F.**, **Emmanuelle P.**, **Simon B.**, **Romain L.**, **Anne D.**, **Léo B.**, **Emmanuel C.**, **Pauline C.**, **Julie P.**, **Florence D.**, **Denis C.**, **Thibault F.**, **Guillaume T.**, **Benjamin Y.**, **Léo M.**, **Camila A.**, **Julie M.**, **Benoit F.**, **Sébastien T.**, **Marine L.**, **Maud M.**, **Marine R.**, **Anne-Caroline P.** et enfin à **Romain J.** pour m'avoir acceptée dans ce labo.

Au meilleur laboratoire d'écologie que je n'échangerais pour rien au monde - Merci pour l'ambiance, pour la chaleur, pour les histoires, pour les soirées. Merci **CESCO** !

Comment j'aurais fait sans mes amis pour qui j'étais là de façon chaotique certes, mais vous vous y étiez toujours quand il le fallait. **Jérôme M.** te rappelles-tu que c'était toi qui a lu le mail de résultats des auditions pour avoir cette thèse. Je n'oublierai jamais l'ascenseur émotionnel que tu m'as fait avec une loooongue pause avant de le lire à haute voix, merci pour tout; **Pilou**, mon voisin préféré, je te remercie d'être là dans les moments compliqués de ma vie, tu sais que ça va mieux maintenant et t'y es pour quelque chose. Il y a des personnes qui s'est sont transformée de collègues en coloc puis en amis, oui je parle de vous **Nicolas D.** et **Simon V.**, même si nous ne sommes plus à la coloc family à part Simon ça nous fait un.e 4ème docteur.e. à la maison ☺. **Lesia**, **Moussa**, **Sabina**, **Nasra**, **Guillaume**, **César** – cette coloc m'a transformé également, alors merci pour les soirées, des repas et plein d'autres folies !

Petite pensée à **Badis**, toujours en vadrouille.

J'ai fait des belles rencontres lors de cette thèse, certaines sont devenues des très bonnes amitiés. Merci **Christian** de m'avoir appris comment fonctionnent des ailes des avions – au moins la formation doctorale a servi à quelque chose !

A part de faire avancer mes recherches, j'ai passé deux ans à gérer une association étudiante « Bureau des Doctorants et Etudiants du Muséum » et ça ne serait pas possible sans une équipe (enfin j'en ai connue deux) au taquet. Je voudrais remercier **Margot B.** pour des heures passées sur des dossiers de subventions, des mails et encore des dossiers - tu fais les meilleurs igloo en carton !; **Adrienne et Typhaine** pour rendre la trésorerie plus facile; **Antoine H., Juliette W.**, ainsi que l'équipe de l'année suivante et tout particulièrement **Charly F.** qui a repris enfin ma place et je sais que ne trouverais pas mieux.

Mes grands remerciements de fond du cœur à **ma mère** qui a semé en moi une graine d'amour de la nature depuis que j'étais toute petite et qui m'a soutenue lors de cette aventure française... depuis déjà dix ans ! Merci à **mon père** pour l'attention et des apparitions brèves mais riches en souvenirs. Je remercie tout particulièrement **mon petit frère** pour avoir relu le manuscrit (même si tu avais du mal à comprendre), Андрюша спасибо за проверку английского, даже если ты не всё понял ! Nous ne voyons vraiment pas souvent et un océan nous sépare, mais vous avez su me reconforter même à distance.

Enfin, un dernier mot pour une personne qui était là pour moi tout au long de cette thèse. Merci infiniment chéri, d'avoir cru en moi, de m'avoir soutenu et supporté mes absences, mes caprices, mes folies. J'espère que tu auras encore de la patience de rester à mes côtés pour nos futurs échanges à refaire le monde toute la nuit.

## Summary

Remerciements.....	3
Introduction. News from the Anthropocene.....	9
1. Biodiversity crisis of Anthropocene.....	9
1.1 Biodiversity Definition.....	11
1.2 Components of Biodiversity.....	13
1.3. Scales of Biodiversity.....	14
2. Humanity and Biodiversity relations.....	14
2.1. Biodiversity and ecosystem functioning.....	15
2.2. Biodiversity contributions to humanity.....	19
2.3. What threats to biodiversity?.....	20
2.4. Value of biodiversity.....	21
2.5. Tools for biodiversity conservation: from global problems to local solutions.....	22
3. Is Biodiversity measurable?.....	24
3.1. What unit for biodiversity measure?.....	25
3.2. Measurement of Biodiversity: how to proceed?.....	27
4. Objectives of the thesis.....	33
Chapter 1.....	35
Reconciling the concepts and measures of diversity, rarity and originality in ecology and evolution.....	35
Chapter 2.....	69
Diversity in the city: urbanization affects differently trait-based and phylogenetic plant originality at local and regional scales.....	69
Introduction to chapter 2.....	70
Abstract.....	73
Introduction.....	74
Materials & Methods.....	77
Study Area.....	77
Biological Data.....	78
Land Cover Data.....	79
Species Biological Characteristics and Phylogeny.....	79
Data Analyses.....	81
Community and Species Metrics.....	81
Statistical and Spatial Modeling.....	83
Results.....	84
Community Level Metrics.....	84
Species Level Metrics.....	85
Dataset without Non-native Species.....	87
Discussion.....	87
Trait-based Originality of Urban Communities.....	88
Phylogenetic Originality of Urban Communities.....	89
Species Richness and Non-native Species Originality.....	90
Conclusion and Directions for Further Research.....	91
Conflict of Interest Statement.....	93
Author Contributions.....	93
Funding.....	93
Acknowledgements.....	93
Data Availability.....	94
References.....	95
Supplementary results and commentaries to chapter 2.....	155

Variation of local species-specific originality .....	155
Comparing data issues and methodological choices with study of Veron et al. ....	157
General discussion .....	163
All is a matter of choice and availability .....	163
1. Choosing and obtaining the data for originality measures .....	166
1.1 On the benefits and limits of using different aspects of species' biology .....	166
1.2 Issues related to the availability and to the quality of the data.....	177
2. The matter of scales for originality measures .....	185
2.1 Importance of the spatial scale .....	186
2.2 Taxonomical scale.....	188
3. Choosing the mathematical formula to measure originality.....	188
3.1 Relations between species originality measures.....	189
3.2 On the ubiquity of originality measures in general.....	192
Conclusion.....	195
References of the thesis .....	199
Afterword .....	219
List of publications .....	220
List of communications .....	220
Appendix A .....	221
On islands, evolutionary but not functional originality is rare.....	221
Abstract .....	223
Introduction.....	224
Methods.....	227
Occurrence data .....	227
Estimating evolutionary and functional originality .....	228
Are species more original on continents or on islands? .....	229
Relationship between evolutionary and functional originality.....	229
Are insular original species geographically rare?.....	230
Dispersal modes, biogeographic characteristic of islands and species originality .....	230
Results.....	232
Functional and evolutionary originality are decoupled.....	232
Species endemic to islands are more original.....	233
Evolutionary but not functionally original species are unusually rare.....	234
Dispersal abilities of original species.....	235
Possibility of islands to be at reach and distribution of originality .....	236
Discussion.....	238
Decoupling of functional and evolutionary originality.....	238
Higher originality on islands than on continents.....	239
Range-restriction of original species .....	240
Dispersal capacities of evolutionary and functionally original species.....	241
Conservation.....	242
References.....	245
Appendix B .....	273
Vigie-flore protocol.....	273
Appendix C .....	281
Teaching Theory of Evolution at primary school.....	281

*Empty spaces, what are we living for?  
Abandoned places, I guess we know the score  
On and on, does anybody know what we are looking for?*

**The show must go on**  
Queen, from album "Innuendo", 1991

## INTRODUCTION. NEWS FROM THE ANTHROPOCENE

---

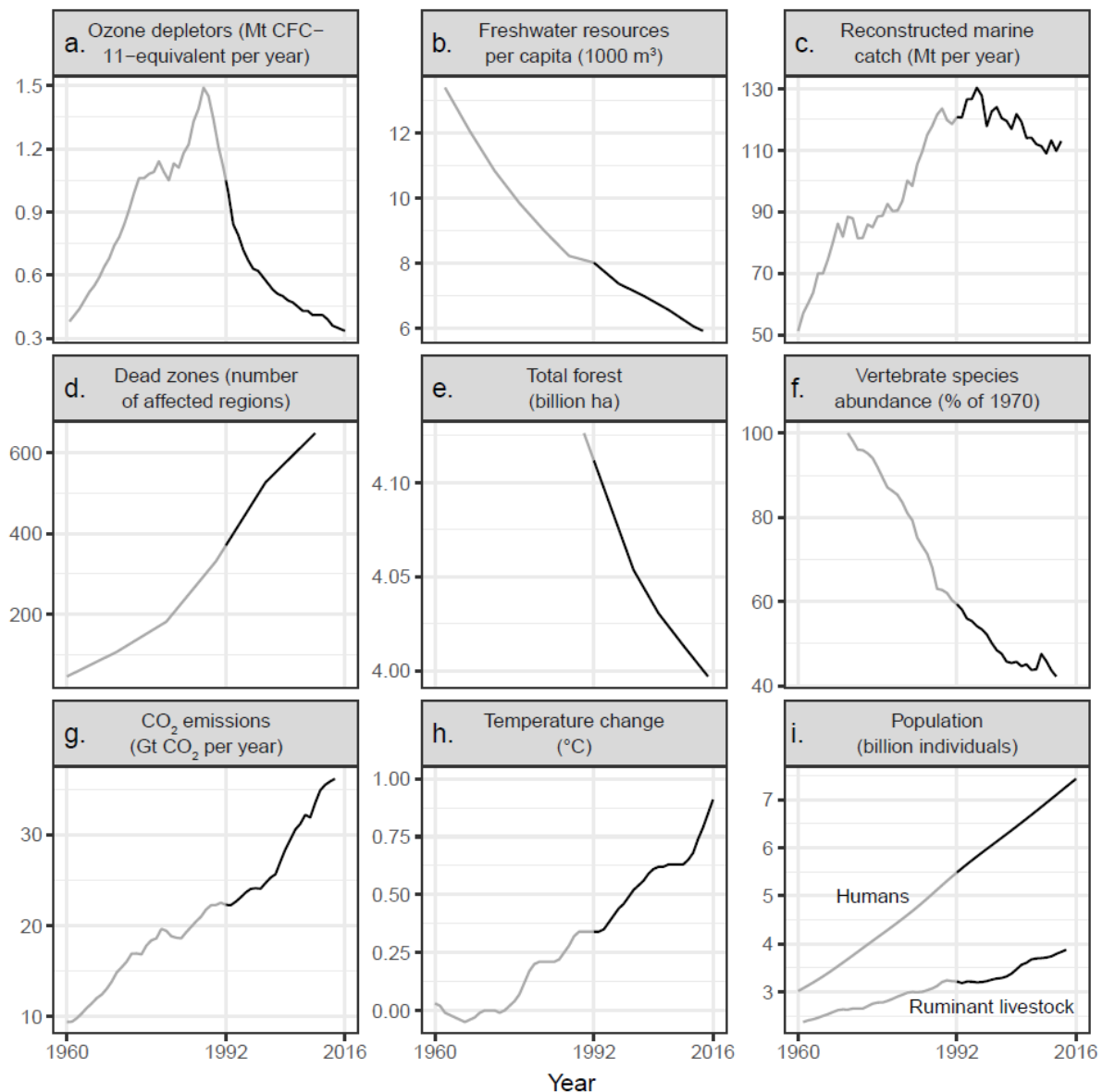
*... it takes a peculiar kind of genius and courage of spirit to be able to envisage nature in the innumerable multitude of its productions without losing one's orientation, and to believe oneself capable of understanding and comparing such productions...*

Buffon, *First discourse*, « Histoire Naturelle » (1749)

### 1. BIODIVERSITY CRISIS OF ANTHROPOCENE

We live in the end of the bottom of material abundancy. In 1992, the Union of Concerned Scientists and many other researchers worldwide edited the first “World Scientists’ Warning to Humanity”. What were those scientists worried about 27 years ago? Essentially, they feared that humanity was about to approach the limits of what natural ecosystems could tolerate from current human practices. Today, in 2019, we can confirm that they did not miss the point. In early 2000’s Paul Crutzen suggested that the Earth had left the Holocene and entered a new epoch, the Anthropocene, demarcated by a footprint of human activities (Crutzen, 2000, 2006). Crutzen did not invent the concept, but he proposed to define the start date of the Anthropocene near the end of the XVIII century, about the time that the industrial revolution began (the invention of the steam engine by James Watt in 1784). Other scientists propose to mark the start of Anthropocene with the agricultural revolution, 15 000 years ago, or else with the Columbian Exchange of Old World and New World species (Waters et al., 2016). Although its beginning date is still under debate, the Anthropocene epoch has certainly established in XXI century.

The ‘Great Acceleration’ graphs usually illustrate the Anthropocene trends of an increasing human imprint on the planet Earth (Figure 1): global rises in temperatures, increasing pollutions, defaunation and intensification of land use. Even though human imprint on Earth continue to rise, there were some signs of improvement in the last decade. For example, the Antarctic ozone hole, extreme poverty, hunger and fertility issues and deforestation of some parts of the Earth have been reduced (Ripple et al., 2017). In the last decades human activity pronounced biotic changes in species assemblages worldwide and accelerated species extinction rate, associated mostly with



**Figure 1.** "Trends over time for environmental issues identified in the 1992 scientists' warning to humanity. The years before and after the 1992 scientists' warning are shown as gray and black lines, respectively. Panel (a) shows emissions of halogen source gases, which deplete stratospheric ozone, assuming a constant natural emission rate of 0.11 Mt CFC- 11-equivalent per year. In panel (c), marine catch has been going down since the mid-1990s, but at the same time, fishing effort has been going up (supplemental file S1 of Ripple et al., 2017). The vertebrate abundance index in panel (f) has been adjusted for taxonomic and geographic bias but incorporates relatively little data from developing countries, where there are the fewest studies; between 1970 and 2012, vertebrates declined by 58 percent, with freshwater, marine, and terrestrial populations declining by 81, 36, and 35 percent, respectively (file S1). Five-year means are shown in panel (h). In panel (i), ruminant livestock consist of domestic cattle, sheep, goats, and buffaloes. Note that y-axes do not start at zero, and it is important to inspect the data range when interpreting each graph. Percentage change, since 1992, for the variables in each panel are as follows: (a) -68.1%; (b) -26.1%; (c) -6.4%; (d) +75.3%; (e) -2.8%; (f) -28.9%; (g) +62.1%; (h) +167.6%; and (i) humans: +35.5%, ruminant livestock: +20.5%. Additional descriptions of the variables and trends, as well as sources for figure 1, are included in file S1 [of Ripple et al., 2017]."

Figure and legend are extracted from "Warning to Humanity: A second notice" (Ripple et al., 2017).

land-use intensification (Pimm et al., 2014; Williams et al., 2015). Specific to Anthropocene, a new branch of ecology developed since several decades, called disturbance ecology (Newman et al., 2016). Disturbance ecology makes historical comparisons of past, current and future disturbance regimes and allows the development of adapted conservation measures.

In 2017, a “Warning to Humanity: A second notice” has been written and signed by over 15 000 researchers all over the world (Ripple et al., 2017). If you are one of those who signed this sort of a scientific petition, you might agree that soon it could be too late to “shift course away from our failing trajectory”. Humanity would certainly need natural resources in the future, and would continue to exploit them. Thus, our existence have to be thought sustainably. However, there are cultural differences between societies (development) and therefore there are differences in their relation to the Nature. There are as many natures as there are cultures. In economically wealthy countries, the relationship of humans to the Nature (and not the reverse) could be described as a relation of a subject to an object (Descola, 2005). In such a society we can imagine a kind of a moral border created between the Humanity with its cultural identities and all the rest assembled under the term of Nature. Seen under human interests perspective, Nature loses its role of a subject and becomes an object of admiration, of exploitation and finally of concern. However, these moral values are not universal and alternatives exist. For example, in Amazonian indigenous cultures, there is no conception of Nature at all (Elands et al., 2015), thus there is no separation of human from Nature (Diaz et al., 2015). However, in a world of scientific research, Nature, denoted as biodiversity, is above all an objects of study. Therefore, I will stick to the scientific definition of biodiversity, which is still often underappreciated in socio-economic and political discourses on sustainable use and conservation of biodiversity.

### **1.1 Biodiversity Definition**

Despite being a common object of study of many researches, biodiversity is a complex concept and there is no univocal way to define it. DeLong (1996) in his review listed 85 definitions of biodiversity and still was not able to define the semantic basis of it. A whole book has been edited



lately, all about defining, measuring and conserving biodiversity (Casetta, da Silva and Vecchi, 2019), that is to say, the concept of biological diversity is large and includes many revealing components. The invention of the term biodiversity was motivated by efforts to stop species extinctions and habitat destruction (Mayer, 2006). It comes as a modern contraction of the term “biological diversity”, derived from Greek “bios” which stands for “life”. Since its appearance in mid-1980’s (Lovejoy, 1980, Haila and Kouki, 1994) major changes happened in the definitions and use of the term biodiversity, referred to the variety and variability among living organisms. Today, the most widely cited definition comes from the Convention on Biological diversity at the “Earth Summit”, held in Brazil in 1992 (UNEP, The United Nations Convention on Biological Diversity) and which says that biodiversity is:

*“the variability among living organisms from all sources including, inter alia, terrestrial, marine, and other aquatic ecosystems as well as the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems”.*

DeLong (1996) offered a more comprehensive definition:

*“Biodiversity is an attribute of an area and specifically refers to the variety within and among living organisms, assemblages of living organisms, biotic communities, and biotic processes, whether naturally occurring or modified by humans. Biodiversity can be measured in terms of genetic diversity and the identity and number of different types of species, assemblages of species, biotic communities, and biotic processes, and the amount (e.g., abundance, biomass, cover, or rate) and structure of each. It can be observed and measured at any spatial scale ranging from microsites and habitat patches to the entire biosphere”.*

The definition of DeLong allows for modification according to the context in which it is used, but stays vague about the diversity of abiotic components and ecological processes, mentioned as diversity of ecosystems. It is astonishing that in XXI century, after all efforts of conservation and management made for biodiversity, we are still trying to define a universally agreed definition of it, which probably does not exist (Swingland, 2013). Perceived differently among scientists, policy-makers or conservationists, a biodiversity definition has to go with pragmatic objectives targeting

specific biodiversity issues (Casetta, da Silva and Vecchi, 2019).

## 1.2 Components of Biodiversity

The definition of biodiversity is organized by three intimately linked main components, proposed by Gaston and Spicer (2004):

- genetic diversity, referring to the heritable variation of genetic information within and between species and populations,
- species diversity (or organismal diversity), referring to inter-species diversity or more largely to taxonomic diversity from individuals up towards to species, genera and so on,
- ecosystem diversity, referring to the diversity in species' communities at ecosystem level.

The molecular diversity and genetic richness of life (Campbell, 2003) should be considered as the basis of biodiversity evolution. Besides, Boon (2019) proposed to investigate intra-individual genetic heterogeneity, accessible today with accumulated data on genetic information contained in DNA sequences. Concerning species diversity, there are three main aspects: taxonomic, trait-based and phylogenetic. Taxonomic diversity is usually presented as a number of different taxa (Humphries et al., 1995). It is the simplest way to describe biodiversity at a given location. Trait-based diversity represents the richness of character types of species or else their biological traits (Petchey and Gaston, 2002). I will define species traits in the section 3.2.1. Finally, phylodiversity represents the evolutionary potential of a given set of species, hidden behind the branches of a phylogenetic tree.

Biodiversity studies usually integrate the up-mentioned “holy trinity” of components (genes, species and ecosystems). This entity-based approach in conservation is particularly used in species-centered studies, trying to describe and to protect each of them. However, species are not isolated from other ecosystem components, biotic and abiotic, thus, entity-approach should be integrated to the process-based approach (Vecchi and Mills, 2019). Indeed, species are at the base of many ecosystem processes insuring efficient ecosystem functioning, and linking species to the underlying ecosystem processes can strongly enhance conservation action efficacy. Such

investigations can be made at several scales of biodiversity.

### **1.3. Scales of Biodiversity**

Living organisms are organized following a nested hierarchy of units with increasing complexity. We can distinguish two levels of organization: biological systems, from genome to species population sharing genetic homogeneity, and ecological systems, which include different interacting species organized into communities at different spatial and temporal levels (André et al., 2003). The level of ecological organization is in the focus of this thesis.

The diversity of ecological systems can be approached at a broad range of spatial levels. Usually researchers separate it on three levels: 1) alpha diversity, which corresponds to intra-habitat diversity between species composing one community, 2) beta diversity of species composition between communities and 3) gamma diversity of a global landscape in a region (Whittaker, 1972; Magurran, 2004). Those three levels of biodiversity are interdependent as gamma diversity is a function of alpha and beta diversities. Moreover, with changing seasons and habitats, biodiversity vary not only in space (Escuedro et al., 2003) but also in time (Pachepsky et al., 2001). Actual biodiversity crisis is a global phenomenon. For example, current practices of species extraction resulted in quasi total disappearance of some plant species (Prance, 2002). However, conservation efforts are usually made at the local scale. Moreover, global species extinctions have been rarely observed, as human impact usually leads to local extinctions and biodiversity changes (Hooper et al., 2005). To be sustainable, we thus need to consider social, political and economic influences that shape biodiversity management at local scale (Williams, 2004). This requires more integrative research on the variation of biodiversity across spatial and temporal scales, guided by local priorities of conservation.

## **2. HUMANITY AND BIODIVERSITY RELATIONS**

There is an evidence for global changes across the world that causes planet's biotic impoverishment (Reid and Miller, 1989). Nature didn't ask to be protected and it causes

contradictions in conservation priorities and “ecological” objectives (Papaux and Bourg, 2015). However, humanity does have big reasons to take care about biodiversity. According to the Food and Agriculture Organization of United Nations, 40% of the world's economy is based directly and indirectly on the use of biological resources that emanate from biodiverse systems. Taking care of biodiversity turns out to taking care of the well-being of our-selves. Unfortunately, humanity's growing needs surpasses biodiversity capacities to satisfy them and if we want to continue benefiting from natural resources, the question of biodiversity use has to be taken seriously. It can be accomplished by the sustainable use of biodiversity, which implies the consideration of eventual future generations' needs. The sustainable use of biodiversity is one of the three main objectives set by the Convention on Biological Diversity (CBD, 1992<sup>1</sup>). In 2011, in Nagoya, Japan, a Strategic Plan for Biodiversity of CBD has been updated with elaboration of 20 Aichi Biodiversity Targets<sup>2</sup>. *“The path to solving this dilemma involves changing patterns of behavior and social norms that influence them; but such norms also typically change only on those longer time scales. Sustainability in the new millennium will depend upon our ability to affect with sufficient dispatch the cultural norms and legal instruments that govern individual behaviors in the global commons”* (Levin, 2000).

That is, as current global changes overwhelms the speed of adaptation of biological systems, we probably should accept current environmental conditions as a new normal and adjust conservation actions to the future trajectories of anthropogenic disturbances (Newman, 2019) that will certainly continue. But first of all, we need to understand how natural ecosystems function and how they depend on biodiversity.

## **2.1. Biodiversity and ecosystem functioning**

Major advances have been made in describing the relationship between biodiversity and ecosystem functioning – a broad concept including a variety of phenomena happening in ecosystems and including ecosystem proprieties, goods and services (Hooper et al., 2005). Biodiversity is a result

---

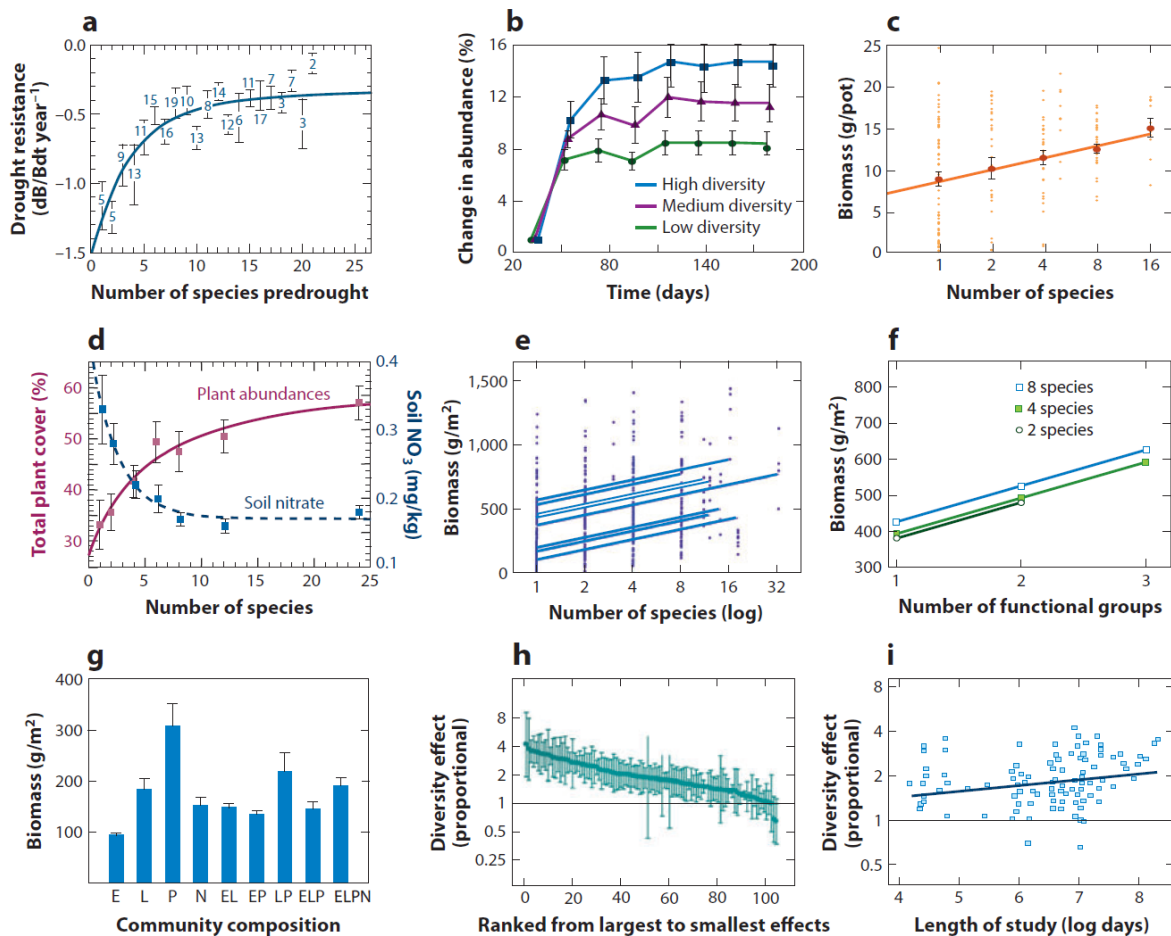
<sup>1</sup> <https://www.cbd.int/doc/legal/cbd-en.pdf>

<sup>2</sup> <https://www.cbd.int/sp/targets/>

of dynamical interactions between its components in space and time, which obey partly different sets of constraints, and hence partly different laws (Loreau, 2010). When studying effects of biodiversity changes on ecosystem functioning one should define which component of biodiversity is in focus. Ecosystem ecologists have often studied species and their biological characters, as a key element for understanding the link between biodiversity and its role in shaping ecosystems (Lavorel et al, 2011).

Since 1960's, ecology studies combining theoretic, field and experimental methods succeeded to gradually develop the debate on causes and consequences of biodiversity loss (Hooper et al, 2005; Tilman et al., 2014). Those studies were investigating three fundamental ecosystem processes: primary productivity, stability (and resilience) and invasibility. Some examples are presented in Figure 2 extracted from Tilman et al. (2014). Concerning productivity-diversity relationship, the major result showed that a total biomass of ecosystem linearly increase with a number of species in ecosystem (see Tilman et al., 2014). Living organisms, through their metabolism and everyday life activities generate an adaptive feedback with their abiotic environment, thus regulating abiotic conditions (Kylafis & Loreau 2008). Further, many natural processes that are happening spontaneously in ecosystem rely on direct interactions between species such as predation, competition, mutualism, facilitation or disease (Cardinale et al., 2002) influencing diversity-stability relationship. Ecosystem stability is a multifaceted concept with many definitions but which generally determines ecosystem capacity to return to an equilibrium state in terms of species densities and interactions after a perturbation (Holling, 1973; Hooper et al., 2005; Ives and Carpenter, 2007). Together stability and diversity determine ecosystem functioning fluctuations. A variety of observational studies confirmed the theory that greater species diversity leads to greater ecosystem stability, especially if species are redundant in their functions (Lehman and Tilman, 2000, Gross et al, 2015). Finally, diversity-invasibility relationship was found positive. For example, invader success of exotic species decreased with higher native species diversity (Naeem et al., 2000; Loreau 2010).

Scientific community came up with a broad consensus that many species have generally a positive impact on ecosystem functioning (Hooper et al., 2005, Ives and Carpenter, 2007). However, the majority of studies that investigated links between biodiversity and ecosystem functioning were carried out in terrestrial plant communities, which are also the primary producer at the base of every trophic chain. Moreover, contradictory effects of plant diversity on higher trophic levels were found (Mittelbach et al., 2001; Proulx et al., 2010). Thus, it is unclear whether the results found with terrestrial plant species can be generalized to other taxa, different trophic levels and different habitat types (aquatic, subterranean). It has been shown that, depending on the biodiversity components studied, on environmental constraints and on spatial scale of study diverse biodiversity patterns can be observed (Mittelbach et al., 2001). Indeed, different components of species diversity follow different dynamics (Brum et al., 2017) and may have contrasting effects on multiple ecosystem processes (Tilman et al., 2014). Trait-based component of biodiversity, for example, strongly influences ecosystem properties (Diaz and Cabido, 2001, Hooper et al., 2005, Loreau, 2010). Le Bagousse-Pinguet et al. (2019) found that taxonomic, phylogenetic and trait-based diversity of plant species taken together were better predictors of multifunctionality of 123 dryland ecosystems than if taken separately. Therefore, biodiversity relations with ecosystem functioning cannot be summarized by a simple diversity metric and have to be addressed at multiple scales for several levels of biological organization (Tilman et al., 2014).



**Figure 2.** "Relationships among species richness, functional composition, and ecosystem function. Increasing species number (a) increases drought resistance; high resistance is associated with a low rate of biomass decline during drought (Tilman & Downing 1994). Increasing species number also (b) increases vegetation cover (Naeem et al. 1994) and (c) increases biomass in a greenhouse experiment (Naeem et al. 1995). (d) Plant abundances (summed percent covers, solid line) increase, whereas soil nitrate (dashed line) decreases (more is taken up by plants) with increasing plant species numbers (Tilman et al. 1996). Aboveground biomass increases with both (e) the number of species and (f) the number of functional groups (Hector et al. 1999). (g) Aboveground biomass depends on plant community composition (Hooper & Vitousek 1997). Abbreviations: E, early-season annuals; L, late-season annuals; P, perennial bunchgrasses; N, nitrogen fixers. (h,i) Diversity effect is plant productivity at high diversity divided by that of monocultures for ~100 experiments, with diversity effects (h) ranked from largest to smallest effects and (i) graphed against length of study (ln of number of days, where "7" = 3 years) (Cardinale et al. 2007). Error bars indicate 1 SEM (standard error of the mean)".

Figure and legend are extracted from Tilman et al. (2014)

## 2.2. Biodiversity contributions to humanity

We saw that biodiversity is involved into many natural ecosystem processes and that understanding of the biodiversity-ecosystem functioning relationships would help to prioritize relevant biodiversity components in conservation management actions (Diaz et al., 2006). Humanity's reliance upon biodiversity for its welfare and survival is enormous, and biodiversity loss has unquestionable socio-economic consequences. Ecological process can be seen through the utilitarian prism making biodiversity of a vital importance to humanity. By influencing their environment different species are involved into primary production, decomposition or biogeochemical cycles of nutrients, water and other elements, soil generation, pest control and climate regulation. A plenitude of ecosystem processes were assembled into subgroups according to the nature of benefits they can offer and called "ecosystem services". The Millennium Ecosystem Assessment (MEA, 2005), that gathered around 1700 researchers all over the world, listed four roles of biodiversity in ecosystem services supply:

1. provisioning role of biodiversity by supplying materials, water, food etc.;
2. regulatory role through the influence of biodiversity on production, stability and resilience of ecosystems;
3. supporting role through structural, compositional and trait-based (functional) diversity
4. cultural role through the non-material benefits (spiritual, educational).

Going further than traditionally described ecosystem services, biodiversity provides an inspiration for new energy sources, materials and technologies, commonly called biomimeticism. In practice, however, it is quasi impossible to clearly differentiate interrelated ecosystem services.

MEA founded their framework on two pillars: ecological and economic, referring to the multiple ecological functions that serves humanity. In the meanwhile, all societies value nature in very different ways for their own sake (Wilson, 1984) and MEA recommendations failed to be



operational locally, as the definition of ecosystem services lacks socio-cultural values along with ecological and economic ones. This intrinsic value of nature is of primary importance in some societies, especially for indigenous people. Therefore, Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES, Diaz et al., in press) scientists proposed more inclusive, yet not approved to this day, notion of “nature’s contributions to people” instead of MEA’s ecosystem services framework (Díaz et al., 2018), embracing the complex relationship between humans and nature.

### **2.3. What threats to biodiversity?**

As described above, our planet’s living mantle – biosphere – suffers from human intervention, unprecedented in Earth’s history (Tittensor et al., 2014). Indeed, global changes accelerated by human activity impose selective pressures on the organisms and therefore influence adaptation and evolutionary trajectories of taxa (Williams et al., 2015). The most recent classification of direct causes of biodiversity decline was made by the IPBES (Diaz et al., in press). The list consist of five causes (with a decreasing level of impact): changes in land and sea use; direct exploitation of organisms; climate change; pollutions; and invasion of non-native species. There are more precisely agricultural overexploitation, urbanization and industrial development threatening wild populations by habitat fragmentation, loss and degradation (Krauss et al., 2010); invasive non-native species taking over native populations in resource use (Walther et al., 2009); introduced pathogens altering biochemical processes in the water, air and soil; climate change modifying species habitats and also their behavior, etc. (Diaz et al., 2015; Williams et al., 2015). Those five direct drivers of biodiversity change are underpinned by societal values and behaviors that include production and consumption patterns, human population dynamics and trends, trades, technological innovations and local to global governance. However, the rate of change in the direct and indirect drivers of biodiversity loss differs among regions and countries. Moreover, new types of anthropogenic disturbances will interact with natural disturbance regimes resulting in completely different threats, proper to Anthropocene (Newman et al., 2019). In order to minimize its negative effect on biodiversity,

humanity established red lists of endangered species, created protected natural areas and is trying to develop a sustainable way of life. Would that be enough?

One of the most obvious and gravest consequences of biodiversity crisis are species extinctions. Species extinction is a natural process occurring without human intervention on large time scales, and all living organisms have its limited geological time of existence (Soulé, 1985). However, extinction rates caused by humans exceeds natural pace of things. The International Union for the Conservation of Nature (IUCN) estimated that out of 1.7 million of described species around 25% are threatened and about 1 million are under threat of extinction (Mittermeier et al., 2011). As species vanish, so does the benefits they offer to humanity. Therefore, because species extinctions are irreversible, their irreplaceable value in terms of services should be a mandatory part of ecosystem assessments.

#### **2.4. Value of biodiversity**

Based on the ecosystem service notion, many scientists and economists tried to quantify the real value of biodiversity loss (Costanza et al. 1997; Balmford et al. 2002). However, it is difficult to establish a meaningful and relevant value of biodiversity. There is no single valuation process of biodiversity and three main perspectives - economic, societal and ecological - should be accounted for (Laurila-Pant et al., 2015; Sukhdev et al., 2014). The monetary valuation of biodiversity changes can still serve as a useful link between environmental problems and political decision-making processes (Braüer, 2003) but there is no univocal, unambiguous monetary indicator (Nunes and van der Bergh, 2001). How much would cost the restoration measures compared to the loss of economic benefits? What does it really costs to replace services provided for free by nature? This economical valuation has a drawback of not accounting for social visions of nature. Therefore, socio-cultural way of biodiversity valuation may be more adapted to the cultural and spiritual relations to the nature (Sukhdev et al., 2014). Unfortunately, many ecosystem services cannot be bought or sold, thus they have been called "non-use values" of biodiversity (Pagiola et al., 2004).

Furthermore, each component of biodiversity has its unique, intrinsic value or “the value of nature regardless of the services and economic benefits it provides” (Pearson, 2016). Some researchers proposed the idea of biodiversity option value, which represents the unanticipated capacity of biodiversity to provide potential future benefits (Faith, 1992, MEA, 2005). Thus, it is not only about focusing on ecosystem current uses, but also about making provisions for future generations.

Today, socio-ecological processes are increasingly shaping the world’s natural ecosystems. Humanity exploits a minor part of the world’s biological diversity (Swingland, 2013) and undiscovered option and intrinsic non-use values should support the precautionary principle of biodiversity maintaining. In order to conserve biodiversity and prevent its further loss, there is a need to complement the economic and socio-cultural valuation approaches of biodiversity with another one based on precise scientific knowledge of ecosystem functioning (Laurila-Pant et al., 2015). This signifies determining all possible threats, direct or indirect, hanging over biodiversity and measuring its well-being.

## **2.5. Tools for biodiversity conservation: from global problems to local solutions**

Prevention of biodiversity loss is becoming of the paramount importance for all nations, societies and individuals. Biodiversity must be urgently assessed in an operational way in order to facilitate conservation targets setting. Since the United Nations Convention on Biodiversity, multiple intergovernmental assessments, conservation programs and agencies flourished in order to understand and reduce biodiversity loss (e.g., European Environment Agency, IUCN, etc). Many of them adopted IPBES as their science-policy interface (Pimm et al., 2014). One of the simplest ways to assure the protection of individual species and entire ecosystems consists in setting aside entire areas, such as natural reserves and parks (Margules and Pressey, 2000), just by isolating a part of biodiversity from any possible threat (“nature for itself” and “nature despite people” policy of conservation; Mace, 2014). Such approaches attempt to preserve native species and their natural habitats they evolved in, excluding any area that has been in contact with human activity (the

concept of wilderness, Mittermeier et al., 2003). However, there are areas around protected spaces, where damage to biodiversity could be so big that species would not be able to re-colonize some parts of their native habitat without human intervention. Conservation translocation and *ex situ* conservation, are another tool for biodiversity conservation (Sarrazin and Barbault, 1996; Taylor et al., 2017). Conservation translocation is an intentional human-mediated movement and release of species in order to re-establish viable populations within species distribution range (Seddon, 2010). *Ex situ* conservation tries to maintain a viable population in cultivation or captivity (in seed banks, embryos, sperm and oocyte storages). Such conservation tools often involves a loss of genetic diversity through inbreeding and founder effect (Milner-Gulland and Mace, 1998). However, species do not influence ecosystem processes equally (Tilman, 1997), nor do they have the same evolutionary history and potential. Thus, no single taxon is a good proxy for biodiversity protection (MEA, 2005).

Poiani et al. (2000) proposed a practical framework of biodiversity conservation intended to make a transition from single-species conservation of rare and endangered species to ecosystem-oriented conservation, taking into account four ecosystem attributes: composition and structure of the focal ecosystems; dominant environmental regimes, including natural disturbance; minimum dynamic area; and connectivity. They recommend to use this framework several times during conservation actions for more effectiveness. Such ecological management aims to conserve and restore historical ecosystems and ensure maintenance of ecosystem services and goods. Therefore, biodiversity protection often requires human assistance. Moreover, once biodiversity passed the tipping point of no return because of human intervention, the “nature and people” policy should be developed linking nature to human well-being (Mace, 2014). Thus, as a final step in human relation to biodiversity, the idea of wilderness should be abandoned and framework of biodiversity conservation should concentrate on socio-natural areas, integrating both, natural biodiversity dimension and socio-economic human dimension as partners (Bourg and Papaux, 2015). Such policy was recently adapted in ecology of landscapes and urban ecology, making accent on ordinary

biodiversity (nature that can be experienced in every-day life).

For sure, no single methodology is appropriate for multiple biodiversity loss issues. Effects of global change differ between locations, spatial scales and time periods. It is not possible to protect all biodiversity components simultaneously and some of them have to be prioritized for conservation. Hence, conservation efforts are usually made at the local scale (Pimm et al., 2001). First, because local species extinctions are more common (e.g., large changes in species abundancies) than global ones (Hooper et al., 2005). Second, because people care more about their close environment and because localized actions are easier to implement. Developing biodiversity conservation demands a deep knowledge of its state in a first place. Therefore, biodiversity monitoring has to be aligned with local and national priorities of biodiversity conservation (see Durant, 2014). Organizations such as IUCN, play a key role in facilitating effective communication between actors of biodiversity management, creating tools useful to all sides and ensuring that information provided is credible, and gathered via a legitimate process. It is important that current conservation efforts follow relevant scientific progress and that policy-makers use up-to-date methods to ensure better future for nature and people (Mace, 2014; Newman, 2019). One efficient way to evaluate biodiversity state is to quantify it, in other words to give biodiversity a numerical value (do not confuse with monetary value). To move from assessing to conserving biodiversity, many biodiversity measures were developed (Purvis and Hector, 2000).

### **3. IS BIODIVERSITY MEASURABLE?**

Measuring such a complex thing as biological diversity, where each level of biological organization varies in composition, structure and function (Noss, 1990; Mayer, 2006), is a hard but achievable task. In spite of such organizational complexity of biodiversity, one have to choose which biodiversity component to measure. Hence, biodiversity conservation has to focus on the dynamic ecological processes at different spatial and temporal scales (Levin, 2000; Poiani et al., 2000). Nevertheless, components of biodiversity are interdependent and accessing biodiversity gradually through its hierarchical organization complements a view of its state. Local biodiversity variations

may lead to biodiversity generation and maintenance on larger spatial scales and longer time periods (Levin, 2000). Therefore, multiscale conservation effort offers a more comprehensive way for protecting a greater amount of biodiversity (Poiani et al., 2000). In ecological sciences a vast range of biodiversity metrics exist, each approaching one of the facets of biodiversity. The choice for the most adapted biodiversity metric will depend on the underlying ecological question, spatial scale of the study and nature of the ecological data (Magurran, 2004, Mayer, 2006). In the Chapter 1 we will compare several biodiversity measures and discuss their possible uses in ecology and conservation biology.

### 3.1. What unit for biodiversity measure?

*No one definition has as yet satisfied all naturalists; yet every naturalist knows vaguely what he means when he speaks of a species.*  
Darwin (1859/1964)

Generally, biodiversity measures are made in a way to evaluate a state of a biological entity. Such entities can be categorized in sets by what they have in common. In ecology, those entities can be represented by a chosen level of biological organization, from genes to landscapes passing by species and individuals. In my thesis and more generally in community ecology, an entity would be composed of individual organisms from the same species and a community would be a set of these entities. In conservational biology, species became a common currency for studying biodiversity variance and from species distribution and abundance (Agapow et al., 2004). For example, species are used to define biodiversity hotspots (Myers et al., 2000) selected upon IUCN species' threat classification. Public concern plays a great role in budgetary allocations for conservation programs, and easily recognized, charismatic species are highly valued, regardless of their real ecological significance (Agapow et al., 2004; Pearson, 2016; Albert et al., 2018). Would you prefer to participate into conservation of a puma or of a dangerous insect? And besides, what hides behind the term of "species" and how is it defined?

John Ray (1627-1705), an English naturalist, was the first who introduced the term of "animal

species". Then, Carl Von Linné introduced binominal nomenclature (genera + species), and biologists adopted this naming system for any living organism. For example, *Canis lupus* stands for the wolf, where "Canis" is a genus and "lupus" is a species belonging to that genus with other Canidae species. Thus, species is the smallest and the most commonly used unit of taxonomical hierarchy starting from a kingdom (which would be animal kingdom for *C. lupus*). Even though sub-species may be named sometimes, they often remain controversial (Wilson and Brown, 1953). However, there is no universal way to define species, and no other subject in biology received so much controversy as the species concept. The extensive literature on species notion shows controversial definitions of species concept which are still debated by scientists. Over 20 definitions were described, and two integrated frameworks of species concepts have been proposed by Mayden (1999) and by de Queiroz (2005b, 2007). Unfortunately, the problem of species concept definition is often confused with issues of how species is delimited and compartmentalized as a meaningful entity in biology (Andersson, 1990; Sites and Marshall, 2004; Naomi, 2011). Globally, there are four partly overlapping definitions of macroorganism species concept:

- the biological (or reproductive) species concept defining species as populations that cannot interbreed successfully (Mayr, 1942, 1963);
- the evolutionary species concept defining species as lineages of ancestral descendant populations evolving separately with its own roles and tendencies (Simpson, 1961; Wiley, 1978);
- the ecological species concept defining species as lineages evolving in minimally different adaptive zones, depending on species ecological preferences (Van Valen, 1976);
- phylogenetic species concept, delimiting species as a group of organisms that share at least one uniquely derived character, perhaps with a shared pattern of ancestry and descent or monophyly (e.g., Nixon and Wheeler 1990).

Which definition to choose when boundaries of species are confused by horizontal gene transfer, hybridization, and recent evolutionary isolation (Agapow et al., 2004)? Furthermore, very different approaches are used in microbiology to define species concept. Microbial taxa definition is

essentially based on DNA sequencing and hybridization, especially when nothing is known about natural history and ecology of those organisms (see Green and Bohannan, 2006, Box1. and 2.). Being fundamentally different from described above “pragmatic” attempts of species definition, the method of 97% - similarity of ‘operational taxonomic units’, based on ribosomal gene sequences, would join all primates from humans to lemurs in one species (Staley and Gosnik, 1999). Inversely, the phylogenetic species concept would split the biological species concept and increase an extant species number (Agapow et al., 2004). Finally, biological species concept definition cannot be applied to species with asexual reproduction or to hybrids between species. Therefore, the best solution is to see species as an operational entity, clearly defined with sufficient data and accepted in a given area of biology. In this thesis I will use an operational species concept used in ecology, delimited by molecular information and species biological properties as a unit of biodiversity measurement (see Sites and Marshall, 2004).

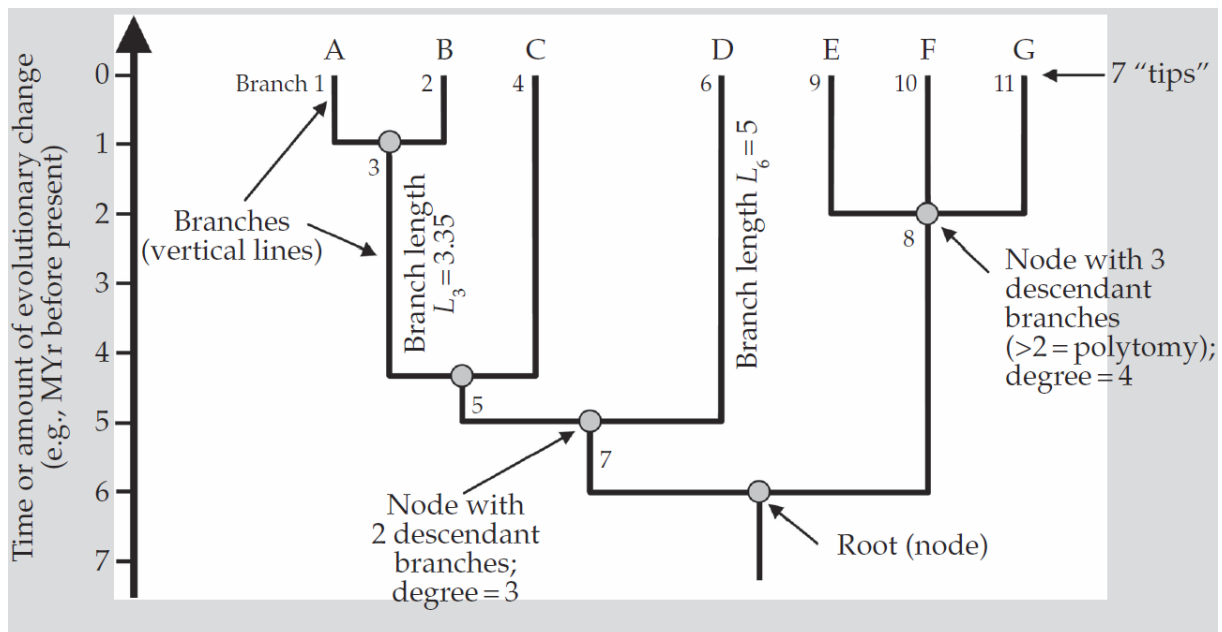
### **3.2. Measurement of Biodiversity: how to proceed?**

#### 3.2.1 Criteria to describe species diversity

The key to quantify species diversity is to consider species differences (dissimilarity) that create that diversity. But which criteria should be used to properly describe species differences? The most frequently used characteristic is species abundance. Species abundance is usually measured as a population size, or a number of individuals in a population, although other means to measure species abundance exist (Rabinowitz, 1982). However, species abundance does not tell us much about species biological characteristics, what gives them the capacity to exploit ecosystem resources, reproduce, interact with their environment and consequently influence ecosystem functioning. Thus, more integrative way to describe differences between species is to use their biological traits. Under the “trait” term have been listed many different types of species’ characteristics, from individually measured features (biochemical, evolutionary, genetic, morphologic, physiologic or behavioral) to environmental conditions of species populations. In order to avoid any confusion and ambiguity Violle et al. (2007) proposed to define traits at the



individual level only. A species trait is a character measurable on every individual of one species and that influence the performance of species (and the ecosystem functioning) and is usually used to compare species between them (McGill et al., 2006; Violle, 2007). As species trait is measured on a sample of species individuals, there can be inter-specific variability. Indeed, species trait values will be different depending on the environment where they were measured (e.g., altitude, latitude, climate or season; e.g., Albert et al., 2012), however mean trait values per species are usually found in global databases. I would like to emphasize the notion of a functional trait, which is often employed to describe many kinds of species traits as a surrogate for species performance in ecosystem allowing them to respond to environmental changes and to influence ecosystem properties. Yet, not all traits are functional, and probably not all of them impact species fitness and survival. As the substantial meaning of biological trait varies among authors and ecological disciplines, from here on I will employ more inclusive term of “biological trait” or just “trait”. Thereby, we can say that every species (or individual) is an assemblage of traits, whose values can be unique features of that specific species (or individual) or commonly shared by many species. Another way to describe species is to use its position on a phylogenetic tree representing the evolutionary relationships between extant species. Phylogenetic tree is the most used principle to organize the classification of organisms. Important progress has been made in building techniques of more reliable and meaningful phylogenetic trees (Roquet et al., 2013). Today, phylogenies are usually made with species molecular data, by aligning multiple genetic sequences in the most parsimonious way (Roquet et al., 2013). Thus, most of the phylogenetic trees used in ecology are molecular, usually built for a specific taxonomic level, and then are assembled into a supertree composed of several molecular phylogenies. Figure 3 represents a structure of a phylogenetic tree (extracted from Vellend et al., 2011).



**Figure 3.** The components of a phylogenetic tree. Tips (or leaves) of a tree represent species, nodes represent the most recent common ancestor of all species descending from it (branches split), the root node (or just a root) represent a single point from which all species descend; branches are characterized by branch length, representing the accumulation of evolutionary history of species changes; polytomies are the nodes with more than 2 descending branches. *Extracted and modified from Vellend et al., 2008, Box 14.1.*

A largely assumed hypothesis says that shared phylogenetic information between species captures on average (and it is important to underline) shared species features, which for the most of them remains, however, unmeasured (Faith, 1992). We will never be able to measure all species biological traits, thus, molecular phylogenies could be a good surrogate for the conservation of both, known (or easily observed) and unknown (hardly observed) characters of organisms. Species unmeasured features are referred as "future options" and represents a potential unanticipated future benefits from features of species in response to future needs (Faith, 1992; Faith, 2017). A recent contradictory debate about phylogenies capturing badly species trait diversity shows that all depends on the traits measured and on a spatial scale considered (see Mazel et al., 2018, 2019 VS Owen et al., 2019). For example, convergent traits, which are independent evolutionary events experienced by species under the same constraints, makes an exception, and convergent trait values are not captured by the phylogenies (Faith and Walker, 1996b). However, an increased use of

phylogenies to describe species in ecological studies reflect the shared recognition of the utility of phylogenetic information for biodiversity representation and conservation.

### 3.2.2 Indices of species diversity – mathematics can save the planet

The eternal problem of “what to protect” could be resolved if conservation efforts were more inclusive using multiple aspects of species biological traits: genetic, morphological, physiological and behavioral. Indeed, species are not equivalent nor exchangeable in their biological characteristics and magnitude of their differences varies across the set of species observed (Vellend et al., 2011). Thus, the choice of appropriate tool to measure biodiversity will depend on the biological discipline, its methods and data availability on species. In general, a biodiversity measure is a mathematical formula, created to give a value to a specific entity. Before being operational, it has to be accepted by scientific community, to be reusable and understood by all actors of biodiversity conservation (Poiani et al., 2000; Balmford, 2005).

We defined that in this thesis the unit of biodiversity measures will be species, studied within the same taxonomic group (e.g., plants, birds, butterflies) at different spatial scales. Measures that determine the amount of biological diversity represented by a set of species are naturally called species diversity measures. Species diversity is the most commonly used measure of biodiversity, mainly because of the ease of working at species level of organization and due to species’ recognized roles in ecosystem functioning (Sullivan and Swingland, 2006). In changing environmental conditions, conservation priorities are set upon specific localities and their specific composition with a common goal to minimize the extinction risk by maintaining the greatest number of species and with them, greater amount of ecosystem services. The simplest diversity measure is species richness, which reflects a number of species in a set (Magurran, 2004) but gives no information on species abundances. To overcome this shortcoming other diversity measures have been developed, such as Shannon index (Shannon, 1948), based on the respective proportions of species abundances or Gini-Simpson index combining species richness and evenness (Gini, 1912; Simpson, 1949). However, these commonly used measures also have several drawbacks, as

they do not allow including species biological differences described above. Thus, in recent decades phylodiversity (Faith, 1992) and trait-based diversity (Petchey and Gaston, 2002) measures have been developed, using accordingly species phylogenetic and trait information. The greater species dispersion on a phylogenetic tree would represent a greater phylodiversity diversity of species. Furthermore, a location with many different higher taxa than species (genus, order, etc.) would be more phylogenetically diverse than a site with higher species richness belonging to the same taxonomical level. For example, two New Zealand's endemic species of Tuatara (genus *Sphenodon*) is the only surviving members of its lineage (*Rhynchocephalia*, Hay et al., 2010), dates to the time of the dinosaurs and looks like lizards (but they are not). These species represents more phylodiversity than some species-rich taxa with a recent common ancestor. Phylodiversity measures, created to capture evolutionary relationships between species were successfully adopted in community ecology to describe community phylogenetic structure (Webb, 2000; 2002; Cadotte et al, 2010). Finally, more recently, trait-based (or functional) diversity has been used to reveal links between species traits and species response to global changes and ecosystem functioning (Petchey and Gaston, 2002).

But what about species response to global change and the amount of diversity this species support? A measure that provides a value for every species in a set would complement a complex picture of biological diversity. In the Chapter 1 we will discuss the most largely used measures for estimating the contribution of each species to biodiversity. Several studies argued that ecological role of species is proportional to their abundance, which corresponds to the “mass ratio” hypothesis of Grime (1998) where the most abundant dominant species represent the most of the biomass and have the most important role in ecosystem functioning (Lavorel et al., 2008; Diaz et al., 2007b; Pakeman et al., 2011). Yet, it is not always true if differences in species' biological characters are considered.

### 3.2.3. Species originality measures based on species' biological characters

Forty years ago, a particular group of measures appeared, in order to characterize each species in a set, and not just the species set (May, 1990; Vane-Wright, 1991; Nixon and Wheeler, 1992). Those measures use inter-species differences to evaluate their specific contribution to the diversity of a finite set of species. A large variety of such indices, even if they measure the same concept, have many different names. In my thesis I refer to them as originality measures, simply because the general principle of all those measures is to evaluate how unusual or original each species is, compared to the others in a set (Pavoine et al., 2005). As for diversity measures, species abundance alone is not always a good predictor of the contribution of species into ecosystem functioning and even relatively rare species can harbor important ecosystem functions and can be even irreplaceable (Mouillot et al., 2013a, Brandl et al., 2016). Thus, species biological characters would allow to measure species originality as a rarity of species character states (Pavoine et al., 2017).

Originality measures can be separated into three large groups, depending on the data they use to give each species a value: species taxonomical relations, species phylogenetic position and species trait values (see Appendix 1 in Chapter 1). From there, to be operational, an originality index needs a species pairwise differences structure, which can be rather a tree structure (for taxonomic and phylogenetic data) or a matrix of differences or dissimilarities between species (usually adopted for trait-based data but could be used for the others too). Originality measures will thus compare species through the shared amount of characteristics (see Figure 2 in General Discussion). Species originality concept can help in our understanding of biodiversity response to global changes which is a step towards better prediction of communities change under future scenarios. I will describe species originality concept and associated methods of its quantification in the Chapter 1.

#### **4. OBJECTIVES OF THE THESIS**

The main objective of this research work is to bring the light on the benefits of species' originality concept and measures to different biological disciplines, particularly to conservation biology and community ecology. In Chapter 1, we start by reviewing the state of the art of three concepts primordial to any biodiversity study: species diversity, species rarity and species originality. We disentangle various terms, approaches and measures associated with each of the three concepts. We also discuss analytical links between measures of diversity, rarity and originality and discuss some examples of their joint use. Above, we described two main types of species biological characters, phylogenetic and trait-based. Thus, in this thesis we consider both of them, as it allows having a more integrated vision of species diversity and originality aspects.

After a vast theoretical review, we demonstrate which new applications of originality measures are possible using a real dataset and ecological context. Indeed, in Chapter 2 we investigate plant species originality variation along an urbanization gradient. Especially, we propose a conceptual framework of species originality measuring at two spatial scales that can help to identify the most relevant spatial scale for trait-based diversity and phylodiversity evaluation in cities and which species contribute the most to such a variation. We also discuss some of methodological and data problems met in Chapter 2 and compare them with another research work (Veron et al., in prep., Appendix A) that proposed several solutions.

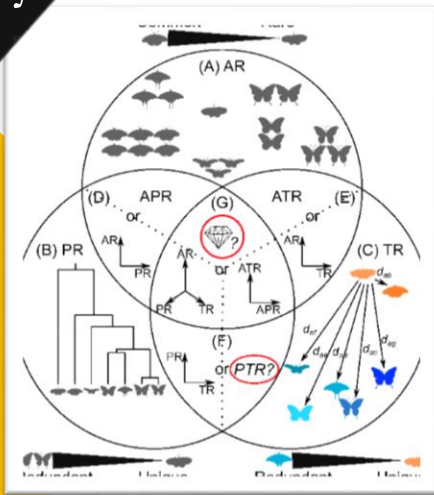
In the end, General Discussion makes an overview of the precedent chapters. We discuss possible drawbacks and advantages of the originality measures use in ecology. Indeed, the data needed for the calculation of originality is not always available but when it does some shortcomings exist.

**INTRODUCTION**

Why and how to protect and measure Biodiversity?

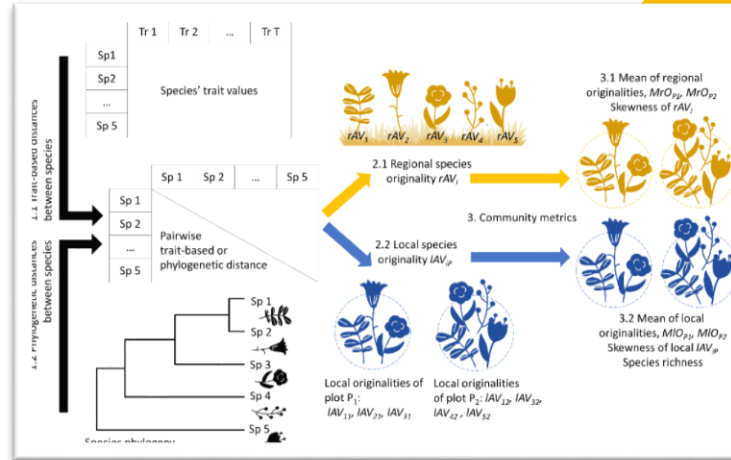
In theory

**Article 1**  
Concepts and measures of species diversity, rarity and originality



In practice

**Article 2**  
Diversity and originality of urban plant species



The matter of choices: data, scale, measures; Drawbacks and advantages

**DISCUSSION**

+Appendix A  
**Article 3**

Figure 4. Mindmap of this thesis

CHAPTER 1.

**RECONCILING THE CONCEPTS AND  
MEASURES OF DIVERSITY, RARITY AND  
ORIGINALITY IN ECOLOGY AND  
EVOLUTION**

---



The concept of biological diversity, or biodiversity, is at the core of many evolutionary and ecological studies. One of the main aims of such studies is to identify the processes that shape biodiversity patterns in space and time. To reach that aim, a myriad of biodiversity measures have been developed especially in the last four decades. Species is one of the central units of these measures. Often species diversity has been quantified not only in terms of species number but also considering species abundance distribution. The abundance of a species is inversely related to an aspect of its rarity, so that traditional diversity measures were shown to be functions of abundance-based rarities. Then, in the last decade, there has been an exponential increase in the number of research studies aimed at determining other aspects of the contribution of each species to diversity. This contribution expresses how original a focal species is compared to all others in an assemblage as regards its position on a phylogenetic tree and the values of its functional traits. Despite their fundamental links, these three concepts of biodiversity, rarity and originality have been mostly independently treated in the past in ecological and evolutionary studies, with only few attempts to show how different they are and how they complement each other. Moreover, there are numerous terms, used currently in the literature, that are synonymous with originality. Those synonyms especially obscure the concept of originality.

In this chapter, presented as a systematic review published in *Biological Reviews* in March 2019, we bring attention to the fact that a conceptual explanation of the links between the three concepts of diversity, rarity and originality is still missing. We thus provide a semantic and historical overview of the definitions of the diversity of an assemblage of species and of the rarity and originality of a species. Furthermore, we clarify some aspects of the complementarities between these concepts and their measurements bringing together ideas developed in evolutionary biology and community and functional ecology.

Regarding how the concepts of diversity, rarity and originality are measured, our aim here was not to review all mathematical formulas developed so far. Instead, we analysed a small set of representative biodiversity measures and demonstrate their links with rarity and originality

measures. We prove, for example, that numerous biodiversity measures can be written as functions of species rarity and originality.

Finally, synthesizing the results of many published studies, we demonstrated why biodiversity, rarity and originality measures should be used together in evolutionary biology, ecology and conservation biology. At a large scale, the joint use of these concepts could help clarify general patterns of evolutionary events such as trait evolution, extinction, speciation, and adaptive radiation of species. At a local scale, it could aid in understanding community assembly and ecosystem functioning. At any scale, it could refine conservation strategies.

# Reconciling the concepts and measures of diversity, rarity and originality in ecology and evolution

Anna Kondratyeva<sup>1,2,\*</sup>, Philippe Grandcolas<sup>2</sup> and Sandrine Pavoine<sup>1</sup>

<sup>1</sup>Centre d'Ecologie et des Sciences de la Conservation (CESCO), Département Homme et Environnement, Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, 57 Rue Cuvier, CP 135, 75005, Paris, France

<sup>2</sup>Institut Systématique Evolution Biodiversité (ISYEB), Département Origines et Evolution, Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université EPHE, 57 Rue Cuvier, CP 50, 75005, Paris, France

## ABSTRACT

The concept of biological diversity, or biodiversity, is at the core of evolutionary and ecological studies. Many indices of biodiversity have been developed in the last four decades, with species being one of the central units of these indices. However, evolutionary and ecological studies need a precise description of species' characteristics to best quantify inter-species diversity, as species are not equivalent and exchangeable. One of the first concepts characterizing species in biodiversity studies was abundance-based rarity. Abundance-based rarity was then complemented by trait- and phylo-based rarity, called species' trait-based and phylogenetic originalities, respectively. Originality, which is a property of an individual species, represents a species' contribution to the overall diversity of a reference set of species. Originality can also be defined as the rarity of a species' characteristics such as the state of a functional trait, which is often assumed to be represented by the position of the species on a phylogenetic tree. We review and compare various approaches for measuring originality, rarity and diversity and demonstrate that (i) even if attempts to bridge these concepts do exist, only a few ecological and evolutionary studies have tried to combine them all in the past two decades; (ii) phylo- and trait-based diversity indices can be written as a function of species rarity and originality measures in several ways; and (iii) there is a need for the joint use of these three types of indices to understand community assembly processes and species' roles in ecosystem functioning in order to protect biodiversity efficiently.

*Key words:* biodiversity measure, community assembly, conservation biology, distinctiveness, extinction risk, functional diversity, originality, phylodiversity, species abundance, trait-based diversity.

## CONTENTS

I. Introduction .....	1318
II. Rarity and Originality: Two Entangled Concepts .....	1319
(1) The concept of species rarity in ecology and evolution .....	1319
(2) The concept of originality or phylo- and trait-based rarity .....	1320
(3) The measurement of originality .....	1321
(4) How should the originality measure be chosen? .....	1323
(5) Spatial scale matters .....	1323
(6) Directions for future research on species' originality .....	1324
III. Links Between Measures of Diversity, Abundance-, Phylo- and Trait-Based Rarity .....	1324
(1) Biodiversity as a mean of abundance-based rarities .....	1325
(2) Explicit link between originality and diversity indices .....	1325
(3) Most diversity indices only implicitly depend on species' originalities .....	1326

\* Author for correspondence (Tel: +33661150410; Fax: +33140793835; E-mail: anna.kondratyeva@edu.mnhn.fr)

(a)	When diversity equals the sum or mean of originalities	1326
(b)	When diversity equals the sum or mean of originalities weighted by abundance	1326
(c)	When species' abundances maximize diversity	1327
(d)	When diversity and originality depend on a multidimensional space	1327
IV.	Reconciling the Diversity, Rarity and Originality Concepts for Their Useful Applications	1330
(1)	Understanding the evolutionary emergence of species' originality	1330
(2)	Analysing the dynamics of community assembly with rarity, originality and diversity	1332
(3)	Role of originality in ecosystem functioning	1333
(4)	Guiding conservation actions	1333
V.	Conclusions	1334
VI.	Acknowledgements	1335
VII.	References	1335
VIII.	Supporting Information	1337

## I. INTRODUCTION

One of the goals of the biological sciences is to identify ecological and evolutionary mechanisms driving community assembly that vary in space and time (Walker, 1992; Myers *et al.*, 2000; Pavoine & Bonsall, 2011). Indices of biological diversity (or biodiversity) are commonly used with this aim. Biodiversity has been defined in many ways. According to the Convention on Biological Diversity published in 1992, biodiversity means 'the variability among living organisms from all sources including, *inter alia*, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part: this includes diversity within species, between species and of ecosystems'. We will focus here on ecological and evolutionary aspects of inter-species diversity.

Historically, species diversity was measured as a function of species number and abundance, and the concept of diversity was thus connected to those of species abundance, commonness and rarity (Patil & Taillie, 1982). The concept of rarity has generally been used in the last few decades in ecology to describe the low abundance, restricted range size or habitat specificity of a species (Rabinowitz, 1981). However, many researchers have claimed the need also to describe variety in species' attributes since species are not equivalent and exchangeable (Findley, 1973, 1976; Vane-Wright, Humphries, & Williams, 1991; Faith, 1992). For example, we can expect that a set of species with a crocodile, an emu and a lion has higher diversity than does a set with a lion, a cheetah and a domestic cat because the species of the first set are more distant in the tree of life than are the species of the second set and have very contrasting biologies. Indices of phylodiversity and trait diversity have been produced (e.g. see Faith, 1992; Petchey & Gaston, 2002; Webb *et al.*, 2002; Hardy & Senterre, 2007; Helmus *et al.*, 2007; Villéger, Mason, & Mouillot, 2008; Pavoine & Bonsall, 2011). These measures of biological diversity synthesize the variety of species in terms of their different attributes: their phylogenetic relations (Vellend *et al.*, 2011; Tucker *et al.*, 2017) and their traits, including those that are qualified as functional (Petchey & Gaston, 2002).

Functional traits have been defined in the literature in different ways, notably as the traits that influence species' responses to environmental conditions (response traits) or that influence ecosystem properties (effect traits) (Lavorel & Garnier, 2002) and as 'the traits that are associated with species' ability to gain resources, disperse, reproduce, respond to loss and generally persist' (Weiher *et al.*, 2011, p. 2403). The diversity in species' functional traits was naturally named 'functional diversity' (Petchey & Gaston, 2002). As many studies used the term 'functional trait' to describe any measurable character of a species, we will consider here the diversity in species traits more generally, referred to as 'trait-based diversity' (Pavoine & Bonsall, 2011). The expression 'phylogenetic diversity' is used to describe both a concept of diversity in the evolutionary histories and relatedness among any set of taxa and a measure developed by Faith (1992) for conservation purposes. To avoid any confusion, we use 'phylodiversity' to describe this concept.

In parallel, a few measures described the degree of isolation of a species in a phylogeny (May, 1990; Vane-Wright *et al.*, 1991; Nixon & Wheeler, 1992). More recent studies developed many alternative measures accounting for branch lengths leading to the focal species in dated phylogenetic trees (Redding, 2003; Pavoine, Ollier, & Dufour, 2005; Redding & Mooers, 2006; Isaac *et al.*, 2007; Redding *et al.*, 2008; Huang, Mi, & Ma, 2011; Redding, Mazel, & Mooers, 2014; Pavoine *et al.*, 2017). All these measures evaluate the originality of each species in a phylogenetic tree.

An issue of many general concepts used in science is that the word used to designate the concept also has a common use for a general audience and thus could have different meanings. Rarity and originality are no exception. In the most common sense, the words 'rarity' and 'rare' are usually associated with a low probability of encountering some specific entity. The term 'original' is also used to designate an unusual entity. When used in science, these words may refer to different definitions. In ecology, species rarity usually represents the low probability of encountering the species. Species' originality represents the low probability of encountering the species' biological characteristics (phylogenetic position or traits).

Several studies thus used the term ‘originality’ to describe a concept, a measure or both (e.g. Pavoine *et al.*, 2005; Buisson *et al.*, 2013; Redding *et al.*, 2014) to quantify the degree to which species may harbour or actually harbour rare characters. Since Webb *et al.* (2002), phylodiversity and originality have largely been used in ecological studies as proxies for trait-based diversity and originality, respectively. Recent approaches now recommend considering phylo- and trait-based measures as complementary: the former to reveal historical and evolutionary processes, and the latter to reveal ecological processes (e.g. Grandcolas, 1998; Pavoine & Bonsall, 2011; Kelly, Grenyer, & Scotland, 2014; Gerhold *et al.*, 2015; Mazel *et al.*, 2017).

Indeed, as ecology and evolution share a wide range of scientific questions about biodiversity, there is an increasing interest in inter-disciplinary studies. Ecologists try to understand how species interact with each other and with their environment as well as how these interactions can influence the assembly patterns of multispecies communities (e.g. Weiher, Clarke, & Keddy, 1998; Emerson & Gillespie, 2008; Cavender-Bares *et al.*, 2009). Evolutionary biologists work at different time scales and seek to understand the origins and history of biodiversity and its variation (e.g. Mergeay & Santamaria, 2012). At the frontier between these disciplines, studies in evolutionary ecology are developing frameworks to explain biodiversity patterns and dynamics, combining ecological causes of evolution and evolutionary implications in community assembly and ecosystem processes (e.g. Mouquet *et al.*, 2012). Such developments are now also needed to explain originality patterns. The use of phylogenies in community ecology, macroecology, and conservation biology reflects the shared recognition that accumulated evolutionary differences may explain trait variation and thus predict biological and ecological processes. Phylogenetic approaches have revolutionized these disciplines (Mouquet *et al.*, 2012; Tucker *et al.*, 2017).

Although the concepts of diversity, originality and rarity are fundamentally connected, they have often been treated independently in the past two decades in ecological and evolutionary studies. Even if attempts to bridge two of these concepts do exist (e.g. Pavoine *et al.*, 2005; Redding & Mooers, 2006), very few studies have tried to combine them all (e.g. Rosauer *et al.*, 2009; Cadotte & Davies, 2010). As a result, our three main objectives are: (i) to bring attention to the fact that a conceptual explanation of the links between the three concepts is still missing; (ii) to highlight the need to clarify some aspects of the complementarities between these concepts and their measurement; and (iii) to demonstrate why these three concepts should be used together in evolutionary, ecological and conservation studies.

## II. RARITY AND ORIGINALITY: TWO ENTANGLED CONCEPTS

Since the first ecological studies, one of the main challenges has been to determine the contribution of each species to

biodiversity. Those contributions are measured *via* species characteristics that can be expressed by species rarity or originality. Here, we provide an overview of the definitions of the rarity and the originality of a species. We highlight that these two words refer to the same core concept of being unusual and point out the importance of spatial scale for their definitions.

### (1) The concept of species rarity in ecology and evolution

The measurement of species rarity is relative, as its definition and units depend on the context, nature, quality and quantity of data, constraining every study to define what they mean by ‘rare species’ (Magurran, 2004; Hessen & Walseng, 2008). Many different definitions and viewpoints on rarity exist in the literature with biological aspects (e.g. abundance), threat aspects (e.g. extinction risk) and value aspects (e.g. ‘how special species are’) (Gaston, 1997). In a seminal paper, Rabinowitz (1981) proposed a typology of rare plant species by crossing three characteristics: local population size (high or low), geographic range (large or small) and habitat specificity (wide or narrow). She proposed seven forms of rarity by combining these three dichotomized criteria, excluding the case where a species has high local population size, large geographic range and wide habitat specificity. If local population size, geographic range and habitat specificity are all scarce (the most drastic form of rarity), then a species will be prone to be the most endangered and to extinction.

Because Rabinowitz’s (1981) classification requires a considerable amount of information for a given taxon that is often not available, many studies use only one criterion or a combination of two to determine species rarity (Kunin & Gaston, 1993). Most past studies have favoured a definition of rarity relying on abundance and range size (Gaston, 1997). However, even if a broad consensus has been reached on these two aspects of rarity, abundance and range size may be measured by many different approaches (Gaston, 1997). For example, the geographic range may be analysed in terms of extent of occurrence (EOO, total range extent even if unevenly occupied: Gaston, 1991; Kunin & Gaston, 1993), area of occupancy (AOO, amount of sites or grid squares inhabited: Gaston, 1991; Kunin & Gaston, 1993), or both. This distinction allows the identification of species that occur only in a restricted, localized area (low EOO) from species occupying a low proportion of the area within their otherwise large range boundaries (high EOO but low AOO) (Hartley & Kunin, 2003).

Regarding species abundance, Rabinowitz (1981) considered local abundance in terms of population size, which can be understood as the number of individuals in a population. In the context of biodiversity measurement, the number of individuals is also currently the most frequently used aspect of species abundance. Abundance, however, can be measured by different means (e.g. absolute and relative density, biomass, per cent cover), and these could also be included in species diversity analyses (e.g. Lyons, 1981).



In addition, the relevance of considering a single aspect of rarity as an indicator of all others is likely dependent on the phylogenetic and spatial scales considered. For example, although a positive correlation between species' range size and mean local population size at occupied sites has been reported by many studies across many different taxa and habitats (see Brown, 1984; Gaston, 1996), this correlation was sometimes found to be rather weak or even negative, particularly when the studied species were phylogenetically distant (Brown, 1984; Gaston, 1997; Johnson, 1998). Population size also varies depending on where in the range of the species it is measured, introducing a spatial contingency into defining rarity (Brown, 1984). Finally, part of this correlation may be due to sampling artefacts, as species with low local abundance are likely to be recorded from fewer sites than the number at which they actually occur (Gaston, 1997; Hessen & Walseng, 2008). Rarity is thus a multifaceted concept; and it is reasonable to argue that when data are available, several facets should be considered and compared (Gaston, 1997).

More recently, the concept of rarity has been extended to the functional and phylogenetic characteristics of species (Pavoine *et al.*, 2005). In his paper on the definition(s) of rarity, Gaston (1997) reported that taxonomic distinctness had been considered to define rare species. Pavoine *et al.* (2005) highlighted that phylogenetic distinctness may represent how rare a species' traits are, and Pavoine *et al.* (2017) reported that many recent studies measured distinctness directly from a finite set of traits (e.g. Moullot *et al.*, 2013). Violle *et al.* (2017) subsequently proposed an expanded framework for rarity to study what they called 'functional trait rarity': a species is rare if it has a low abundance and distinct traits relative to the other species of an assemblage. They also stressed the need to study functional rarity from an evolutionary perspective by examining the phylogenetic signal of trait rarity. Overall, species rarity can thus be based on a variety of species attributes. Species' originality works with the phylogenetic and trait-based aspects of rarity.

## (2) The concept of originality or phylo- and trait-based rarity

Rarity is a concept widely used to determine a species' contribution to the diversity of a finite set of species. Nevertheless, fully to evaluate a species' contribution to phylo- and trait-based diversity, an additional measure of species characters is needed that can be made by means of species' originality. Here, we discuss the definition of species' originality, different types of originality measures and their link to abundance-based rarity.

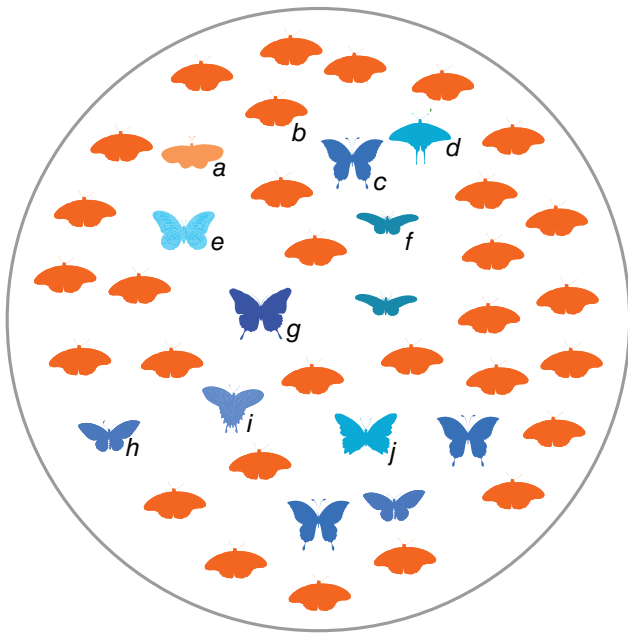
The vocabulary employed to describe species' characters is continuously evolving and differs between evolutionary and ecological studies, leading to potential confusion when a term is employed without a clear definition or reference (Pavoine & Bonsall, 2011). Since the 1990s, various terms, such as 'originality', 'distinctness', 'distinctiveness', 'uniqueness' and 'isolation', have sometimes been used to refer to the same concepts but sometimes not. Thus, there is an inconsistency

in the literature in the terms used to describe measures and concepts of species rarity and in how they should be used.

The term originality describes species' general rarity using characteristics linked to traits and phylogeny. Pavoine *et al.* (2017) defined originality as the rarity of species' characteristics in a given set of species, where a characteristic can be a position on a phylogenetic tree or the state of a functional trait. As recommended by Pavoine *et al.* (2017), we define 'phylogenetic originality' as synonymous with the terms 'evolutionary distinctiveness' (Isaac *et al.*, 2007), 'evolutionary isolation' (Redding *et al.*, 2014), 'phylogenetic rarity' (Winter, Devictor, & Schweiger, 2013) and 'phylo-based rarity'. A species without close sister species in a phylogenetic tree is likely to have high phylogenetic originality and thus a high contribution to phylodiversity (Redding *et al.*, 2008). In recent decades, particular attention has also been paid to species' functional traits that indicate species' roles in ecosystem functioning (e.g. Lavorel & Garnier, 2002; Moullot *et al.*, 2008; Schmera, Erős & Podani, 2009; Buisson *et al.*, 2013; Moullot *et al.*, 2013; Rosatti, Silva, & Batalha, 2015; Brandl *et al.*, 2016). The assumption that phylogenies indirectly comprise the evolutionary changes in species' characters may explain why the first originality approaches were inferred from phylogenetic trees.

These approaches were adapted to the analysis of species' functional traits, to measure what we hereafter refer to as 'functional originality'. We consider 'functional originality' a synonym of 'functional rarity' and 'rarity of functional traits'. More generally, we consider below 'trait-based originality' a synonym of 'trait-based rarity' and 'rarity of traits'. Hence, a species with very distinct trait values compared to those of other species is expected to have a higher contribution to trait-based diversity (Jarzyna & Jetz, 2016). Similar to abundance-based rarity, species' originality is relative to the values of the other species in a set. Following Pavoine *et al.* (2017), we define the uniqueness of a species as an unshared part of the species characteristics in a set. Strict uniqueness is thus an extreme case of originality in which a species does not share any of its characteristics. Finally, redundancy is antonymous with uniqueness and measures the number of shared characteristics of species. Therefore, originality is the full contribution of a species to the diversity of the set composed of species' unshared (uniqueness) and shared (redundancy) characteristics.

Commonness and rarity traditionally have been presented as the extremes of a gradient of the abundance-based rarity of species (Kunin & Gaston, 1993). Redundancy and uniqueness are the extremes of a gradient of species' originality, measured in terms of species' phylogenies and traits (Redding *et al.*, 2014). Some studies have analysed potential correlations between the two gradients, asking whether original species are rare in terms of abundance at different spatial scales (Mi *et al.*, 2012; Pigot *et al.*, 2016). Researchers found variable results showing that abundance-based rarity and phylo- and trait-based originality are not always correlated depending on the



**Fig. 1.** Theoretical illustration of the concept of abundance-weighted functional originality. Here, we consider an assemblage of butterflies. The drawings represent butterfly species of different shapes, for which colours are used to represent a functional trait: two species with similar colours have similar values for the trait. Species *a*, *d*, *e*, *g*, *i* and *j* are each represented by one individual; species *b* by 30 individuals; species *c* by three individuals; and species *f* and *h* by two individuals each. The species *a* and *b* are original because they are the sole species with shades of orange. The other eight species have different shades of blue and green. However, given that species *b* is very abundant (being represented by 30 individuals), the abundance-weighted originality of species *a* can be considered low, and if originality is measured at the individual level, instead of at the species level, then the originality of any individual of species *b* can be considered low.

spatial resolution of a study. Violle *et al.* (2017) combined both gradients to define what they called the ‘functional rarity’ of species (see Fig. 1 in Violle *et al.*, 2017): a species is functionally rare if it is both original and scarce. Their concept of functional rarity is not equal to our concept of functional rarity, which expresses the scarcity of species’ traits rather than the scarcity of the species themselves.

The measure of originality indeed can be weighted by abundance, leading to a third gradient of rarity where originality and abundance-based rarity are entangled. Along this gradient, a species is original in an assemblage if its characteristics are rare among the individuals of the assemblage (Fig. 1). Similarly, Violle *et al.* (2017) suggested an abundance-weighted measurement of functional trait distinctiveness, making it dependent on species scarcity (Box 2 in Violle *et al.*, 2017). Hence, originality (trait-based and phylo-based rarities) and abundance-based rarity are closely related concepts but use different types of data for characterizing species based on their abundance, traits or phylogeny (Fig. 2). More generally, other aspects of species

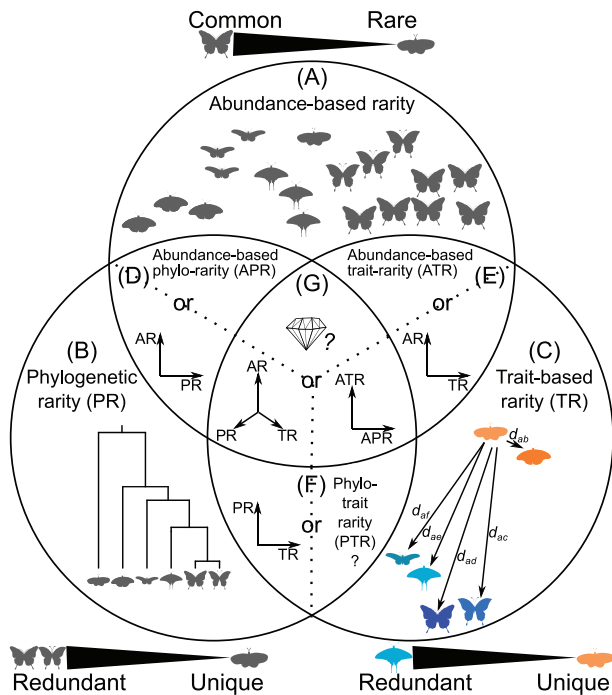
rarity could also be added to this framework, as, for example, the link between originality and range size has also been studied (Mouillot *et al.*, 2013; Grenié *et al.*, 2018).

### (3) The measurement of originality

The concept of species’ originality can be evaluated by different measures that rely on various types of data (see online Supporting information, Appendix S1). The first originality indices were developed on a tree structure and thus measured phylogenetic originality. However, they were applied to undated phylogenies ignoring the evolutionary time of change in taxa (May, 1990; Vane-Wright *et al.*, 1991; Nixon & Wheeler, 1992). Those indices used the number of internal nodes or the number of branches descending from each node to compute species’ originality. Unfortunately, they can hardly distinguish the originality between species from the same clade (Huang *et al.*, 2011; Redding *et al.*, 2014). These indices are useful when the branch lengths on a phylogenetic tree cannot be estimated.

With the improved access to dated phylogenetic trees, measures of phylogenetic originality have been developed that account for branch lengths. Several of these measures are derived from the Phylogenetic Diversity (*PD*) index of Faith (1992): the sum of branch lengths on a phylogenetic tree. Some of them are partly redundant, and their formulae are very similar [the indices of fair proportion (*FP*) (Redding, 2003), equal splits (*ES*) (Redding & Mooers, 2006) and evolutionary distinctiveness (*ED*) (Isaac *et al.*, 2007)]. Species with few relatives and deep terminal branches would be more original than species with many close relatives (Frishkoff *et al.*, 2014). Those indices can thus be used when one wants to know which species subtends most of the evolutionary history in a given set of species. According to Redding *et al.* (2014), technically, all tree-based indices of species’ originality combine two main aspects of originality: the average distance to another species on the tree [average phylogenetic distance (*APD*) (Redding *et al.*, 2014)] and the terminal branch lengths that connect the species to the rest of the tree [pendant edge (*PE*) (Redding *et al.*, 2008)]. Another framework of tree-based measures of species’ originality has been proposed that uses the genome evolution model of species characters called ‘character rarity’ (*CHR*; Huang *et al.*, 2011). Its advantage over other indices is that it incorporates models of dynamic processes, such as character changes and distribution along lineages during evolution. However, applications of this framework are still scarce.

With the accumulation of species’ trait information in large databases, a new type of study appeared that considered a limited number of traits (Violle *et al.*, 2007; Mouillot *et al.*, 2008; Hidasi-Neto, Loyola, & Cianciaruso, 2015; Ricotta *et al.*, 2016). As measures of originality already existed for phylogenetic trees, trait data were transformed into dendrograms (trees) to measure trait-based originality (Buisson *et al.*, 2013; Pavoine *et al.*, 2017). A new challenge appeared: how best to construct trait-based dendrograms from a finite set of traits and how to avoid the distortion of the original trait data (Petchey & Gaston, 2002; Mouchet



**Fig. 2.** Diagram illustrating the three types of rarity discussed herein. Each black circle represents one type of rarity. The areas where the circles cross designate studies where two or more aspects of originality are compared or combined. We illustrate each core concept with a theoretical data set, represented here by seven theoretical butterfly species. (A) Abundance-based rarity (AR) (as an illustration, it is here inversely related to the number of individuals of each species); (B) phylo-based rarity (PR) representing species evolutionary histories (here exemplified by a theoretical phylogenetic tree); (C) trait-based rarity (TR) reflecting species' trait dissimilarities. In C, butterfly colours represent species' traits. We illustrate the trait-based rarity (or originality) of species *a* by the dissimilarities *d* between this species and all the others. The gradient above the top circle represents the two extremes of AR: commonness and rarity. The gradients below the bottom left and right circles represent phylogenetic and trait-based originality, respectively, from redundancy to strict uniqueness. (D, E) PR and TR can be compared with AR using two-dimensional graphs. The phylogenetic and trait data can also be weighted by species' abundances, leading to abundance-weighted trait-based rarity (ATR) and abundance-weighted phylo-based rarity (APR) (see, e.g. Fig. 1). (F) Studies that analyse both TR and PR thus far have compared the two aspects by two-dimensional graphs. No mathematical measures have been developed that combine these two aspects, which could lead to phylo- and trait-based rarity (PTR). (G) The area where the three circles cross would permit identification of the rarest species. Depending on whether different aspects of species rarity are treated independently or are combined, such rare species could be defined as having high AR, TR and PR (identified by the three-dimensional plot), high ATR and APR (identified by the two-dimensional plot), or high abundance-weighted PTR (identified by the diamond). The red circles and question marks show where further research is critically needed to determine how and why phylogenetically and functionally original species emerged in the course of evolution and how this emergence has influenced species abundance.

*et al.*, 2008), as most traits do not possess any structured hierarchy. A common approach is to first calculate pairwise trait-based dissimilarities between species and then construct the dendrogram from these dissimilarities with a clustering algorithm. Numerous mathematical equations exist to calculate the dissimilarity between two species using various traits (Pavoine *et al.* 2009). Additionally, the choice of the clustering method influences the shape of the dendrogram and eventually influences trait-based originality values (Mouchet *et al.*, 2008; Maire *et al.*, 2015). Transforming trait values into dendrograms introduces the risk of distorting the information on traits (Pavoine *et al.*, 2017) but could be justified as a way to compare trait-based originalities to phylogenetic originalities for the same set of species. Thus, in theory, any of the phylogenetic tree-based originality measures can be applied to trait-based dendrograms.

The problem raised by deforming trait data when building functional dendrograms can be bypassed by using trait-based dissimilarity matrices. Originality measures have therefore been adapted to work with dissimilarity matrices. Indeed, the average (*AV*) and nearest neighbour (*NN*) indices, working directly with trait-based dissimilarities between species, are the alternatives of the tree-based *APD* and *PE* indices (Pavoine *et al.*, 2017; Violle *et al.*, 2017; see also Appendix S1). *AV* is the average dissimilarity to all other species, and *NN* is the lowest dissimilarity to all other species. By extension, these dissimilarity-based indices can, in turn, incorporate not only trait-based but also phylogenetic dissimilarities calculated, for example, as the sum of branch lengths in the smallest path that connects the two species on the phylogenetic tree or as the time since their most recent common ancestor (Pavoine *et al.*, 2017). Computing phylogenetic dissimilarities is useful for comparing equally the trait-based and phylogenetic originalities of the same set of species without taking the risk of distorting trait data.

Clustering methods have also been criticized in favour of the use of multidimensional space (Maire *et al.*, 2015). A multidimensional trait space is a geometrical space representing the distribution of species according to their trait values (Mouillot *et al.*, 2013). It can be constructed in several ways (see Appendix S1). The coordinates of the species projected as points along the axes of such a space could be used for measuring originality. For example, originality has been measured as the distance from a species to the centroid (mean position) of all species (Magnuson-Ford *et al.*, 2009; Buisson *et al.*, 2013). In theory, the multidimensional approach was created for the trait-based context but could also be applied to a phylogenetic space obtained, for example, with a principle coordinate analysis applied to phylogenetic dissimilarities between species (e.g. Sobral, Lees, & Cianciaruso, 2016).

Pavoine *et al.* (2005) introduced another family of originality measures by analysing the composition of theoretical species assemblages that would maximize an index of diversity. The first index of this family, the *QE*-based index (*Q<sub>b</sub>*), relies on Rao's quadratic entropy (*Q*) (1982). *Q* is an index of diversity equal, in our context, to the phylogenetic



or trait-based dissimilarity between two randomly selected species in an assemblage.  $Qb$  uses the abundance values that species should have to maximize the phylogenetic or trait-based dissimilarity between two randomly selected species (Huang *et al.*, 2011). Up to a certain degree, the more abundant the original species are, the more the diversity of a set is increased (Pavoine *et al.*, 2017). The hypothetical abundances of species that maximize  $Q$  thus reflect species' originality. Using the same approach, Pavoine *et al.* (2017) developed the  $Rb$  index using another index of entropy ( $R$  index) closely related to  $Q$ .  $Rb$  turned out to be similar, although not equal, to the  $AV$  index (the mean trait-based or phylogenetic distance to all other species), but it better discriminates original from redundant species. The  $Qb$  and  $Rb$  indices are discussed in detail in Section III.3.

Finally, a measure of originality based solely on species' phylogeny or traits does not account for other aspects of rarity such as population size or species range size. As shown in Fig. 2, species originality can be weighted by species abundance. Regarding phylogenetic data, Cadotte *et al.* (2010), for example, modified the  $FP$  index (see Appendix S1) to account for species abundance when species' originalities are calculated within a local community. To do so, at each tip (species) of a phylogenetic tree, they artificially added as many branches with a length of zero as there were individuals from the corresponding species in the local community. In other words, they weighted each terminal branch by the number of individuals it subtends. They then calculated originality using the  $FP$  index but replaced species with individuals, leading to their abundance-weighted evolutionary distinctiveness ( $AED$ ). Later, they modified their index, replacing individuals with populations or sites (Cadotte & Davies, 2010), leading to the biogeographically weighted evolutionary distinctiveness ( $BED$ ) index.

By doing so, Cadotte & Davies (2010) evaluated the originality of an individual or a population of a species. Other approaches have been proposed using different types of data. For example, Ricotta *et al.* (2016) proposed measuring the originality of a species as the abundance-weighted mean of the trait-based difference between this focal species and all other species in a set (weighted version of  $AV$  index). Laliberté & Legendre (2010) measured the distance in a functional trait space between a species point and the abundance-weighted centroid of all species points, which can be viewed as a measure of abundance-weighted originality. Other indices were also developed that weight originality by species' extinction risk (e.g. Steel, Mimoto, & Mooers, 2007).

#### (4) How should the originality measure be chosen?

Measures of species' originality are expressed by formulae, and their mathematical properties can influence the results and conclusions of a study. Several studies have compared some of these measures, seeking to establish the relations between them (Redding *et al.*, 2008, 2015; Huang *et al.*, 2011; Vellend *et al.*, 2011; Redding *et al.*, 2014; Pavoine *et al.*, 2017). According to these studies, some measures are more redundant than others in capturing species-specific

information and in their contribution to phylodiversity and trait-based diversity. Each type of originality measure has its own advantages and drawbacks. In general, tree-based indices would be influenced by tree topology (terminal and deeper branch lengths, unresolved nodes, root consideration), but trees without branch lengths fail to discriminate individual species.

If one wants to compare species' phylogenetic and trait-based originalities, then dissimilarity-based indices are a better choice than tree-based indices, as the former avoid the potential distortion of trait data. The use of a multidimensional trait space permits potential original species and their traits to be visualized. However, all axes need to be retained. For example, a common practice when measuring trait-based diversity is to apply a principal coordinate analysis to trait dissimilarities and then select the first few axes to calculate diversity, a constraint that may be forced by the mathematical properties of the diversity indices (e.g. Villéger, Mason, & Mouillot, 2008). This dimension reduction can hide some important species if their originalities are explained by traits on the non-visualized axes. Finally, originality indices, such as  $Qb$  and  $Rb$  (Appendix S1), predict species' originality expressed as the theoretical abundance of a given species that maximizes the diversity of a set. Note that, as originality is context dependent, its application to a set of two species would give equal originality to both species; the concept of originality is thus useful as soon as there are at least three species in a set.

#### (5) Spatial scale matters

Spatial resolution is essential in species' originality analyses that are relative to a reference set of species. Phylogenetic originality is usually measured at the level of an entire clade in an evolutionary context (e.g. Jetz *et al.*, 2014). This allows the identification of the most phylogenetically original species on Earth and the analysis of their biogeography (e.g. Safi *et al.*, 2013) and extinction risks (e.g. Isaac *et al.*, 2007). Such global phylogenetic originalities have been typically summed for all species present in a region to evaluate the worth of the region for conservation: how many globally original species occur in the region (e.g. Pollock, Thullier & Jetz, 2017), and how original are these species (e.g. Safi *et al.*, 2013; Jetz *et al.*, 2014)?

Such an approach has also been applied at local levels. In this case, phylogenetic originality was calculated within the species pool present at a regional (or continental) or global level. Then, these regional or global originalities were summed or averaged for all species within a local site to identify priority sites for conservation (e.g. Veron, Clergeau, & Pavoine, 2016) or to determine variations in the presence of original species among environments (e.g. Morelli *et al.*, 2018). Most of these studies on phylogenetic originality focused on the preservation of evolutionary history.

Another approach, which Redding *et al.* (2015) explored, is to calculate species originalities directly within a local site and to compare these local values with regional originalities of the species. They demonstrated that the correlation between

local and regional originality depends on the originality index used. Analysing Nearctic and Neotropical birds, they found that this correlation is spatially structured and may depend on the habitat, being notably higher in forested habitats. More studies of this type are needed to evaluate how local conservation projects based on local originalities could complement regional and global projects relying on large-scale originalities.

In contrast to phylogenetic originality, trait-based (especially functional) originality is rarely measured at the global level (see, e.g. Grenié *et al.*, 2018). Instead, it is usually evaluated at the regional and local levels, as species' characteristics are directly related to ecosystem functioning and depend mostly on field data (Da Silva, Silva & Batalha, 2012; Mouillot *et al.*, 2013; Rosatti *et al.*, 2015). Several of these studies have underlined the potential vulnerability of functionally original species (Mouillot *et al.*, 2013) and the negative impact of human disturbance on these species, with potential consequences on the sustainability of ecosystems (Rosatti *et al.*, 2015). However, these conclusions may depend on the taxa, traits and geographic areas analysed (Da Silva *et al.*, 2012; Brandl *et al.*, 2016).

Future studies could thus focus on determining and explaining scale-dependent patterns in both phylogenetic and trait-based originalities as local, regional and global originality scores may or may not be correlated depending on the processes that drive diversity patterns (Redding *et al.*, 2015).

### (6) Directions for future research on species' originality

We have shown that studies on species rarity recently have been improved by the addition of aspects of trait-based and phylogenetic originalities. Although many originality measures have been developed at the species level, they can be applied to other units of biodiversity, notably, to individual organisms (Pavoine *et al.*, 2017). For example, replacing trait-based dissimilarities between species with trait-based dissimilarities between individuals and replacing species points in multidimensional trait space by individual points allow all originality indices presented above to be transferred from the measure of the originality of a species to the measure of the originality of each individual (de Bello *et al.*, 2011; Violle *et al.*, 2012). For example, if information on estimated among-individual divergence is available on a phylogenetic tree, Cadotte *et al.* (2010, Appendix S2) proposed replacing species with individuals in the *FP* originality index to measure the originality of each individual. Thus, from a mathematical viewpoint, it is possible to apply all originality indices at an individual level and access the intraspecific variability of genes and traits. Applying such detail is limited by the scarcity of relevant data and by the relative cost of obtaining these data; however, this approach could provide more accurate results when trait variation is large within species.

Although we focused our review on inter-species diversity with aspects of phylo- and trait-based diversity, the concept of originality was also treated in genetics. For example, genetic

originalities can be calculated by replacing phylogenetic distances between species with genetic distances based on a genome evolution model (Thaon d'Arnoldi, Foulley, & Ollivier, 1998; Laval *et al.*, 2000; Huang *et al.*, 2011). Connecting these different fields of research could improve biodiversity analyses, allowing more continuous multiscale analyses, better to connect patterns to underlying processes in ecological and evolutionary studies.

In Section III, we demonstrate how the indices of originality relate to indices of diversity. We develop potential applications of originality indices with real case studies in Section IV.

### III. LINKS BETWEEN MEASURES OF DIVERSITY, ABUNDANCE-, PHYLO- AND TRAIT-BASED RARITY

A biodiversity measure is a calculation method expressed by a mathematical formula that allows specific values for the amount of variety in a biological system to be computed; in our case, these values are computed for a set of species. The oldest and most intuitive biodiversity measure is species richness, computed as the number of species (Magurran, 2004). Nevertheless, species richness has several drawbacks. First, it is strongly dependent on sampling effects: in a highly diverse community, the observed number of species may greatly underestimate the real number of species because species with very low abundance frequently will be absent from even very large samples (Lande, 1996). Second, it would give equal diversity to a region dominated by a single species with two rare species and a region with three species having even abundances (e.g. Magurran, 2004). Third, it does not include any information on species' traits and evolutionary histories. Fourth, it depends strongly on the definition of the concept of species (Gaston & Spicer, 2004; Groves *et al.*, 2017) and on taxonomic incompleteness (e.g. Delić *et al.*, 2017). To overcome these drawbacks, alternative diversity indices have been developed that include species' abundances, phylogenies and traits and thus incorporate various facets of species rarity. Diversity is measured at the level of the species set, while rarity and originality indices provide a value for each species that is linked to its contribution to the diversity of the set.

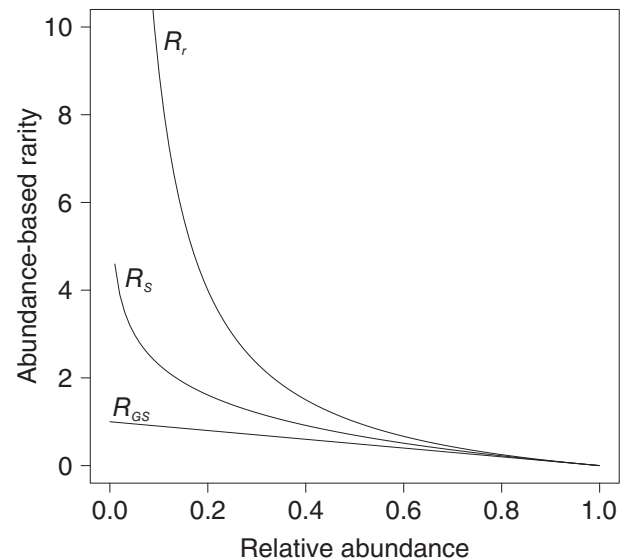
In the last two decades, the development of trait-based diversity and phylodiversity indices has expanded. This variety of indices is partly explained by the fact that one single mathematical formula cannot alone encompass all aspects of biodiversity in a set, especially phylogeny, functionality and abundance (Mason *et al.*, 2005; Mazel *et al.*, 2016). Several previous studies tried to define semantic frameworks for measurements of species' phylo- and trait-based diversities, in which they specified whether each measure of diversity was weighted by abundance data (e.g. Ricotta, 2007; Pavoine & Bonsall, 2011; Vellend *et al.*, 2011; Tucker *et al.*, 2017). Our aim here is not to review all mathematical indices of species' phylo- and trait-based diversities but to analyse a small set of

previous approaches for which we can try to distinguish the relative importance of abundance-based, phylogenetic and trait-based rarities. Although most of the diversity indices we discuss in this part were developed either for phylo- or trait-based diversity, all our reasoning could easily be adapted to be applied to both contexts.

### (1) Biodiversity as a mean of abundance-based rarities

Indices that relied on species number and species' abundances were named indices of species diversity. A myriad of such indices have been developed in the literature (e.g. Magurran, 2004). The most famous and most commonly used are the Gini–Simpson index (Gini, 1912; Simpson, 1949) and the Shannon index (Shannon, 1948). In Fig. 3, we provide Patil & Taillie's (1982) demonstration that species richness, the Shannon index, and the Gini–Simpson index can be written as the mean of species abundance-based rarity values. The three indices differ in their sensitivities to the presence of rare species. The most sensitive measure is species richness, as it gives equal importance to all species regardless of their abundance. The least sensitive is the Gini–Simpson index, which overweights common species relative to rare species (Lande, 1996). Thus, using abundance in indices of species diversity is actually a way to give a different value to each species. For example, giving the same value to each individual of a species forces us to give different values to the species, as species are represented by different numbers of individuals in a community. These well-known diversity indices illustrate that the first links between diversity measurement and rarity measurement date back to around 40 years ago, but these links concerned only abundance-based rarity. They did not refer to the phylogenetic or functional trait characteristics of the species.

The simplest approach for measuring aspects of phylo- and trait-based diversity naturally consisted of replacing species in the traditional diversity indices discussed above with trait-based or phylogeny-based entities. When the analysis of species' functional traits as an aspect of diversity measurement started to emerge in the ecological literature, the number of functional groups represented by the species in a community was the main measure of trait-based diversity (Petchey & Gaston, 2002). A functional group is a group of species united by their similarities in a given trait value or set of traits (Brandl *et al.*, 2016). The traditional diversity measures can be applied to the abundance of each functional group instead of to the abundance of each species (Hooper *et al.*, 2002). In that case, trait-based diversity can be related to the abundance-based rarity of functional groups. Such an approach could also be applied to clades in a phylogenetic tree, replacing functional groups with clades (Pavoine, Love, & Bonsall 2009). Although such approaches allow phylodiversity and trait-based diversity indices to be connected to abundance-based rarities, they do not connect them directly to measures of originality, that is, with trait-based and phylogenetic rarities.



**Fig. 3.** Functions of rarity associated with three traditional indices of diversity. A simple index to measure the diversity of a set is the number of species, or 'species richness'. A related index is the number of species minus one, for which an assemblage with only one species has null diversity. Indeed, the minimum value of many indices of species diversity is zero and is reached when a single species dominates the assemblage (its relative abundance is 1). Let  $S$  be the number of species in the assemblage;  $p_i$ , the relative abundance of species  $i$ ; and  $\mathbf{p}$ , the vector of species' relative abundances. Patil & Taillie (1982) emphasized that a richness-related ( $H_r$ ), the Shannon ( $H_s$ ) and the Gini–Simpson ( $H_{GS}$ ) indices can be written as follows:  $H_r(\mathbf{p}) = S - 1 = \sum_{i=1}^S p_i R_r(p_i)$ ,  $H_s(\mathbf{p}) = -\sum_{i=1}^S p_i \log(p_i) = \sum_{i=1}^S p_i R_s(p_i)$ , and  $H_{GS}(\mathbf{p}) = 1 - \sum_{i=1}^S p_i^2 = \sum_{i=1}^S p_i R_{GS}(p_i)$ , where  $R_r(p_i) = 1/p_i - 1$ ,  $R_s = -\log(p_i)$ , and  $R_{GS} = 1 - p_i$ . The functions  $R_r$ ,  $R_s$  and  $R_{GS}$  represent how rare a species is. In Patil & Taillie's (1982) framework, rarity means low abundance (for example, in terms of biomass or number of biological organisms).  $H_{CDT}$  entropy generalizes all these indices of diversity (Havrdá & Charvat, 1967; Daróczy, 1970; Tsallis, 1988). Its formula is  ${}^q H_{HCDT}(\mathbf{p}) = \left(1 - \sum_{i=1}^S p_i^q\right) / (q - 1) = \sum_{i=1}^S p_i [{}^q R_{HCDT}(p_i)]$ . If  $q = 0$ , then the  $H_{CDT}$  index is  $H_r$ ;  $q$  tending to 1 gives  $H_s$ , and  $q = 2$  gives  $H_{GS}$ . The general form of the function of rarity associated with the  $H_{CDT}$  index is  ${}^q R_{HCDT}(p_i) = \left(1 - p_i^{q-1}\right) / (q - 1)$ . The sensitivity to rare species decreases with  $q$ . The figure shows how the functions of rarity  $R_r$ ,  $R_s$  and  $R_{GS}$  vary if  $p_i$  increases from zero to one.

### (2) Explicit link between originality and diversity indices

Because a myriad of species diversity indices have been developed since the 1970s, a myriad of phylo- and trait-based diversity indices could also be imagined by replacing species abundance in species diversity indices, such as the Gini–Simpson and Shannon indices, with species originality or abundance-weighted originality. This approach was adopted, for example, by Cadotte *et al.* (2010), Scheiner

(2012) and Scheiner *et al.* (2017). With such an approach, the diversity would be high if the number of species was high and if species' originalities were even. If originalities are weighted by abundance (Figs 1 and 2), then the diversity would be high if the distribution of abundance was imbalanced such that the most original species have the highest abundances (Cadotte *et al.*, 2010).

Ricotta (2004) previously proposed a different approach in which abundance and originality were treated as independent information on species (instead of using abundance-weighted originalities) and were combined into one measure of diversity. Further, he used taxonomy instead of a phylogeny, leading to taxonomic originality, where a species is original if there are no or relatively few species from the same genus, family, order, etc. We will see in Section III.3 that many diversity indices can also be considered as functions (e.g. sum, mean or abundance-weighted mean) of originality values, even if the indices were not developed with that goal in mind. The approach in which species' originality values and potential species abundances are combined in a mathematical formula is thus central to the measurement of biodiversity.

### (3) Most diversity indices only implicitly depend on species' originalities

Contrary to the few approaches presented in the previous Section III.2, species' originality, obtained through phylo- and trait-based rarities, was often only implicitly incorporated into measures of phylo- and trait-based diversity. We will show here that many of these measures of diversity can instead be seen as functions of species' originality or are largely linked in their formulae to a measure of species' originality, even if they were not developed explicitly as functions of species' originality.

#### (a) When diversity equals the sum or mean of originalities

Early on, Faith (1992) suggested replacing species in diversity measurements with features (or character states) of species. He then developed, as a proxy for feature richness, the *PD* index, which is the sum of branch lengths on a phylogenetic tree. His approach assumes that branch lengths reflect feature richness. Ten years later, Petchey & Gaston (2002) similarly proposed representing trait diversity as the sum of branch lengths on a trait-based dendrogram, an index named 'functional diversity' (*FD*). Two of the most used phylogenetic originality measures, the *FP* and *ES* indices (Appendix S1), can be written as additive decompositions of *PD* (Redding & Mooers, 2006; Isaac *et al.*, 2007). Thus, the sum of the originality values over all species in a set is equal to the *PD* of the set. If these originality indices, *FP* and *ES*, are calculated from a trait-based dendrogram rather than from a phylogenetic tree, then the sum of species' trait-based originality values over all considered species equals the *FD* index.

Two other commonly employed indices of community phylogeny, the mean pairwise distance (*MPD*) index

and the mean nearest taxon distance (*MNTD*) (Webb *et al.*, 2002), are means of species' originality values. The mean of the *APDs* (average phylogenetic distance to all other species) over all species in a set is equal to the *MPD* or the average phylogenetic distance between any two species in a set (Redding, Mazel, & Mooers, 2014). The minimum phylogenetic distance to another species is an index of strict uniqueness (see indices *NN* and *PE* in Appendix S1). Its mean over all species in a set is equal to the *MNTD* index used during the last two decades to detect the effects of competition in community assembly (Kraft *et al.*, 2007). Such reasoning could also be applied to trait-based dissimilarities between species. Indeed, indices such as *MPD* and *MNTD* were used early on by Findley (1973, 1976) to analyse the morphometrical diversity of bats. These indices do not integrate information on how abundant species are. However, other diversity indices can be seen as functions of abundance-weighted originalities.

#### (b) When diversity equals the sum or mean of originalities weighted by abundance

Regarding phylodiversity, the sum over all individuals of the *AED* value (individual's phylogenetic originality) of Cadotte *et al.* (2010) is equal to Faith's *PD* (phylogenetic diversity) of the species pool. Similarly, the sum over all populations of the *BED* value (the phylogenetic originality of each population of a species) of Cadotte & Davies (2010) is also equal to Faith's *PD* of the species pool. Furthermore, Cadotte & Davies (2010) suggested summing the *BED* values over all species in a site, leading to a measure of the relative originality of each site, and thus to its relative conservative worth within the studied regional area. This *BED* index was inspired by the index of phylogenetic endemism (*PE*) of Rosauer *et al.* (2009), which measures the spatial restriction of the evolutionary history in a site. *PE* is the sum of the ratios branch length/clade range for each branch of the phylogenetic tree pruned to retain only the species of a site. The clade range is the union of the ranges of all species descending from the branch. As such, *PE* combines notions of phylodiversity (if species were ubiquitous, *PE* would be maximal, which means that regional and local phylodiversity would be equal), range-based rarity (each branch length is divided by clade range), and originality (if a species with a small range descends from a long branch, it is likely to have a high contribution to *PE*). Cadotte & Davies (2010) developed *BED* to complement the concept of phylogenetic endemism with a measure that can be calculated at different levels: at the species, population, local-site or regional levels.

Pavoine *et al.* (2009) used clades (groups of all species descending from a specified ancestor) in a phylogenetic tree to obtain measures of phylodiversity. They divided a tree into time periods defined between two consecutive speciation events. They then applied the Havrda–Chavát–Daróczy–Tsalis (*HCDT*) diversity index (a generalization of the well-known species richness, Shannon and Gini–Simpson indices; Fig. 3) at each period, using the number of clades that descend from the period and their



relative abundances (the abundance of a clade is the sum of abundances of all its species). At the most recent period, clades are species. However, from the second period, at least one clade has more than one species. The associated index of diversity is equal to the length of the period (in million years of evolution) times the clade diversity (as measured by the *HCDT* index applied to clades). Such an approach brings an evolutionary dimension to traditional diversity analysis and has the advantage of considering species abundance when measuring the phylodiversity of a local community.

This approach is a generalization of the *HCDT* index for phylogenetic studies, which was at the core of the theory of Patil & Taillie (1982) to express diversity in terms of the mean of species' rarities (Fig. 3). In Patil & Taillie's (1982) theory, rarity was expressed as low abundance. In Table 1, we propose a new theory using the generalization of the *HCDT* index from Pavoine *et al.* (2009). In this new theory, rarity is measured as an abundance-weighted phylogenetic originality, and the measure of phylodiversity is quantified as the mean of these abundance-weighted originality values. This theory could also apply to trait-based diversity if a dendrogram is available.

Another currently widespread approach to measure biodiversity consists of defining trait-based or phylogenetic dissimilarities between species and taking their sum or their mean (e.g. Rao, 1982; Walker, Kinzig, & Langridge, 1999; Schmera, Erös, & Podani, 2009). The quadratic entropy index ( $Q$ ; Rao, 1982) is a diversity measure that can handle dissimilarities between species and species abundance. Shimatani (2001) demonstrated how to decompose  $Q$  into three underlying components: (i) the Gini–Simpson index (an index of species diversity that itself is a combination of species richness and species evenness (Fig. 3)); (ii) the non-weighted mean of the (trait-based or phylogenetic) dissimilarities between species; and (iii) a measure of the balance between species' abundances and the (trait-based or phylogenetic) dissimilarities between species (which can be related to a covariance-like measure). Examining this decomposition reveals that  $Q$  increases with abundance evenness and with the number of trait-based or phylogenetic dissimilarities between species, and it increases when the most redundant species are rare and the most abundant ones have the highest trait-based or phylogenetic dissimilarities between them (Fig. 4).

We showed in Section III.1 that the first component, the Gini–Simpson index, can be viewed as a mean of abundance-based rarities and, in Section III.3a, that the second component, the non-weighted mean of the trait-based (or phylogenetic) dissimilarities between species, can be viewed as a mean of trait-based or phylo-based rarities. The third component, the balance component, relates abundance to trait-based or phylogenetic differences. More generally, if we measure the originality of a species as the abundance-weighted mean of the (phylogenetic or trait-based) dissimilarity to all species (including the species itself), then the mean of species' abundance-weighted originalities equals Rao's quadratic entropy (Ricotta *et al.*,

2016). Therefore, the quadratic entropy can be decomposed into independent values of abundance-based rarity and trait-based/phylogenetic originality or into more integrative values of abundance-weighted originalities. In Section III.3c, we also develop a distinct link between quadratic entropy and the notion of originality. We were thus able to extract multiple links between abundance-based rarity, originality and diversity indices from some of the mathematical formulae used to measure these concepts.

(c) *When species' abundances maximize diversity*

To measure the trait-based diversity and phylodiversity of a species assemblage, Pavoine *et al.* (2005) and Pavoine *et al.* (2017) used the quadratic entropy index,  $Q$ , and a related index,  $R$ , with the following formulae:

$$Q = \sum_{i=1}^S \sum_{j=1}^S p_i p_j d_{ij} \quad (1)$$

$$R = \sum_{i=1}^S \sum_{j=1}^S \sqrt{p_i} \sqrt{p_j} d_{ij} \quad (2)$$

where  $p_i$  is the relative abundance of species  $i$  among  $S$  species and  $d_{ij}$  is a measure of phylogenetic or trait-based dissimilarity between species [see Pavoine *et al.*, 2005 for the use of  $Q$  with a restriction on dissimilarities, i.e. use of ultrametric dissimilarities]. As a measure of species' originality, the authors proposed the abundance that species should have in order to maximize the diversity (according to  $Q$  or  $R$ ) of a theoretical assemblage. The associated indices of species' originality, named  $Qb$  and  $Rb$ , thus correspond to the values of  $p_i$  that maximize  $Q$  and  $R$ , respectively, given that the trait-based or phylogenetic dissimilarities cannot vary (for details, see Pavoine *et al.*, 2017). The  $Qb$  and  $Rb$  indices illustrate that, up to a certain degree, trait-based diversity and phylodiversity are high if the most original species in a set (with the highest trait- and phylo-based rarities) have the highest abundance and thus the lowest abundance-based rarity (Fig. 5, Table 2). The indices  $Q$  and  $R$  thus reveal an important property that abundance-weighted indices of trait-based diversity and phylodiversity should have: a positive correlation between species' abundances and species' originalities tends to increase trait-based diversity and phylodiversity.

(d) *When diversity and originality depend on a multidimensional space*

The previously discussed diversity indices either used a tree structure or used phylogenetic or trait-based dissimilarities among species directly. However, several diversity indices were also developed specifically for use in a multidimensional space, where axes reflect traits and points are species' positions according to their traits, as described in Section II.3. The distance of a species' point to the centroid of all points (mean position of all species from the reference set

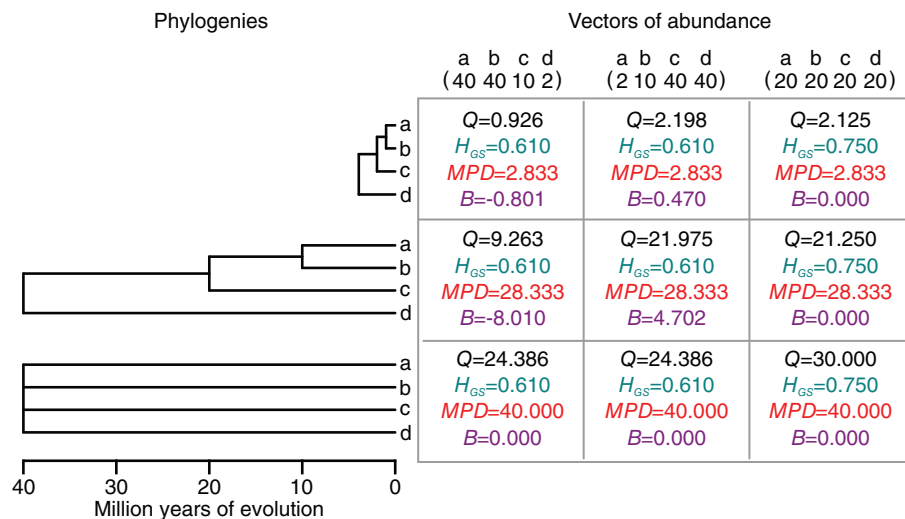
Table 1. Phylodiversity as a mean of phylogenetic rarities

	Function of diversity	Function of rarity <sup>a</sup>
General formula <sup>b</sup>	$I_q = \sum_{k=1}^K T_k \left[ \frac{1 - \sum_{b \in P_k} p_b^q}{q-1} \right] = \sum_{i=1}^S p_i O_q$	$O_q = \left[ \sum_{b \in C(i, \text{Root})} L_b \left( \frac{1 - p_b^{q-1}}{q-1} \right) \right]$
$q = 0^c$	$I_0 = PD - T = \sum_{i=1}^S p_i O_0$	$O_0 = \sum_{b \in C(i, \text{Root})} L_b \left( \frac{1}{p_b} - 1 \right)$
$q \rightarrow 1$	$I_1 = - \sum_{k=1}^K T_k \left[ \sum_{b \in P_k} p_b \ln(p_b) \right] = \sum_{i=1}^S p_i O_1$	$O_1 = \sum_{b \in C(i, \text{Root})} L_b \ln(1/p_b)$
$q = 2$	$I_2 = \sum_{k=1}^K T_k \left( 1 - \sum_{b \in P_k} p_b^2 \right) = \sum_{i=1}^S p_i O_2$	$O_2 = \sum_{b \in C(i, \text{Root})} L_b (1 - p_b)$

<sup>a</sup>The functions  $O_q$ ,  $O_0$ ,  $O_1$  and  $O_2$  represent the abundance-weighted originality of a species.

<sup>b</sup> $q$  is the parameter of the diversity index. It controls the importance given to species lineage abundances (from presence/absence if  $q = 0$  to the overweighting of the most abundant lineage if  $q \rightarrow \infty$ ).  $p_i$  is the relative abundance of species  $i$ .  $b$  is a branch of the phylogenetic tree.  $k$  is a period in the phylogeny,  $T_k$  is its length (time elapsed between two speciation events) and  $P_k$  is the set of branches that cross period  $k$ .  $K$  is the number of periods.  $C(i, \text{Root})$  is the shortest path from species  $i$  (tip) to the root of the tree.  $L_b$  is the length of branch  $b$ .  $p_b$  is the summed relative abundance of all species descending from branch  $b$ .  $S$  is the number of species. See Appendix S2 for a proof of this.

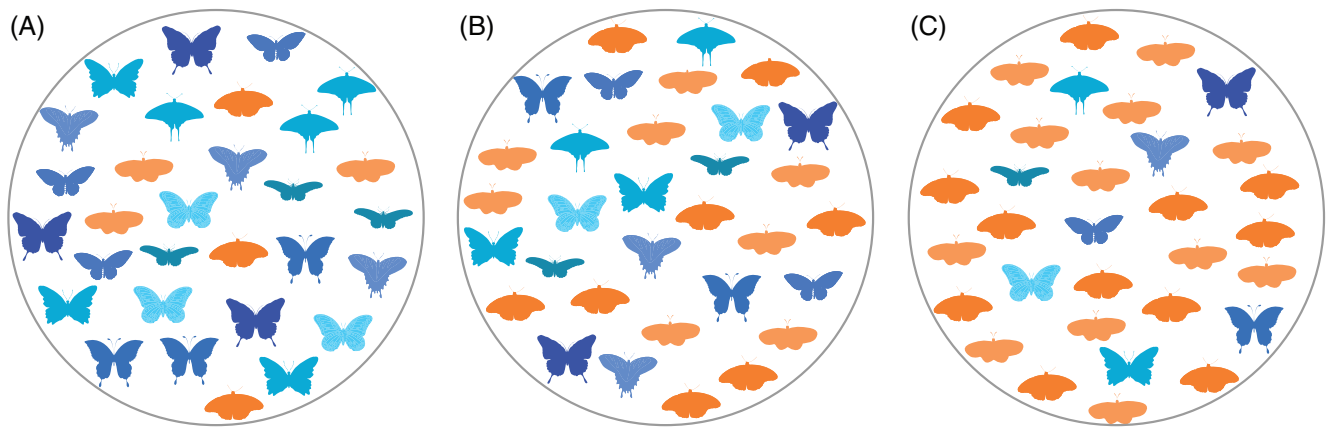
<sup>c</sup>For  $q = 0$ ,  $PD$  is the sum of branch lengths on the phylogenetic tree, and  $T$  is the height of the tree. The function of rarity is linked to the  $AED$  index of abundance-weighted phylogenetic originality in Cadotte *et al.* (2010): if  $p_i = N_i/N$ , with  $N_i$  being the number of individuals of species  $i$ , and  $N$  is the number of individuals across all species, then  $O_0 = AED \times N - T$ .



**Fig. 4.** The quadratic entropy ( $Q$ ) and its main components according to Shimatani (2001). Here, we considered nine theoretical examples formed by crossing three phylogenetic trees with species as tips and three vectors of species abundance. In each example, species are named a, b, c and d. We defined the dissimilarity  $d_{ij}$  between any two species  $i$  and  $j$  as the time since their first common ancestor. For example, in the top tree,  $d_{ab} = 1$  million years. The vectors of abundance give the number of individuals of each species in a community (the  $N_i$  values). We calculated the relative abundance of any species  $i$  as  $p_i = N_i / \sum_{i \in \{a, b, c, d\}} N_i$ . According to Shimatani (2001), the quadratic entropy ( $Q = \sum_{i, j \in \{a, b, c, d\}} p_i p_j d_{ij}$ ) is a function of the Gini–Simpson index of species diversity ( $H_{GS} = 1 - \sum_{i \in \{a, b, c, d\}} p_i^2$ ), the mean dissimilarity between two species ( $MPD = (\sum_{i, j \in \{a, b, c, d\}} d_{ij}) / (4 \times 3)$ ) and a balance component measuring a link between species’ dissimilarities and species’ abundances ( $B = \frac{1}{2} \sum_{i, j \in \{a, b, c, d\}} (d_{ij} - MPD) / [p_i p_j - H_{GS} / (4 \times 3)]$ ):  $Q = H_{GS} * MPD + B$ . The figure shows that  $B$  is negative if the most abundant species are closely related and that it is positive if the most abundant species are also the most dissimilar. It shows that  $B$  is zero in two cases: when the abundances are even and when the dissimilarities between species are even. The top and middle trees have similar topologies, but we multiplied branch lengths by 10 in the middle tree, which allowed us to emphasize that  $Q$ ,  $MPD$  and  $B$  are multiplied by any factor  $X$  (here, 10) if the dissimilarities between species are all multiplied by  $X$ .

in the space) is the main originality measure derived from this method (e.g. Magnuson-Ford *et al.*, 2009). Laliberté & Legendre (2010) proposed using the average distance to the centroid over all species as a measure of trait-based diversity (functional dispersion index,  $FDi_s$ ). They defined two versions of this index: one unweighted (presence/absence data) and

the other weighted by species’ relative abundances. In the unweighted version,  $FDi_s$  values are high if species are all very original, i.e. have the highest distances to the centroid. In the weighted version, they used abundance-weighted distance to the centroid and thus abundance-weighted originality values. In this weighted version,  $FDi_s$  will be high if the most original



**Fig. 5.** Theoretical illustration of the  $Qb$  and  $Rb$  originality measures. Up to a certain degree, increasing the abundance of the most original species increases diversity. As in Fig. 1, we use different drawings of butterflies as theoretical species and differences in colours as indicators of functional/phylogenetic dissimilarities between species. Let us first consider a community in which each species is represented by three individuals (A). This butterfly community appears mostly blue. However, if we increase the abundance (in terms of number of individuals) of orange species, which are more original, then the colour diversity of the new community appears much higher (B). If we increase the abundance of orange species too much, then the butterfly community appears mostly orange, and the diversity decreases (C). The relative originalities of the species do not change from A to C if abundance data are discarded, but the relative abundance-weighted originalities of the orange species decrease, and those of the blue species increase from A to C. In other words, an orange species in C is original (species level), but an orange individual is not (organism level). The indices  $Qb$  and  $Rb$  determine precisely the abundance that species should have to maximize diversity (see Table 2 for an example with the index  $Rb$ ).

Table 2. Illustration of the link between species' originality ( $Rb$ ) and the index of diversity named  $R$  in Pavoine et al. (2017)

Species <sup>a</sup>	Mass <sup>a</sup> (kg)	S1 (%) <sup>b</sup>	S2 (%) <sup>b</sup>	S3 (%) <sup>b</sup>	S4 (%) <sup>b</sup>	S5 (%) <sup>b</sup>
<i>U. arctos</i>	266.50	10	25	49.60	70	85
<i>V. vulpes</i>	5.60	30	25	16.61	10	5
<i>M. erminea</i>	0.90	30	25	16.89	10	5
<i>M. nivalis</i>	0.04	30	25	16.90	10	5
		↓	↓	↓	↓	↓
Body-mass diversity ( $R$ ) <sup>c</sup>		282	402	462	422	328

<sup>a</sup>We considered a theoretical set of four New World terrestrial carnivora species (*Ursus arctos*, *Vulpes vulpes*, *Mustela erminea* and *M. nivalis*) with estimates of each species' body mass obtained from Diniz-Filho & Tôrres (2002).

<sup>b</sup>We considered five case studies corresponding to five scenarios of species' relative abundances, from S1 to S5. S3 corresponds to the species' relative abundances that lead to the highest possible value for  $R$  and thus to the values of species' relative originalities ( $Rb$ ). This table shows that, up to a certain threshold (S3), increasing the abundance of the most different (original) species (here, *U. arctos*) more than that of others increases diversity.

<sup>c</sup>We consider the following formula for the index  $R$  applied to this data set:  $R = \sum_{i=1}^4 \sum_{j=1}^4 \sqrt{p_i} \sqrt{p_j} |M_i - M_j|$ , where  $p_i$  is the relative abundance given to species  $i$  in the set (for example, in terms of number of individuals), and  $M_i$  is the body mass of species  $i$  (in kg).

species have the greatest abundance, again highlighting that an inverse correlation between abundance-based rarity and originality increases diversity.  $FDis$  is closely related to quadratic entropy (their mathematical formulae are very similar, as shown in Pavoine & Bonsall, 2011).

Villéger, Mason, & Mouillot (2008) developed indices to describe three facets of trait-based diversity: functional richness, functional evenness, and functional divergence. None of these indices can be divided into continuous values of species' originality. However, they also implicitly depend on the notion of originality. The first one, functional richness ( $FRic$ ), is the volume of the convex hull of species' points (minimal space that includes all species' points) in the trait-based multidimensional space. From a certain

viewpoint, one can consider that the most original species support the convex hull, with less-original species lying inside the convex hull. The second index, functional evenness ( $FEve$ ), increases if species and their abundances are evenly distributed in the multidimensional space delimited by the most original ones (those at the vertices of the convex hull). The third index, functional divergence ( $FDiv$ ), is linked to the mean distances from species' points to a centroid, but this centroid is measured differently than in the  $FDis$  index developed by Laliberté & Legendre (2010). Villéger, Mason, & Mouillot (2008) considered only the species at the vertices of the convex hull, i.e. from a certain viewpoint, the most original species when finding the centroid coordinates. Their  $FDiv$  index relies on the difference between the

unweighted mean of the distance to the centroid and the abundance-weighted mean of the distance to the centroid. In this index, Villéger, Mason, & Mouillot (2008) did not use species abundance to define the centroid – originality and abundance-based rarity are thus treated separately – while Laliberté & Legendre (2010) did use species abundance to define the centroid of all points, thus obtaining an abundance-weighted originality.

$FDiv$  increases if the most peripheral species have the highest abundances (Fig. 6). Here, we can again observe an inverse link between abundance-based rarity and originality, if we consider the most peripheral species as the most original. However,  $FDiv$  is not an index of biodiversity: we identified that  $FDiv$  can be close to its maximum when a single species dominates in abundance, with the remaining species having residual abundance; for this to happen, this dominant species must be at the largest distance to the centroid of the convex hull (Fig. 6B). Diversity means variety. If intraspecific variations are omitted and if an assemblage is composed of a single species, then any measure of the trait diversity of the assemblage must attain its minimum. Similarly, if intraspecific variations are omitted and if a single species represents almost all the abundance of an assemblage, then, whatever its trait values, the trait diversity of the assemblage is bound to be close to the minimum, whereas in contrast,  $FDiv$  may be close to its maximum.  $FDiv$  thus describes a particular pattern of community functional compositions but is not an index of trait-based diversity.

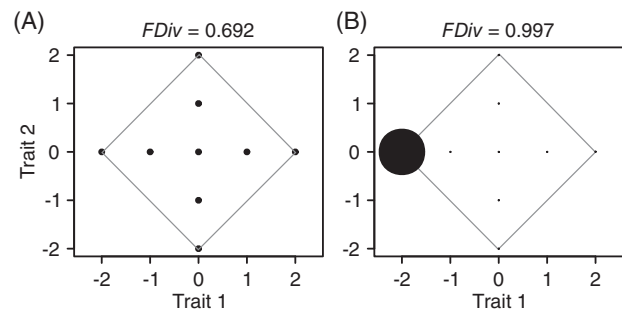
#### IV. RECONCILING THE DIVERSITY, RARITY AND ORIGINALITY CONCEPTS FOR THEIR USEFUL APPLICATIONS

In the previous sections, we have discussed the definitions of the concepts of rarity and originality. We have also demonstrated that fundamental links exist in the definition and in the measurements of the concepts of rarity, originality and diversity. Now we highlight, through many examples, how these concepts can be explored conjointly for the benefits of ecology and evolution.

Every organism is a product of its own individual evolutionary history and is shaped by the environmental conditions and interactions experienced by its ancestors (Cadotte & Davies, 2016). We observed that originality and rarity indices are designed to capture species-specific features, while diversity indices are applied to sets of species. Diversity, rarity and originality thus complement each other in describing the amount of biological variability in species assemblages. Here, we briefly develop some examples of fields where diversity, rarity and originality indices can be applied and underline the importance of their joint analysis.

##### (1) Understanding the evolutionary emergence of species' originality

No approaches have been developed thus far to analyse how trait-based and phylogenetic aspects of originality

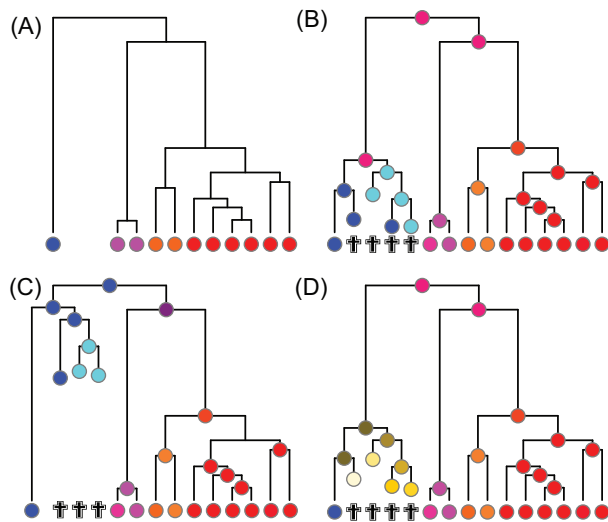


**Fig. 6.** A key property of the functional divergence ( $FDiv$ ) index of Villéger *et al.* (2008) is that it tends to the maximum (equal to 1) when a single original species dominates in abundance. Here, we use the theoretical example of Villéger *et al.* (2008) including a set of nine species (points) characterized by two traits (axes). The traits (positions in the functional space) of the nine species are similar in A and B. Each species is represented by a point, the size of which is proportional to the species' abundance. In A, species have even abundances: the abundance of a species equals  $1/9$ . In B, one of the species dominates in abundance: it has a relative abundance of 0.99, while other species have relative abundances equal to 0.00125 (given that  $0.99 + 0.00125 \times 8 = 1$ ). The convex hull of the set of points is represented by the grey square in the two panels, and the centroid of its vertices is simply its centre. As noted in the two panels, the  $FDiv$  value in case A is lower than that in case B, where the value of  $FDiv$  is close to the maximum (1). These relations hold because the dominant species is one of the original species located on the convex hull of all points, and this dominant species is one with the largest distance to the gravity centre of this convex hull.

are entangled. Often researchers have measured the two aspects independently and have then searched for statistical correlations between them (two-dimensional graph in Fig. 2F). A given species could have evolved, for example, a specific set of characteristics that are individually rare or rare in their combination. Such a species could be both phylogenetically original and original according to its traits, in which case the overall originality reflects the history of a species in terms of trait evolution and species relations [the phylo-trait rarity (PTR) case in Fig. 2F]. The correlation between phylogenetic and trait-based originality is expected if traits have a phylogenetic signal, which means that closely related species tend to share similar values of traits, while distantly related species tend to have different traits.

Inferring that a species is original both in the phylogeny and according to its traits requires observing a particular pattern on a phylogenetic tree (Fig. 7A). However, we must remember that past species extinction could blur the phylogenetic history of a trait state by suppressing many species displaying it (Fig. 7B). This inference also depends on the relevance of our present-day taxonomic sampling (Grandcolas, Nattier, & Trewick, 2014). Rare, distinct traits can occur either on old, persisting branches (Fig. 7C) or on rapidly and recently evolved branches of a given phylogenetic tree (Fig. 7D). The combination of phylogenetic and trait-based originalities of a species is then

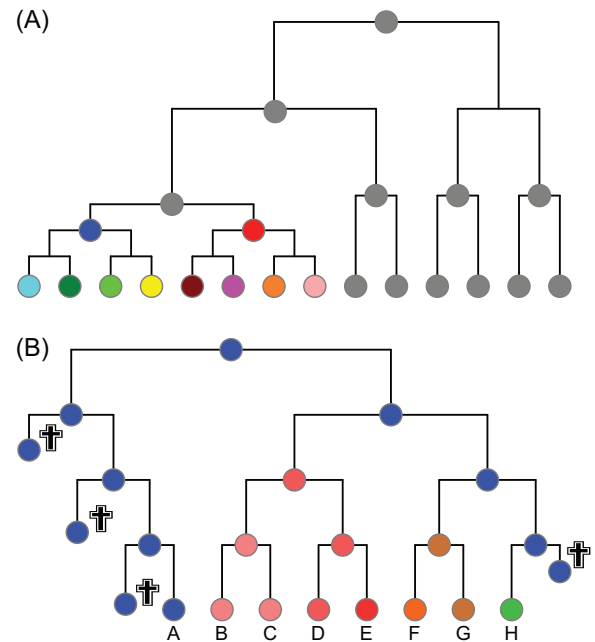




**Fig. 7.** Theoretical examples where a species is original according to both its traits and phylogenetic position. We considered a theoretical phylogenetic tree with 11 extant and up to four extinct species as tips and branch lengths expressed in time of evolution. A shows the tree with only extant species, and B, C and D represent potential scenarios of complete trees with extinct species and trait values for all species and their theoretical ancestors (interior nodes). Circle colours express different values of a theoretical trait (or a complex of traits). (A) Among all extant species, the current species with the blue trait is both the most phylogenetically original, because it is the most isolated, and the most trait-based original, because it is very different in its trait (colour) value from the other extant species. (B) A scenario where past extinction events have suppressed many species displaying the blue trait. (C) A scenario where the species with the blue trait is a relictual species that has kept an ancestral value of the trait (that at the root of the tree). (D) The case where the species with the blue trait has become original among extant species, although its trait value is a recent autapomorphy. In B to D, crosses indicate extinct species.

not explained only by the species' relictiness if it alone remains from a group that is mainly extinct (Grandcolas, Nattier, & Trewick, 2014). Species' originality can emerge from evolutionary autapomorphy, that is, from the appearance in a given taxon of a distinctive derived value of a trait that is unique to this taxon (Fig. 7D).

Species originality may also be used to replace intuitive characterizations that are still too often employed in evolutionary biology, such as 'evolved' or 'primitive' (Crisp & Cook, 2005; Grandcolas & Trewick, 2016). The comparison of a species to a reference species group could benefit from being carried out while considering speciation or extinction rates. For example, adaptive radiations are associated with both an increase in speciation rates and the adaptation of constituent species to a diversity of ecological niches (Gavrillets & Losos, 2009). This adaptation usually leads to high levels of diversity in the trait(s) on which selection for local adaptation acts. As a consequence, adaptation leads to species being original for the trait(s), relative to each other and to the remaining species of the clade. For example, the



**Fig. 8.** Theoretical examples of the potential effects of speciation and extinction events on species' originality and diversity. (A) Adaptive radiation. (B) Phylogenetically structured extinctions. As in Fig. 7, we considered theoretical phylogenetic trees with species as tips and with branch lengths expressed in time of evolution. Circle colours express different values of a theoretical trait (or a complex of traits). Each species and each of their hypothetical ancestors (interior nodes) has a defined value for the trait(s). In A, the ancestor at the root of the phylogeny has a grey colour. All of its descendants kept this trait value up to a given period when adaptive radiation yielded speciation events with rapid trait evolution. In B, the crosses indicate that four species are extinct. Extant species are named A to H. The colours in the circles represent the assumption that some extant species might have acquired traits (pink, orange and green) that render them either tolerant or adapted to new environmental conditions. We used different colours for the acquired traits as there may be multiple ways of being either tolerant or adapted to some specific environmental conditions. This shows that species A, which is phylogenetically distant from the other extant species, is also likely to be functionally original (here by remaining with the ancestral character state), but this may depend on the considered traits.

adaptive radiation in Darwin's finches led to an increase in the diversity of beak size and shape, making these species more original than their relatives (Grant & Grant, 2008). If speciation events occur in a short time period during such a radiation, then trait-based and species diversities are likely to increase by a much larger extent than phylodiversity. In such a case, trait-based originality is not always accompanied by phylogenetic originality (Fig. 8A).

Inversely, high extinction rates in a particular lineage could lead to the phylogenetic isolation of a relict species and thus to its phylogenetic originality among extant species. The trait-based originality of the relict species is likely dependent on the traits considered and on their evolution

(Fig. 8B). An example of a relict species is the present-day ginkgo (*Ginkgo biloba*) (Crisp & Cook, 2011). Phylogenetically clumped extinctions could thus decrease the diversity of a regional pool in terms of species number and phylogeny. At the same time, they could lead to a skewed distribution of phylogenetic originality across the study region, with a few very original species belonging to the species-poor lineages where extinctions occurred. Future research could thus concentrate on searching for how and why phylogenetically and functionally original species emerge in the course of evolution and at which spatial and taxonomic scales they emerge. Additionally, future studies could evaluate how the emergence of original species has influenced current patterns of diversity and distributions of species' abundances.

## (2) Analysing the dynamics of community assembly with rarity, originality and diversity

Species evolve in continuously changing communities governed by numerous ecological, evolutionary and stochastic processes. The ecological processes that have been most studied to explain species diversity are competition, in which biotic interactions between species regulate species' abundances; environmental filtering, where abiotic forces act to constrain certain species' traits within limits; and density dependence, where abundant species have lower individual performance than do rare species (e.g. Hubbell, 2001; Holyoak, Leibold, & Holt, 2005). Mathematical models have been developed in recent decades to synthesize knowledge on community assembly and diversity patterns (Hubbell, 2001; Holyoak, Leibold, & Holt, 2005; Munoz & Huneman, 2016). So far, most of them have focused on measuring and/or predicting patterns of species abundance (number of individuals) and patterns of species richness in communities (MacArthur & Wilson, 1967; Magurran, 2004; McGill *et al.*, 2007). These models have thus rarely focused on patterns of phylo- and trait-based diversity and originality, despite copious numbers of empirical and conceptual studies on community assembly in terms of functional and phylodiversity. A few recent models have, however, started to give insights into functional and phylodiversity (e.g. Münkemüller & Gallien, 2015; Munoz *et al.*, 2018).

There is a broad consensus now on the fact that several distinct processes, including niche-based and neutral ones, interact in species assemblages (e.g. Chase & Myers, 2011). However, despite this consensus, it was also shown that community assembly models based on contrasting underlying processes (e.g. niche-based *versus* neutral processes) may predict the same patterns of species abundance and diversity equally well (e.g. McGill *et al.*, 2007). Such equivalence hampers the direct inference of assembly processes from observed patterns of species abundance and diversity. In fact, mechanistic insights into patterns of community assembly may rely more on the originality of species rather than on their abundance alone (Cadotte & Davies, 2016). One of the oldest identified patterns in ecology is the hollow curve of the distribution of species abundance, with few dominant and many rare species (McGill *et al.*,

2007). The shape of this curve is usually described in terms of skewness. The skewness of an abundance curve is thus inversely related to species diversity, with the latter increasing with an increase in abundance evenness. A hollow curve can also describe the distribution of species' originality in the case of many redundant and few original species (Da Silva *et al.*, 2012). Both species abundance and originality could thus be described by the same tool: a hollow curve.

In parallel, since Webb *et al.* (2002), ecologists often use statistical approaches to associate the trait and/or phylogenetic structure of a community with scenarios of community assembly. This association can be accomplished by comparing observed patterns of trait-based diversity and phylodiversity with those expected by chance using a null model of randomly assembled communities from a regional species pool (Emerson & Gillespie, 2008; Hardy, 2008; Cavender-Bares *et al.*, 2009). Trait-based clustered communities, with many redundant species and low average originality, would reveal environmental filtering, while trait-based overdispersed communities, with the presence of highly original species, would indicate limiting similarity and interspecific competition (Webb *et al.*, 2002). These paradigms may, however, be an oversimplification as competition can sometimes lead to a reduction in functional diversity, particularly when traits are linked to species fitness (Mayfield & Levine, 2010).

The same statements were made for phylodiversity patterns under the assumption of the phylogenetic signal mentioned in Section IV.1 (Webb *et al.*, 2002; see also Saito *et al.*, 2018). Phylogenetic signal tends to decrease at local spatial scales in communities and for narrow taxonomic levels (Cavender-Bares, Keen, & Miles, 2006). Differences between trait-based diversity and phylodiversity patterns for the same community would indicate labile, convergent traits, character displacement between lineages (Gerhold *et al.*, 2015), or environmentally determined traits that vary more than others (Cadotte & Davies, 2016). Species' trait evolution must thus be quantified (see Section IV.1) to connect community trait-based diversity and phylodiversity patterns to the assembly processes (Losos, 2008).

Therefore, to date, the processes that shape community assembly have mostly been analysed based on the concepts of diversity, abundance and thus abundance-based rarity. The roles of trait-based (functional) and phylo-based rarities and thus of species' originality in community assembly have been far less studied. For example, during a colonization event from a regional pool to a local community, the local originality of a colonizing species relative to resident species could be decisive for the successful colonization of the species. The local originality could enhance the colonization success of those species if they can avoid competition with natives for resources due to their high originality ('Darwin's naturalization hypothesis'; Strauss, Webb, & Salamin, 2006; Pearson, Ortega, & Sears, 2012). Inversely, the species' local originality could hamper its colonization success if its high originality makes it less adapted to the local environmental conditions than are the native species (Strayer, 2012).

The local coexistence of species could influence the lineage diversification of the species (Gerhold *et al.*, 2015) and the distribution of the species' originalities. Overall, analyses of community species, phylo- and trait-based diversity, rarity and originality patterns at various spatial scales could therefore better indicate underlying ecological and evolutionary processes that influence species coexistence than analyses of species abundance could alone. A given process could impact only part of the species of an assemblage, which could result in patterns in the distributions of species' originalities. The joint use of rarity, originality and diversity indices could thus provide insights into which interactions among stochastic, ecological, biogeographical and evolutionary processes shape local communities and their dynamics.

### (3) Role of originality in ecosystem functioning

Many studies of biodiversity patterns are focused on common species because they consider common species to represent the largest part of the biomass of a community and thus to be essential for ecosystem functioning (Gaston, 2012). By contrast, rare species may be difficult to include in community-level analyses due to sampling limits. In cases in which knowledge of these species is poor, rare species can be even more difficult to include in trait-based and phylogenetic originality analyses. However, originality and abundance-based rarity were found to be correlated in some cases (e.g. Mi *et al.*, 2012). If low-density species possess the most distinct functional traits, then they could support vulnerable functions in ecosystems and be of particular importance to the community (Gaston, 1998; Mouillot *et al.*, 2013; Leitão *et al.*, 2016). The removal of such species from a community could result in a considerable reduction in ecosystem functioning (Mouillot *et al.*, 2008; Bender *et al.*, 2017), with no possibility for other species to compensate for their loss (Leitão *et al.*, 2016). For example, using species extinction simulations, Rosatti *et al.* (2015) showed that, under extinction scenarios based on abundance and fire tolerance, the probability of losing the most functionally original woody cerrado species was higher than that expected by chance, while the loss of phylogenetically original species was random. In this case study, trait-based functional originality could thus be an indicator of species vulnerability defined based on species rarity and fire sensitivity.

By contrast, two species are functionally redundant if they overlap in their functional niches, i.e. they maintain similar ecosystem functions (Brandl & Bellwood, 2014; Carmona *et al.*, 2016). However, any two species in a set are unlikely to be perfectly redundant but likely to be complementary, i.e. sharing parts of their functional niches and their roles in the ecosystem (Rosenfeld, 2002; Loreau, 2004). Contrary to trait-based originality, trait-based redundancy does not increase trait-based diversity, but it may increase the stability and resilience of communities of species and ecosystem functions. Determining redundant species can thus guide conservation measures. Indeed, local species extinctions caused by perturbations could be compensated for by the

persistence of species that are functionally similar to the lost species but that differ from them in their responses to changes in environmental factors or disturbances (Walker, 1992; Pillar *et al.*, 2013). For example, an increase in trait-based functional originality has been found in coral reef communities after a cyclone disturbance due to the local extinction of redundant species that had trait values similar to those of the surviving species (Brandl *et al.*, 2016). Ecosystem functioning, stability and resilience are dependent on the composition of species and those species' abundance-, phylo- and trait-based rarities. Inversely, species' characters, abundances and distribution patterns are shaped by many ecological processes. We should continue to consider this mutuality in future studies to better comprehend the complexity of ecosystem functioning.

### (4) Guiding conservation actions

Species are among the key units of biodiversity measurement. Their conservation is at the core of many national and international programmes that request effective methods for habitat prioritization and species preservation. Criteria related to ecosystem services evaluated as economic costs, aesthetic value, contribution to well-being, and species richness and rarity are often parts of these programmes. Should such programmes also include phylo- and trait-based diversity and originality? Even if conservation planning alternates between preserving particular units of biodiversity and preserving the processes that shaped those units, the most commonly employed method to design conservational priorities uses species richness ranked by species' abundance-based rarity and endemism (Mace, 2003). Conservation values are thus often given to geographical units, or biodiversity hotspots, that are not chosen with regard to species' traits and evolutionary histories (Veron *et al.*, 2017).

For example, Brum *et al.* (2017) reviewed currently protected areas determined by the International Union for Conservation of Nature (IUCN) risk classification system. They demonstrated that those areas do not harbour more phylo- or trait-based diversities of threatened mammals than would be expected if they were randomly selected. As shown in the previous sections, species are not equivalent, and the phylo- and trait-based diversities of an area are the products of numerous stochastic, evolutionary and ecological processes. Conservation planning would thus benefit from considering biodiversity as multidimensional by including the phylogenetic and trait-based originalities of species (Rodrigues & Gaston, 2002; Pellens, Faith, & Grandcolas, 2016). For example, Pollock, Thuiller, & Jetz (2017) showed that a 5% expansion of protected areas could more than triple the protected range of species or trait-based or phylo-based units.

Originality indices could be useful as a complement for conservation actions that target species rather than areas (Pavoine *et al.*, 2005; Isaac *et al.*, 2007; Jetz *et al.*, 2014; Laity *et al.*, 2015; Grenié *et al.*, 2018). Species are often ranked for conservation attention by their patrimonial,

abundance-based rarity and threat status, which is not sufficient to determine priorities (Mouillot *et al.*, 2008). These conservation attentions could nevertheless indirectly (non-intentionally) preserve original species. For example, Thévenin *et al.* (2018) showed that, although evolutionary considerations are unlikely to have driven explicitly the allocation of reintroduction efforts, reintroduced birds and mammals in Europe, North America and Central America tend to be more phylogenetically original than expected by chance.

Originality indices allow treating species within multiple facets of biological diversity. The explicit inclusion of such indices in conservation actions would be useful to evaluate each species' contribution to local and global phylo- and trait-based diversity (Jensen *et al.*, 2016; Pavoine *et al.*, 2017). The recommendation to safeguard species that are both original and threatened inspired some authors to extend pre-existing indices of originality to include additional species attributes: abundance, range size and probability of extinction (Isaac *et al.*, 2007; Rosauer *et al.*, 2009; Cadotte & Davies, 2010; Hidasi-Neto *et al.*, 2015). Although alternatives exist (Steel *et al.*, 2007; Jensen *et al.*, 2016), the most widely cited originality measure that also integrates species extinction risk is the evolutionary distinct and globally endangered (EDGE) species index (Isaac *et al.*, 2007). The EDGE index combines the phylogenetic originality of a focal species with its IUCN threat status. Given that it was designed for large-scale analyses, the EDGE index does not consider species' traits. However, conservation actions are often developed locally. Hidasi-Neto *et al.* (2015) thus modified the EDGE index and proposed the EcoEDGE index, which combines phylogenetic and functional components of species' originalities with those species' extinction risks.

Some studies found that low abundance and narrow range size are the characteristics of functionally original species that are threatened to extinction at the local scale in French breeding birds (Calba, Maris, & Devictor, 2014) and freshwater bivalve molluscs (Burlakova *et al.*, 2011) and at both local and regional scales in coral reef fishes, alpine plants and tropical trees (Mouillot *et al.*, 2013). At a global scale, phylogenetically original primate species were found to be more threatened with extinction than were other primates (Verde Arregoitia, Blomberg & Fisher, 2013; but see Redding, deWolff, & Mooers, 2010), but this was not true for the whole class of mammals (Verde Arregoitia *et al.*, 2013). The endangerment of original species could, however, lead to a drastic decrease in phylo- and trait-based diversity if it were also associated with the phylogenetic clustering of species extinction (Parhar & Mooers, 2011).

Therefore, the spatial distribution of original species may provide information on whether they are concentrated in low-diversity areas that are not targeted by current conservation actions (e.g. Jetz *et al.*, 2014; Veron, Clergeau & Pavoine, 2016). Additionally, using simulations, Redding *et al.* (2008) showed that prioritizing species by different originality indices measured globally tends to safeguard more local phylodiversity than is expected by the selection

of random species. Similarly, prioritizing Neotropical and Nearctic bird species with high average global phylogenetic originality scores could allow local phylodiversity to be safeguarded (Redding *et al.*, 2015). Clearly, species' originality identifies 'key' species for preservation priority that could be overlooked by classical abundance-based rarity and diversity methods.

## V. CONCLUSIONS

(1) Due to the emergence of myriad terms related to biodiversity, rarity, and originality in the ecological literature, it has become difficult to determine whether terms used by different studies refer to the same concepts and measures. Each concept encompasses different aspects of species assemblages and can be measured by several mathematical indices. We provided a semantic and historical overview of these three concepts – biodiversity, rarity and originality – at the level of an assemblage of species. We showed that mathematical links exist between their associated indices.

(2) Historically, rarity was explicitly integrated into diversity measures by means of species abundance. Later, diversity measures incorporated species' biological (trait or phylogenetic) differences, sometimes weighted by the species' abundances. Species identities can be interchanged without affecting diversity measurement based on the number and abundances of species. Phylo- and trait-based diversities have thus gone far beyond the vision of species diversity.

(3) The contribution of individual species to the phylo- and trait-based diversities of a reference species set is captured by the concept of originality. This concept brings together ecologists and evolutionary biologists because it can combine species' evolutionary histories, traits and abundances. It can be measured by various approaches based on phylogenetic trees and trait-based dendrograms, dissimilarity matrices and multidimensional space. Focused at the individual species level, originality complements the diversity measures that can then sometimes be written as simple functions of species' trait-, phylogeny-, and abundance-based rarities.

(4) The joint use of the concepts of diversity, rarity, and originality could aid in the understanding of the multiple mechanisms shaping communities at different spatial and temporal scales. At a large scale, the joint use of these concepts could help clarify general patterns of evolutionary events such as trait evolution, extinction, speciation, and adaptive radiation of species. At a local scale, it could aid in understanding community assembly and ecosystem functioning. At any scale, it could refine conservation strategies.

(5) It is widely accepted that no single mathematical formula could alone encompass all aspects of biodiversity. Here, we have shown that the joint use of diversity, rarity and originality measures has the potential to recompose accurately the complex picture of the diversity of a species assemblage.



## VI. ACKNOWLEDGEMENTS

A.K. and S.P. conceived of the overarching ideas and wrote the manuscript; P.G. contributed to idea development and manuscript editing. A.K., P.G., and S.P. acknowledge the support of the French State through the Agence Nationale de la Recherche under the reference project 10-LABX-0003, within the framework of the programme 'Investing for the future' (11-IDEX-0004), and thank the anonymous referees for their very constructive comments and Simon Veron for his useful comments on an earlier version of the paper.

## VII. REFERENCES

References marked with asterisk have been cited within the supporting information.

- \*ALTSCHUL, S. F. & LIPMAN, D. J. (1990). Equal Animals. *Nature* **348**, 493–494.
- BENDER, I. M. A., KISSLING, W. D., BÖHNING-GAESE, K., HENSEN, I., KÜHN, I., WIEGAND, T., DEHLING, D. M. & SCHLEUNING, M. (2017). Functionally specialised birds respond flexibly to seasonal changes in fruit availability. *Journal of Animal Ecology* **86**, 800–811.
- BRANDL, S. J. & BELLWOOD, D. R. (2014). Individual-based analyses reveal limited functional overlap in a coral reef fish community. *Journal of Animal Ecology* **83**, 661–670.
- BRANDL, S. J., EMSLIE, M. J., CECCARELLI, D. M. & RICHARDS, Z. T. (2016). Habitat degradation increases functional originality in highly diverse coral reef fish assemblages. *Ecosphere* **7**, e01557.
- BROWN, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist* **124**, 255–279.
- BRUM, F. T., GRAHAM, C. H., COSTA, G. C., HEDGES, S. B., PENONE, C., RADELOFF, V. C., RONDININI, C., LOYOLA, R. & DAVIDSON, A. D. (2017). Global priorities for conservation across multiple dimensions of mammalian diversity. *Proceedings of the National Academy of Sciences* **114**, 7641–7646.
- 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 Biology* **19**, 387–400.
- BURLAKOVA, L. E., KARATAYEV, A. Y., KARATAYEV, V. A., MAY, M. E., BENNETT, D. L. & COOK, M. J. (2011). Endemic species: contribution to community uniqueness, effect of habitat alteration, and conservation priorities. *Biological Conservation* **144**, 155–165.
- CADOTTE, M. W. & DAVIES, T. J. (2010). Rarest of the rare: advances in combining evolutionary distinctiveness and scarcity to inform conservation at biogeographical scales. *Diversity and Distributions* **16**, 376–385.
- CADOTTE, M. W. & DAVIES, T. J. (2016). *Phylogenies in Ecology: A Guide to Concepts and Methods*. Princeton University Press, Princeton.
- CADOTTE, M. W., DAVIES, T. J., REGETZ, J., KEMBEL, S. W., CLELAND, E. & OAKLEY, T. H. (2010). Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. *Ecology Letters* **13**, 96–105.
- CALBA, S., MARIS, V. & DEVICTOR, V. (2014). Measuring and explaining large-scale distribution of functional and phylogenetic diversity in birds: separating ecological drivers from methodological choices. *Global Ecology and Biogeography* **23**, 669–678.
- CARMONA, C. P., DE BELLO, F., MASON, N. W. H. & LEPS, J. (2016). Traits without borders: integrating functional diversity across scales. *Trends in Ecology and Evolution* **31**, 382–394.
- CAVENDER-BARES, J., KEEN, A. & MILES, B. (2006). Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* **87**, S109–S122.
- CAVENDER-BARES, J., KOZAK, K., FINE, P. V. A. & KEMBEL, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters* **12**, 693–715.
- CHASE, J. M. & MYERS, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 2351–2363.
- \*CHIU, C.-H., JOST, L. & CHAO, A. (2014). Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs* **84**, 21–44.
- CRISP, M. D. & COOK, L. G. (2005). Do early branching lineages signify ancestral traits? *Trends in Ecology & Evolution* **20**, 122–128.
- CRISP, M. D. & COOK, L. G. (2011). Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. *New Phytologist* **192**, 997–1009.
- DA SILVA, D. M., SILVA, I. A. & BATALHA, M. A. (2012). Phylogenetic and phenotypic originality and abundance in a cerrado plant community. *Austral Ecology* **37**, 302–307.
- DARÓCZY, Z. (1970). Generalized information functions. *Information and Control* **16**, 36–51.
- DE BELLO, F., LAVOREL, S., ALBERT, C. H., THULLER, W., GRIGULIS, K., DOLEZAL, J., JANEČEK, Š. & LEPS, J. (2011). Quantifying the relevance of intraspecific trait variability for functional diversity. *Methods in Ecology and Evolution* **2**, 163–174.
- DELIĆ, T., TRONTELJ, P., RENDOŠ, M. & FIŠER, C. (2017). The importance of naming cryptic species and the conservation of endemic subterranean amphipods. *Scientific Reports* **7**, 3391.
- DINIZ-FILHO, J. A. F. & TÓRRES, N. M. (2002). Phylogenetic comparative methods and the geographic range size-body size relationship in new world terrestrial Carnivora. *Evolutionary Ecology* **16**, 351–367.
- \*EISWERTH, M. E. & HANEY, J. C. (1992). Allocating conservation expenditures: accounting for inter-species genetic distinctiveness. *Ecological Economics* **5**, 235–249.
- EMERSON, B. C. & GILLESPIE, R. G. (2008). Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution* **23**, 619–630.
- FAITH, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**, 1–10.
- FINDLEY, J. S. (1973). Phenetic packing as a measure of faunal diversity. *The American Naturalist* **107**, 580–584.
- FINDLEY, J. S. (1976). The structure of bat communities. *The American Naturalist* **110**, 129–139.
- FRISHKOFF, L. O., KARP, D. S., M'GONIGLE, L. K., MENDENHALL, C. D., ZOOK, J., KREMEN, C., HADLY, E. A. & DAILY, G. C. (2014). Loss of avian phylogenetic diversity in neotropical agricultural systems. *Science* **345**, 1343–1346.
- GASTON, K. J. (1991). How large is a species' geographic range? *Oikos* **61**, 434–438.
- GASTON, K. J. (1996). The multiple forms of the interspecific abundance-distribution relationship. *Oikos* **76**, 211–220.
- GASTON, K. J. (1997). What is rarity? In *The Biology of Rarity* (eds W. E. KUNIN and K. J. GASTON), pp. 30–47. Chapman & Hall, London.
- GASTON, K. J. (1998). Rarity as double jeopardy. *Nature* **394**, 229–230.
- GASTON, K. J. (2012). The importance of being rare. *Nature* **487**, 46–47.
- GASTON, K. J. & SPICER, J. I. (2004). *Biodiversity: An Introduction*. Blackwell Publishing, Oxford.
- GAVRILETS, S. & LOSOS, J. B. (2009). Adaptive radiation: contrasting theory with data. *Science* **323**, 732–737.
- 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). *Functional Ecology* **29**, 600–614.
- GINI, C. (1912). *Variabilità e Mutabilità*. Studi economica giuridici delle facoltà di giurisprudenza dell'Università di Cagliari III, Parte II.
- GRANDCOLAS, P. (1998). Phylogenetic analysis and the study of community structure. *Oikos* **82**, 397–400.
- GRANDCOLAS, P., NATTIER, R. & TREWICK, S. A. (2014). Relict species: a relict concept? *Trends in Ecology & Evolution* **29**, 655–663.
- GRANDCOLAS, P. & TREWICK, S. A. (2016). What is the meaning of extreme phylogenetic diversity? The case of phylogenetic relict species. In *Biodiversity Conservation and Phylogenetic Systematics: Preserving our Evolutionary Heritage in an Extinction Crisis* (eds R. PELLENS and P. GRANDCOLAS), pp. 99–115. Springer Open, Cham, Switzerland.
- GRANT, P. R. & GRANT, R. (2008). *How and why Species Multiply? The Radiation of Darwin's Finches*. Princeton University Press, Princeton, NJ.
- GRENIÉ, M., MOUILLOT, D., VILLÉGER, S., DENELLE, P., TUCKER, C. M., MUNOZ, F. & VIOLLE, C. (2018). Functional rarity of coral reef fishes at the global scale: hotspots and challenges for conservation. *Biological Conservation* **226**, 288–299.
- GROVES, C., COTTERILL, F. P. D., GIPPOLITI, S., ROBOVSKÝ, J., ROOS, C., TAYLOR, P. J. & ZINNER, D. (2017). Species definitions and conservation: a review and case studies from African mammals. *Conservation Genetics* **18**, 1247–1256.
- HARDY, O. J. (2008). Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *Journal of Ecology* **96**, 914–926.
- HARDY, O. J. & SENTERRE, B. (2007). Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. *Journal of Ecology* **95**, 493–506.
- HARTLEY, S. & KUNIN, W. E. (2003). Scale dependency of rarity, extinction risk, and conservation priority. *Conservation Biology* **17**, 1559–1570.
- HAVRDA, M. & CHARVAT, F. (1967). Quantification method of classification processes: concept of structural a-entropy. *Kybernetika* **3**, 30–35.
- HELMUS, M. R., BLAND, T. J., WILLIAMS, C. K. & IVES, A. R. (2007). Phylogenetic measures of biodiversity. *The American Naturalist* **169**, E68–E83.
- HESSEN, D. O. & WALSENG, B. (2008). The rarity concept and the commonness of rarity in freshwater zooplankton. *Freshwater Biology* **53**, 2026–2035.
- HIDASÍ-NEITO, J., LOYOLA, R. & CIANCIARUSO, M. V. (2015). Global and local evolutionary and ecological distinctiveness of terrestrial mammals: identifying priorities across scales. *Diversity and Distributions* **21**, 548–559.
- HOLYOAK, M., LEIBOLD, M. A. & HOLT, R. D. (2005). *Competing Theories for Competitive Metacommunities*. The University of Chicago Press, Chicago.

- HOOPER, D. U., SOLAN, M., SYMSTAD, A., DÍAZ, S., GESSNER, M. O., BUCHMANN, N., DEGRANGE, V., GRIME, P., HULOT, F., MERMILLOD-BLONDIN, F., ROY, J., SPEHN, E. & VAN PEER, L. (2002). Species diversity, functional diversity and ecosystem functioning. In *Biodiversity and Ecosystem Functioning: Syntheses and Perspectives* (eds M. LOREAU, S. NAEEM and P. INCHAUSTI), pp. 195–208. Oxford University Press, Oxford.
- HUANG, J., MI, X. & MA, K. (2011). A genome evolution-based framework for measures of originality for clades. *Journal of Theoretical Biology* **276**, 99–105.
- HUBBELL, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton.
- ISAAC, N. J. B., TURVEY, S. T., COLLEN, B., WATERMAN, C. & BAILLIE, J. E. M. (2007). Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS One* **2**, e296.
- JARZYNA, M. A. & JETZ, W. (2016). Detecting the multiple facets of biodiversity. *Trends in Ecology and Evolution* **31**, 527–538.
- JENSEN, E. L., MOOERS, A. Ø., CACCONE, A. & RUSSELLO, M. A. (2016). I-HEDGE: determining the optimum complementary 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. *Current Biology* **24**, 919–930.
- JOHNSON, C. N. (1998). Species extinction and the relationship between distribution and abundance. *Nature* **394**, 272–274.
- KELLY, S., GRENYER, R. & SCOTLAND, R. W. (2014). Phylogenetic trees do not reliably predict feature diversity. *Diversity and Distributions* **20**, 600–612.
- KRAFT, N., CORNWELL, W., WEBB, C. & ACKERLY, D. (2007). Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist* **170**, 271–283.
- KUNIN, W. E. & GASTON, K. J. (1993). The biology of rarity: patterns, causes and consequences. *Trends in Ecology and Evolution* **8**, 298–301.
- LAITY, T., LAFFAN, S. W., GONZÁLEZ-OROZCO, C. E., FAITH, D. P., ROSAUER, D. F., BYRNE, M., MILLER, J. T., CRAYN, D., COSTION, C., MORITZ, C. C. & NEWPORT, K. (2015). Phylodiversity to inform conservation policy: an Australian example. *Science of the Total Environment* **534**, 131–143.
- LALIBERTÉ, E. & LEGENDRE, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **91**, 299–305.
- LANDE, R. (1996). Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* **76**, 5–13.
- LAVAL, G., IANNUCELLI, N., LEGAULT, C., MILAN, D., GROENEN, M. A. M., GIUFFRÀ, E., ANDERSSON, L., NISSEN, P. H., JØRGENSEN, C. B., BECKMANN, P., GELDERMANN, H., FOULLEY, J.-L., CHEVALET, C. & OLLIVIER, L. (2000). Genetic diversity of eleven European pig breeds. *Genetics Selection Evolution* **32**, 187–203.
- LAVOREL, S. & GARNIER, E. (2002). Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the holy grail. *Functional Ecology* **16**, 545–556.
- LEITÃO, R. P., ZUANON, J., VILLÉGER, S., WILLIAMS, S. E., BARALOTO, C., FORTUNEL, C., MENDONÇA, F. P. & MOUILLOT, D. (2016). Rare species contribute disproportionately to the functional structure of species assemblages. *Proceedings of the Royal Society B* **283**, 20160084.
- LOREAU, M. (2004). Does functional redundancy exist? *Oikos* **104**, 606–611.
- LOSOS, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters* **11**, 995–1007.
- LYONS, N. I. (1981). Comparing diversity indices based on counts weighted by biomass or other importance values. *American Naturalist* **118**, 438–442.
- MACARTHUR, R. H. & WILSON, E. O. (1967). *The Theory of Island Biogeography*. Princeton University Press, Princeton.
- MACE, G. M. (2003). Preserving the tree of life. *Science* **300**, 1707–1709.
- 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.
- MAGURRAN, A. E. (2004). *Measuring Biological Diversity*. Blackwell Publishing, Oxford.
- MAIRE, E., GRENOUILLET, G., BROSSE, S. & VILLÉGER, S. (2015). How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. *Global Ecology and Biogeography* **27**, 728–740.
- MASON, N. W., MOUILLOT, D., LEE, W. G. & WILSON, J. B. (2005). Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos* **111**, 112–118.
- MAY, R. M. (1990). Taxonomy as destiny. *Nature* **347**, 129–130.
- MAYFIELD, M. M. & LEVINE, J. M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters* **13**, 1085–1093.
- MAZEL, F., DAVIES, T. J., GALLIEN, L., RENAUD, J., GROUSSIN, M., MÜNKMÜLLER, T. & THUILLER, W. (2016). Influence of tree shape and evolutionary time-scale on phylogenetic diversity metrics. *Ecography* **39**, 913–920.
- MAZEL, F., MOOERS, A., VALENTINO, G., RIVA, D. & PENNELL, M. W. (2017). Conserving phylogenetic diversity can be a poor strategy for conserving functional diversity. *Systematic Biology* **66**, 1019–1027.
- MCGILL, B. J., RAMPAL, S. E., GRAY, J. S., ALONSO, D., ANDERSON, M. J., KASSA BENECHA, H., DORNELAS, M., ENQUIST, B. J., GREEN, J. L., HE, F., HURLBERT, A. H., MAGURRAN, A. E., MARQUET, P. A., MAURER, B. A., OSTLING, A., et al. (2007). Species abundance distributions: moving beyond simple prediction theories to integration within ecological framework. *Ecology Letters* **10**, 995–1015.
- MERGEAY, J. & SANTAMARIA, L. (2012). Evolution and biodiversity: the evolutionary basis of biodiversity and its potential for adaptation to global change. *Evolutionary Applications* **5**, 103–106.
- MI, X., SWENSON, N. G., VALENCIA, R., KRESS, W. J., ERICKSON, D. L., PÉREZ, Á. J., REN, H., SU, S. H., GUNATILLEKE, N., GUNATILLEKE, S., HAO, Z., YÉ, W., CAO, M., SURESH, H. S., DATTAARAJA, H. S., SUKUMAR, R. & MA, K. (2012). The contribution of rare species to community phylogenetic diversity across a global network of forest plots. *The American Naturalist* **180**, E17–E30.
- MORELLI, F., BENEDETTI, Y., PERNA, P. & SANTOLINI, R. (2018). Associations among taxonomic diversity, functional diversity and evolutionary distinctiveness vary among environments. *Ecological Indicators* **88**, 8–16.
- 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., 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 Biology* **11**, e1001569.
- MOUILLOT, D., CULIOLI, J. M., PELLETIER, D. & TOMASINI, J. A. (2008). Do we protect biological originality in protected areas? A new index and an application to the Bonifacio Strait natural reserve. *Biological Conservation* **141**, 1569–1580.
- MOUQUET, N., DEVICTOR, V., MEYNARD, C. N., MUNOZ, F., BERSIER, L.-F., CHAVE, J., COUTERON, P., DALECKY, A., FONTAINE, C., GRAVEL, D., HARDY, O. J., JABOT, F., LAVERGNE, S., LEIBOLD, M. S., MOUILLOT, D., et al. (2012). Ecophylogenetics: advances and perspectives. *Biological Reviews* **87**, 769–785.
- MÜNKMÜLLER, T. & GALLIEN, L. (2015). VirtualCom: a simulation model for eco-evolutionary community assembly and invasion. *Methods in Ecology and Evolution* **6**, 735–743.
- MUNOZ, F., GRENIÉ, M., DENELLE, P., TAUDIÈRE, A., LAROCHE, F., TUCKER, C. & VIOLLE, C. (2018). Ecolottery: simulating and assessing community assembly with environmental filtering and neutral dynamics in R. *Methods in Ecology and Evolution* **9**, 693–703.
- MUNOZ, F. & HUNEMAN, P. (2016). From neutral theory to a comprehensive and multiscale theory of ecological equivalence. *The Quarterly Review of Biology* **91**, 321–342.
- MYERS, N., MITTERMEIER, R. A., MITTERMEIER, C. G., FONSECA, G. A. B. & KENT, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858.
- NIXON, K. C. & WHEELER, Q. D. (1992). Measures of phylogenetic diversity. In *Extinction and Phylogeny* (eds M. J. NOVACEK and Q. D. WHEELER), pp. 216–234. Columbia University Press, New York.
- PARHAR, R. K. & MOOERS, A. Ø. (2011). Phylogenetically clustered extinction risks do not substantially prune the tree of life. *PLoS One* **6**, e23528.
- PATIL, G. P. & TAILLIE, C. (1982). Diversity as a concept and its measurement. *Journal of the American Statistical Association* **77**, 548–561.
- PAVOINE, S. & BONSALE, M. B. (2011). Measuring biodiversity to explain community assembly: a unified approach. *Biological Reviews* **86**, 792–812.
- PAVOINE, S., BONSALE, M. B., DUPAIX, A., JACOB, U. & RICOTTA, C. (2017). From phylogenetic to functional originality: guide through indices and new developments. *Ecological Indicators* **82**, 196–205.
- PAVOINE, S., LOVE, M. & BONSALE, M. (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 biogeographic realm. *Ecology Letters* **12**, 898–908.
- PAVOINE, S., OLLIER, S. & DUFOUR, A. B. (2005). Is the originality of a species measurable? *Ecology Letters* **8**, 579–586.
- PAVOINE, S., VALLET, J., DUFOUR, A.-B., GACHET, S. & DANIEL, H. (2009). On the challenge of treating various types of variables: application for improving the measurement of functional diversity. *Oikos* **118**, 391–402.
- PEARSON, D. E., ORTEGA, Y. K. & SEARS, S. J. (2012). Darwin's naturalization hypothesis up-close: intermountain grassland invaders differ morphologically and phenologically from native community dominants. *Biological Invasions* **14**, 901–913.
- PELLENS, R., FAITH, D. P. & GRANDCOLAS, P. (2016). The future of phylogenetic systematics in conservation biology: linking biodiversity and society. In *Biodiversity Conservation and Phylogenetic Systematics: Preserving our Evolutionary Heritage in an Extinction Crisis* (eds R. PELLENS and P. GRANDCOLAS), pp. 375–383. Springer Open, Cham, Switzerland.
- PETCHEY, O. L. & GASTON, K. J. (2002). Functional diversity (FD), species richness and community composition. *Ecology Letters* **5**, 402–411.
- PIGOT, A. L., BREGMAN, T., SHEARD, C., DALY, B., ETIENNE, R. S. & TOBIAS, J. A. (2016). Quantifying species contributions to ecosystem processes: a global assessment of functional trait and phylogenetic metrics across avian seed-dispersal networks. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20161597.

- PILLAR, V. D., BLANCO, C. C., MÜLLER, S. C., SOSINSKI, E. E., JONER, F. & DUARTE, L. D. S. (2013). Functional redundancy and stability in plant communities. *Journal of Vegetation Science* **24**, 963–974.
- POLLOCK, L. J., THULLER, W. & JETZ, W. (2017). Large conservation gains possible for global biodiversity facets. *Nature* **546**, 141–144.
- RABINOWITZ, D. (1981). Seven forms of rarity. In *The Biological Aspects of Rare Plant Conservation* (ed. H. SYNGE), pp. 205–217. Wiley, Chichester.
- RAO, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology* **21**, 24–43.
- REDDING, D. W. (2003). *Incorporating genetic distinctness and reserve occupancy into a conservation prioritisation approach*. Master Thesis: University of East Anglia, Norwich.
- REDDING, D. W., DEWOLFF, C. U. R. T. & MOOERS, A. Ø. (2010). Evolutionary distinctiveness, threat status, and ecological oddity in primates. *Conservation Biology* **24**, 1052–1058.
- REDDING, D. W., HARTMANN, K., MIMOTO, A., BOKAL, D., DEVOS, M. & MOOERS, A. Ø. (2008). Evolutionarily distinctive species often capture more phylogenetic diversity than expected. *Journal of Theoretical Biology* **251**, 606–615.
- REDDING, D. W., MAZEL, F. & MOOERS, A. Ø. (2014). Measuring evolutionary isolation for conservation. *PLoS One* **9**, e113490.
- REDDING, D. W. & MOOERS, A. Ø. (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology* **20**, 1670–1678.
- REDDING, D. W., MOOERS, A. Ø., ŞEKERCIOĞLU, Ç. H. & COLLEN, B. (2015). Global evolutionary isolation measures can capture key local conservation species in Nearctic and Neotropical bird communities. *Philosophical Transactions of the Royal Society B-Biological Sciences* **370**, 20140013.
- RICOTTA, C. (2004). A parametric diversity measure combining the relative abundance and taxonomic distinctiveness of species. *Diversity and Distribution* **10**, 143–146.
- RICOTTA, C. (2007). A semantic taxonomy for diversity measures. *Acta Biotheoretica* **55**, 23–33.
- 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 in Ecology and Evolution* **7**, 1386–1395.
- RODRIGUES, A. S. & GASTON, K. J. (2002). Maximising phylogenetic diversity in the selection of networks of conservation areas. *Biological Conservation* **105**, 103–111.
- 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 Ecology* **40**, 267–274.
- ROSAUER, D., 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. *Molecular Ecology* **18**, 4061–4072.
- ROSENFELD, J. S. (2002). Functional redundancy in ecology and conservation. *Oikos* **98**, 156–162.
- SAFI, K., ARMOUR-MARSHALL, K., BAILLIE, J. E. M. & ISAAC, N. J. B. (2013). Global patterns of evolutionary distinct and globally endangered amphibians and mammals. *PLoS One* **8**, e63582.
- SAITO, V., LAROCHE, F., SQUIERA, T. & PAVOINE, S. (2018). Ecological versatility and the assembly of multiple competitors: cautionary notes for assembly inferences. *Ecology* **99**, 1173–1183.
- SCHNEIDER, S. M. (2012). A metric of biodiversity that integrates abundance, phylogeny, and function. *Oikos* **121**, 1191–1202.
- SCHNEIDER, S. M., KOSMAN, E., PRESLEY, S. J. & WILLIG, M. R. (2017). The components of biodiversity, with a particular focus on phylogenetic information. *Ecology and Evolution* **7**, 6444–6454.
- SCHMERA, D., ERŐS, T. & PODANI, J. (2009). A measure for assessing functional diversity in ecological communities. *Aquatic Ecology* **43**, 157–167.
- SHANNON, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* **27**, 379, 623–423, 656.
- SHIMATANI, K. (2001). On the measurement of species diversity incorporating species differences. *Oikos* **93**, 135–147.
- SIMPSON, E. H. (1949). Measurement of diversity. *Nature* **163**, 688.
- SOBRAL, F. L., LEES, A. C. & CIANGIARUSO, M. V. (2016). Introductions do not compensate for functional and phylogenetic losses following extinctions in insular bird assemblages. *Ecology Letters* **19**, 1091–1100.
- STEEL, M., MIMOTO, A. & MOOERS, A. Ø. (2007). Hedging our bets: the expected contribution of species to future phylogenetic diversity. *Evolutionary Bioinformatics* **3**, 237–244.
- STRAUSS, S. Y., WEBB, C. O. & SALAMIN, N. (2006). Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Science United States of America* **103**, 5841–5845.
- STRAYER, D. L. (2012). Eight questions about invasions and ecosystem functioning. *Ecology Letters* **15**, 1199–1210.
- THAON D'ARNOLDI, C., FOULLEY, J.-L. & OLLIVIER, L. (1998). An overview of the Weitzman approach to diversity. *Genetics Selection Evolution* **30**, 149–161.
- THÉVENIN, C., MOUCHET, M., ROBERT, A., KERBIRIOU, C. & SARRAZIN, F. (2018). Reintroductions of birds and mammals involve evolutionarily distinct species at the regional scale. *Proceedings of the National Academy of Science USA* **115**, 3404–3409.
- TSALLIS, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* **52**, 480–487.
- TUCKER, C. M., CADOTTE, M. W., CARVALHO, S. B., DAVIES, J., FERRIER, S., FRITZ, S. A., GRENYER, R., HELMUS, M. R., JIN, L. S., MOOERS, A. Ø., PAVOINE, S., PURSCHKE, O., REDDING, D. W., ROSAUER, D. F., WINTER, M. & MAZEL, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews* **92**, 698–715.
- VANE-WRIGHT, R. I., HUMPHRIES, C. J. & WILLIAMS, P. H. (1991). What to protect? Systematics and the agony of choice. *Biological Conservation* **55**, 235–254.
- VELLEND, M., CORNWELL, W. K., MAGNUSON-FORD, K. & MOOERS, A. Ø. (2011). Measuring phylogenetic biodiversity. In *Biological Diversity: Frontiers in Measurement and Assessment* (eds A. E. MAGURRAN and B. J. MCGILL), pp. 194–207. Oxford University Press, Oxford.
- VERDE ARREGOITIA, L. D., BLOMBERG, S. P. & FISHER, D. O. (2013). Phylogenetic correlates of extinction risk in mammals: species in older lineages are not at greater risk. *Proceedings of the Royal Society B* **280**, 20131092.
- VERON, S., CLERGEAU, P. & PAVOINE, S. (2016). Loss and conservation of evolutionary history in the Mediterranean Basin. *BMC Ecology* **16**, 43.
- VERON, S., DAVIES, T. J., CADOTTE, M. W., CLERGEAU, P. & PAVOINE, S. (2017). Predicting loss of evolutionary history: where are we? *Biological Reviews* **92**, 271–291.
- VILLÉGER, S., MASON, N. W. & MOUILLOT, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* **89**, 2290–2301.
- VIOLLE, C., ENQUIST, B. J., MCGILL, B. J., JIANG, L., ALBERT, C. H., HULSHOF, C., JUNG, V. & MESSIER, J. (2012). The return of the variance: intraspecific variability in community ecology. *Trends in Ecology and Evolution* **27**, 244–252.
- 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.
- VIOLLE, C., THULLER, W., MOUQUET, N., MUNOZ, F., KRAFT, N. J., CADOTTE, M. W., LIVINGSTONE, S. W. & MOUILLOT, D. (2017). Functional rarity: the ecology of outliers. *Trends in Ecology & Evolution* **32**, 356–367.
- WALKER, B., KINZIG, A. & LANGRIDGE, J. (1999). Plant attribute diversity, resilience, and ecosystem function: the nature and significance of dominant and minor species. *Ecosystems* **2**, 95–113.
- WALKER, B. H. (1992). Biodiversity and ecological redundancy. *Conservation Biology* **6**, 18–23.
- WEBB, C. O., ACKERLY, D. D., MCPHEE, M. A. & DONOGHUE, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics* **33**, 475–505.
- WEIHER, E., CLARKE, G. D. P. & KEDDY, P. A. (1998). Community assembly rules, morphological dispersion, and the coexistence of plant species. *Oikos* **81**, 309–322.
- WEIHER, E., FREUND, D., BUNTON, T., STEFANSKI, A., LEE, T. & BENTIVENGA, S. (2011). Advances, challenges, and a developing synthesis of ecological community assembly theory. *Philosophical Transactions of the Royal Society B-Biological Sciences* **366**, 2403–2413.
- WINTER, M., DEVICTOR, V. & SCHWEIGER, O. (2013). Phylogenetic diversity and nature conservation: where are we? *Trends in Ecology and Evolution* **28**, 199–204.

## VIII. SUPPORTING INFORMATION

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

**Appendix S1.** List of main originality indices cited, with references, abbreviations and some of their key properties.

**Appendix S2.** Illustration associated with Table 1.

(Received 25 June 2018; revised 7 February 2019; accepted 11 February 2019; published online 12 March 2019)



**Appendix S1. List of main originality indices cited in this paper, with references, abbreviations and some of their key properties.**

In Table S1.1, we describe the most used indices that measure species originality, but we are aware that other indices exist and many have been reviewed and compared elsewhere (e.g., Redding *et al.*, 2014; Cadotte & Davies, 2016). Our choice was structured based on the goal of giving an overview of the variety of originality indices and showing their link with diversity, which is a property of a set of species, and with abundance-based rarity, which is a property of individual species. We describe those links in detail in part III of the main text.

**Table S1.1.** Indices developed for measuring originality. The first column distinguishes among four different types of data used to evaluate species originality: tree structures (phylogenetic trees or trait-based dendrograms), dissimilarity matrices (phylogenetic or trait-based distances), multidimensional space axes and indices based on generalized entropy.

Data type	Method description	Reference	Name <sup>1</sup> (Examples of abbreviations used in the literature)	Code used in our paper	Some known properties
Tree-based approach, no branch length <sup>2</sup>	Inversely proportional to the number of branches descending from each internal node between the focal species (tip) and the root	May, 1990	May index ( $MVW$ [Redding <i>et al.</i> , 2014]; $M$ [Pavoine <i>et al.</i> , 2017])	$M$	Compared to $VW$ , it accounts for polytomies in phylogenetic trees (May, 1990)
	Inversely proportional to the number of internal nodes between the focal species (tip) and the root	Vane-Wright <i>et al.</i> , 1991	Vane-Wright <i>et al.</i> index ( $VW$ [e.g., Redding <i>et al.</i> , 2014])	$VW$	Compared to $M$ , it does not account for polytomies in phylogenetic trees (May, 1990)



	Inversely related to the sum of scores attributed to the nodes between the focal species (tip) and the root. A score is attributed to each node depending on whether it has more or less species descending from it than its sister nodes	Nixon & Wheeler, 1992; Redding <i>et al.</i> , 2014	$NWU = BPD^{-1}$ (Redding <i>et al.</i> , 2014; BPD is the unweighted binary phylogenetic diversity of Nixon & Wheeler, 1992)	NW U	Ranks all members of a clade as having the same priority relative to all members of its sister clade (Nixon & Wheeler, 1992)
	Inversely related to the sum of the number of species for each clade to which the focal species belongs		$NWW = WPD^{-1}$ (Redding <i>et al.</i> , 2014; WPD is the weighted phylogenetic diversity of Nixon & Wheeler, 1992)		NW W
Tree-based approach, with branch lengths <sup>2</sup>	Apportions the evolutionary history represented by a branch length equally among descending daughter branches	Redding & Mooers, 2006	Equal splits ( <i>ES</i> [Redding & Mooers, 2006])	<i>ES</i>	Sensitive to phylogenetic revisions and strongly dependent on the terminal branch lengths (Redding <i>et al.</i> , 2014)
	Apportions the evolutionary history represented by a branch length fairly among descending daughter species	Redding, 2003	Fair proportion ( <i>FP</i> [Redding, 2003]); evolutionary distinctiveness ( <i>ED</i> [Isaac <i>et al.</i> , 2007])		
	Distance from a focal species to where it subtends the tree	Altschul & Lipman, 1990	Pendant edge ( <i>PE</i> [Redding <i>et al.</i> , 2008])	<i>PE</i>	Overweights unique species near the tips of a tree (Redding <i>et al.</i> , 2008); ignores the pattern of shared evolution between species

	Expected average rarity of the characters of a species under a framework of genome evolution	Huang <i>et al.</i> , 2011	Character rarity ( <i>CHR</i> [Huang <i>et al.</i> , 2011])	<i>CHR</i>	If the phylogenetic tree is ultrametric, then <i>CHR</i> is equal to the <i>APD</i> and is correlated with <i>NWW</i> ; incorporates models of dynamic processes of trait evolution (Huang <i>et al.</i> , 2011)
	Average pairwise phylogenetic distance from a focal species to all the others in a reference species set	Redding <i>et al.</i> , 2014	Average phylogenetic distance ( <i>APD</i> [Redding <i>et al.</i> , 2014]) <sup>3</sup>	<i>APD</i>	Less sensitive to the terminal branch length changes than are <i>FP</i> and <i>ES</i> (Redding <i>et al.</i> , 2014)
Dissimilarity-based approach	Shortest distance between the focal species and all the others in the set	Pavoine <i>et al.</i> , 2017	Nearest neighbour ( <i>NN</i> [Pavoine <i>et al.</i> , 2017])	<i>NN</i>	With dissimilarity-based approaches, trait-based and phylogenetic originality can be compared with less distortion of trait information than can tree-based approaches
	Average pairwise dissimilarities between the focal species and all the others in the set	Eiswerth & Haney, 1992	Genetic distinctiveness (Eiswerth & Haney, 1992); average <sup>3</sup> ( <i>AV</i> , Pavoine <i>et al.</i> , 2017)	<i>AV</i>	
Originality in a multidimensional space <sup>4</sup>	Distance from a focal species' point to the hypothetical average centroid of all species	Magnuson-Ford <i>et al.</i> , 2009	Magnusson <i>et al.</i> index	— <sup>5</sup>	May discriminate the traits responsible for the originality of identified species

Entropy-based approach	Estimates the abundances species should have to maximize the diversity of a set; these abundances reflect species' originalities (see section III.3.c)	Pavoine <i>et al.</i> , 2005	Quadratic entropy-based (QE-based [Pavoine <i>et al.</i> , 2005]; $Qb$ [Pavoine <i>et al.</i> , 2017])	$Qb$	Depends on ultrametric distances between species and on an entropy index (the quadratic entropy, Rao, 1982) that weights commonness over rarity (Pavoine <i>et al.</i> , 2017)
		Pavoine <i>et al.</i> , 2017	Pavoine <i>et al.</i> , index ( $Rb$ [Pavoine <i>et al.</i> , 2017])	$Rb$	Can handle any matrix of dissimilarity between species (does not depend on ultrametric distances); depends on an entropy index (the $R$ index) that weights rarity over commonness; is similar to $AV$ , but better discriminates original from redundant species than does $AV$ (Pavoine <i>et al.</i> , 2017)

<sup>1</sup> When we did not find a specific name, we used the name of the authors who developed the index.

<sup>2</sup> Tree-based approaches were generally applied to phylogenies. However, they have also been applied to functional dendrograms (e.g.; Sobral *et al.*, 2016; Pavoine *et al.*, 2017). In the latter case, the originality indices are sensitive to the way in which the dendrograms have been built, with a risk of distorting the information contained in the traits.

<sup>3</sup>  $APD$  is actually a special application of  $AV$ . However, in  $APD$ , the dissimilarities between species are calculated from a phylogenetic tree; we thus classified  $APD$  as belonging to the tree-based indices.

<sup>4</sup> Ways of multidimensional space construction: a functional multidimensional space can be obtained by at least three approaches. First, each axis in the functional space can be a functional continuous trait. Second, axes can be calculated with a principal component analysis (PCA) applied

to continuous traits or by a Hill and Smith PCA applied to continuous and categorical traits. Third, axes can be calculated with a principal coordinate analysis (PCoA) or non-metric multidimensional scaling (nMDS) applied to trait-based dissimilarities between species. A phylogenetic space can be obtained by applying the PCoA or nMDS to phylogenetic dissimilarities between species. Magnusson *et al.* (2009) originally used a PCA on a matrix of species' morphological traits; they defined their index as the Euclidean distance from the origin of the PCA (the origin being the centroid of all points; we propose here to define species originality more generally as the distance from a focal species' point to the hypothetical average centroid of all species).

<sup>5</sup> We did not find abbreviation for this index and did not use any in our paper.

### Literature cited

- ALTSCHUL, S. F. & LIPMAN, D. J. (1990). Equal Animals. *Nature* **348**:493–494.
- CADOTTE, M. W. & DAVIES, T. J. (2016). *Phylogenies in Ecology: a Guide to Concepts and Methods*. Princeton University Press, Princeton.
- EISWERTH, M. E. & HANEY, J. C. (1992). Allocating conservation expenditures: accounting for inter-species genetic distinctiveness. *Ecological Economics* **5**, 235–249.
- HUANG, J., MI, X. & MA, K. (2011). A genome evolution-based framework for measures of originality for clades. *Journal of Theoretical Biology* **276**, 99–105.
- ISAAC, N. J. B., TURVEY, S. T., COLLEN, B., WATERMAN, C. & BAILLIE, J. E. M. (2007). Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PloS ONE* **2**, e296.
- 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.
- MAY, R. M. (1990). Taxonomy as destiny. *Nature* **347**, 129–130.
- NIXON, K. C. & WHEELER, Q. D. (1992). Measures of phylogenetic diversity. In *Extinction and Phylogeny* (eds M. J. NOVACEK and Q. D. WHEELER), pp. 216–234. Columbia University Press, New York.
- PAVOINE, S., BONSALE, M. B., DUPAIX, A., JACOB, U. & RICOTTA, C. (2017). From phylogenetic to functional originality: guide through indices and new developments. *Ecological Indicators* **82**, 196–205.
- PAVOINE, S., OLLIER, S. & DUFOUR, A. B. (2005). Is the originality of a species measurable? *Ecology Letters* **8**, 579–586.
- RAO, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology* **21**, 24–43.
- REDDING, D. W. (2003). *Incorporating Genetic Distinctness and Reserve Occupancy into a Conservation Prioritisation Approach*. Master Thesis: University of East Anglia, Norwich.
- REDDING, D. W., HARTMANN, K., MIMOTO, A., BOKAL, D., DEVOS, M. & MOOERS, A. Ø. (2008). Evolutionarily distinctive species often capture more phylogenetic diversity than expected. *Journal of Theoretical Biology* **251**, 606–615.
- REDDING, D. W., MAZEL, F. & MOOERS, A. Ø. (2014). Measuring evolutionary isolation for conservation. *PloS ONE* **9**, e113490.

- REDDING, D. W. & MOOERS, A. Ø. (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology* **20**, 1670–1678.
- SOBRAL, F. L., LEES A. C. & CIANCIARUSO, M. V. (2016). Introductions do not compensate for functional and phylogenetic losses following extinctions in insular bird assemblages. *Ecology Letters* **19**, 1091–1100.
- VANE-WRIGHT, R. I., HUMPHRIES, C. J. & WILLIAMS, P. H. (1991). What to protect? Systematics and the agony of choice. *Biological Conservation* **55**, 235–254.

## Appendix S2. Illustration associated with Table 1.

Consider a community with  $S$  species and a phylogenetic tree with the species as tips. As an illustration, we will consider the data set in Fig. S2.1. Pavoine, Love and Bonsall (2009) developed the following parametric index of phylogenetic diversity for the community:

$$I_q = \sum_{k=1}^K T_k \left[ \frac{1 - \sum_{b \in P_k} p_b^q}{q-1} \right] \quad (\text{eq. 1})$$

where  $q$  is the parameter of the diversity index. It controls the importance given to lineage abundances.  $k$  is a period in the phylogeny. Periods are defined between two successive speciation events (see Fig. S1.1(B) for an example).  $T_k$  is the length of period  $k$  (time elapsed between the two speciation events).  $P_k$  is the set of branches that cross period  $k$ . For example in Fig. S1.1(C), we highlighted in red the set of branches that cross the second period.  $b$  is a branch of the phylogenetic tree.  $L_b$  is the length of branch  $b$ .  $p_b$  is the summed relative abundance of all species descending from branch  $b$ .

Applied to the data set of Fig. S1.1,  $I_q$  is equal to:

$$I_q = T_1 \left[ \frac{1 - (p_A + p_B + p_C)^q - (p_D + p_E)^q}{q-1} \right] + T_2 \left[ \frac{1 - (p_A + p_B)^q - p_C^q - (p_D + p_E)^q}{q-1} \right] + T_3 \left[ \frac{1 - (p_A + p_B)^q - p_C^q - p_D^q - p_E^q}{q-1} \right] + T_4 \left[ \frac{1 - p_A^q - p_B^q - p_C^q - p_D^q - p_E^q}{q-1} \right] \quad (\text{eq. 2})$$

This can also be written as:

$$I_q = \frac{1}{q-1} \left[ \begin{aligned} & (p_A + p_B + p_C)L_1(1 - (p_A + p_B + p_C)^{q-1}) + (p_D + p_E)L_{6,1}(1 - (p_D + p_E)^{q-1}) \\ & + (p_A + p_B)L_{2,2}(1 - (p_A + p_B)^{q-1}) + p_C L_{3,2}(1 - p_C^{q-1}) + (p_D + p_E)L_{6,2}(1 - (p_D + p_E)^{q-1}) \\ & + (p_A + p_B)L_{2,3}(1 - (p_A + p_B)^{q-1}) + p_C L_{3,3}(1 - p_C^{q-1}) + p_D L_{7,3}(1 - p_D^{q-1}) + p_E L_{8,3}(1 - p_E^{q-1}) \\ & p_A L_4(1 - p_A^{q-1}) + p_B L_5(1 - p_B^{q-1}) + p_C L_{3,4}(1 - p_C^{q-1}) + p_D L_{7,4}(1 - p_D^{q-1}) + p_E L_{8,4}(1 - p_E^{q-1}) \end{aligned} \right]$$

(eq. 3)

and thus

$$I_q = \frac{1}{q-1} \left[ \begin{array}{l} (p_A + p_B + p_C)L_1(1 - (p_A + p_B + p_C)^{q-1}) \\ + (p_A + p_B)L_2(1 - (p_A + p_B)^{q-1}) + p_C L_3(1 - p_C^{q-1}) + p_A L_4(1 - p_A^{q-1}) \\ + p_B L_5(1 - p_B^{q-1}) + (p_D + p_E)L_6(1 - (p_D + p_E)^{q-1}) \\ + p_D L_7(1 - p_D^{q-1}) + p_E L_8(1 - p_E^{q-1}) \end{array} \right] \quad (\text{eq. 4})$$

Index  $I_q$  can thus also be written more generally as

$$I_q = \sum_b L_b p_b \left( \frac{1 - p_b^{q-1}}{q-1} \right) \quad (\text{eq. 5})$$

which, for ultrametric trees, is equivalent to the general equation in Chiu and Chao (2014), as follows:

$$I_q = (T_1 + T_2 + T_3 + T_4 - \sum_b L_b p_b^q) / (q-1) \quad (\text{eq. 6})$$

Let  $p_i$  be the relative abundance of species  $i$ , and let  $C(i, \text{Root})$  be the shortest path from species  $i$  (tip) to the root of the tree. In Table 1, we highlighted that  $I_q$  can also be written as a mean of abundance-weighted originalities:

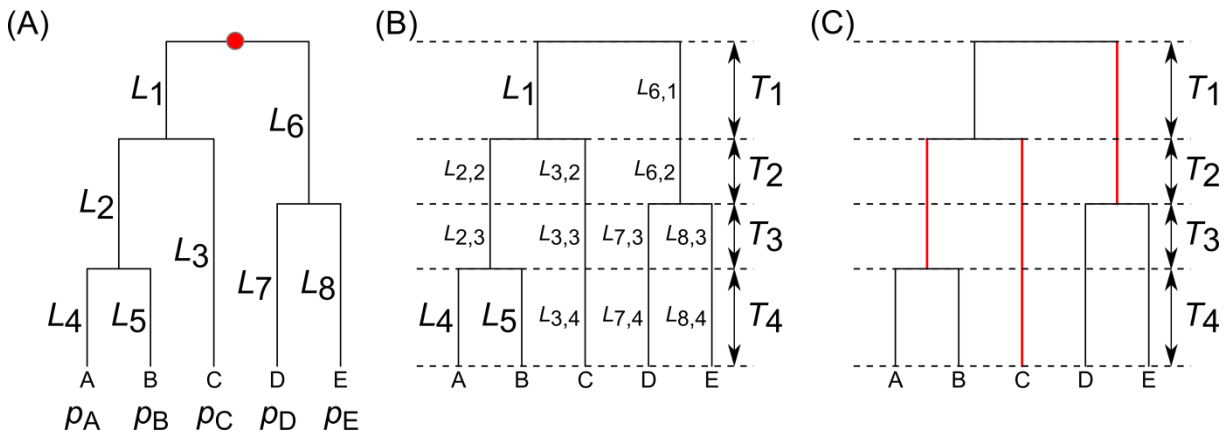
$$I_q = \sum_{i=1}^S p_i O_q \quad (\text{eq. 7})$$

with

$$O_q = \left[ \sum_{b \in C(i, \text{Root})} L_b \left( \frac{1 - p_b^{q-1}}{q-1} \right) \right] \quad (\text{eq. 8})$$

Applied to the theoretical example of Fig. S1.1, eq. 7 corresponds to reorganizing eq. 4 as follows:

$$I_q = \frac{1}{q-1} \left\{ \begin{array}{l} p_A \left[ L_1 (1 - (p_A + p_B + p_C)^{q-1}) + L_2 (1 - (p_A + p_B)^{q-1}) + L_4 (1 - p_A^{q-1}) \right] \\ p_B \left[ L_1 (1 - (p_A + p_B + p_C)^{q-1}) + L_2 (1 - (p_A + p_B)^{q-1}) + L_5 (1 - p_B^{q-1}) \right] \\ + p_C \left[ L_1 (1 - (p_A + p_B + p_C)^{q-1}) + L_3 (1 - p_C^{q-1}) \right] \\ + p_D \left[ L_6 (1 - (p_D + p_E)^{q-1}) + L_7 (1 - p_D^{q-1}) \right] \\ + p_E \left[ L_6 (1 - (p_D + p_E)^{q-1}) + L_8 (1 - p_E^{q-1}) \right] \end{array} \right\}$$



**Fig. S2-1.** Theoretical data set with five species, named from A to E, and an ultrametric phylogenetic tree with the species as tips. In panel (A), the main aspects of the data set are given. The red circle represents the root node. Branch lengths of the tree are indicated and denoted by  $L_1$  to  $L_8$ . The relative abundances of the species in a community are noted from  $p_A$  for species A to  $p_E$  for species E. In (B), the phylogeny is split into periods, from  $T_1$  (the oldest) to  $T_4$  (the most recent). These periods are defined between successive speciation events. The length of the branches that cross several periods is also split per period. The notation  $L_{ij}$  means the length of the section of branch  $i$  in period  $j$ . This leads to  $L_2 = L_{2,2} + L_{2,3}$ ,  $L_3 = L_{3,2} + L_{3,3} + L_{3,4}$ ,  $L_6 = L_{6,1} + L_{6,2}$ ,  $L_7 = L_{7,3} + L_{7,4}$ ,  $L_8 = L_{8,3} + L_{8,4}$ ,  $L_{6,1} = T_1$ ,  $L_{2,2} = L_{3,2} = L_{6,2} = T_2$ ,  $L_{2,3} = L_{3,3} = L_{7,3} = L_{8,3} = T_3$ , and  $L_{3,4} = L_{7,4} = L_{8,4} = T_4$ . In (C), as an example, we overlined the branches that cross the second period of length  $T_2$  in red.



## References

- CHIU, C.-H., JOST, L. & CHAO, A. (2014). Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs* **84**, 21–44.
- PAVOINE, S., LOVE, M. & BONSALL, M. (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. *Ecology Letters* **12**, 898–908.

**DIVERSITY IN THE CITY:  
URBANIZATION AFFECTS  
DIFFERENTLY TRAIT-BASED AND  
PHYLOGENETIC PLANT  
ORIGINALITY AT LOCAL AND  
REGIONAL SCALES**

---

## INTRODUCTION TO CHAPTER 2

It is clear from Chapter 1 that species originality and diversity measures and concepts are tightly related. However, there were very few studies investigating this relation with a real dataset. In this chapter, we investigated species diversity dependence on species originality and their respective variation along an urbanization gradient. Urbanization is currently considered as one of the major threats to biodiversity worldwide, challenging species by habitat fragmentation and land-cover changes. Although, urban ecologists studied the biological diversity of cities, there is still a lack of information about how species contribute to the diversity observed in a given community and how it is affected by urbanization. In this study, we propose a methodological framework to analyze species originality and diversity at two spatial scales: local communities and regional species pool. We assessed species diversity as the mean of species local originalities, which represents a measure of community diversity, and the mean of regional originalities, which is viewed as the amount of the regional diversity expressed by species in the plot. Finally, we calculated the skewness of originalities which represents the fraction of original species in the community, i.e. the distribution of originality along redundancy-uniqueness gradient.

In the light of plant species importance for ecosystem services, we chose to apply our originality-diversity framework to the herbaceous plant communities in Paris and its surrounding region, Ile-de-France. We demonstrate that our framework can determine at which spatial scale, local or regional, urbanization induces variations in the diversity of the biological characteristics and evolutionary history of species, while identifying which species are responsible for these variations depending on their originality values. To our knowledge, this is the first real case study of plant originality measure in an urbanization context. We believe that the results of this study may assist urban policy-makers in efficient management of biodiversity.

This research work is a result of my collaboration with four research structures:

- Research team from Germany, based at **Helmholtz Centre for Environmental Research** – UFZ, Department Community Ecology, Halle (Saale) where I spent two months, in

January and February 2019, working with **Sonja Knapp**, **Walter Durka** and **Ingolf Kühn**. At that stage, I had already started plant data analysis, but there was still some missing data in plant traits and mismatches in taxonomical names between trait databases and species occurrences database. **SK**, **WD** and **IK** considerably improved the trait data, and their botanical expertise allowed to identify accidental species or obviously misidentified species (e.g., alpine species that occur only once in a dataset are more likely to be identification mistakes). **WD** also showed me methods he used to construct DaPhnE phylogeny of vascular plants so I could add species from my dataset missing in the tree. **IK** contributed significantly to the statistical analysis and improved it by adding spatial autocorrelation analysis between species plots.

- **Vigie-flore** citizen science program, which is supervised by researchers from my own research unit and by the Vigie-Nature program: **Gabrielle Martin**, **Nathalie Machon** and **Emmanuelle Porcher** managed Vigie-flore data; **Eric Motard** supervised data collection and input to the database; **GM** substantially helped at the very beginning of this study, by guiding me through data structure and providing ideas for traits use.
- **Conservatoire Botanique National du Bassin Parisien** is represented by **Jeanne Vallet**, who provided several ideas for the originality framework and helped me with plant species taxonomy and occurrence probabilities.
- My PhD supervisors, **Sandrine Pavoine** and **Philippe Grandcolas**, who were invested at every stage of this study, bringing their expertise in statistical analysis and general context information.

# Diversity in the city: urbanization affects trait-based and phylogenetic plant originality at local and regional scales differently

Anna Kondratyeva<sup>1,2\*</sup>, Sonja Knapp<sup>3</sup>, Walter Durka<sup>3,4</sup>, Ingolf Kühn<sup>3,4,5</sup>, Jeanne Vallet<sup>6</sup>, Nathalie Machon<sup>1</sup>, Gabrielle Martin<sup>7</sup>, Eric Motard<sup>8</sup>, Philippe Grandcolas<sup>2</sup>, Sandrine Pavoine<sup>1\*</sup>

<sup>1</sup>*Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum National d'Histoire Naturelle, Centre National de la Recherche Scientifique, Sorbonne Université, Paris, France*

<sup>2</sup>*Institut Systématique Evolution Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Paris, France*

<sup>3</sup>*Helmholtz Centre for Environmental Research – UFZ, Department Community Ecology, Halle (Saale), Germany*

<sup>4</sup>*German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Germany.*

<sup>5</sup>*Martin Luther University Halle-Wittenberg, Geobotany and Botanical Garden, Halle, Germany*

<sup>6</sup>*Conservatoire Botanique National du Bassin Parisien (CBNBP), Muséum National d'Histoire Naturelle, Sorbonne Université, Paris, France*

<sup>7</sup>*Institut Méditerranéen de Biodiversité et d'Ecologie marine et continentale (IMBE), Univ Avignon, Aix Marseille Univ, CNRS, IRD, IUT site Agroparc, Avignon Cedex 09, France*

<sup>8</sup>*Institut d'Ecologie et des Sciences de l'Environnement de Paris (IEES), Sorbonne Université, Paris, France*

## \* Correspondence:

Anna Kondratyeva: [anna.kondratyeva@edu.mnhn.fr](mailto:anna.kondratyeva@edu.mnhn.fr)

Sandrine Pavoine: [sandrine.pavoine@mnhn.fr](mailto:sandrine.pavoine@mnhn.fr)

**Keywords:** biodiversity measure, citizen science, functional trait, originality measure, phylogeny, spatial scale, urbanization

**This version is in revision for *Frontiers Ecology and Evolution***

## **ABSTRACT**

Urbanization is one of the most intensive and rapid human-driven factors that influence biodiversity. Thus, society needs appropriate methods for understanding species responses to urbanization processes. To find the appropriate methods, species diversity must be surveyed and measured. Although a number of biodiversity metrics exist in ecology, little is known about the contributions of different species to urban community diversity. In urban contexts, studies have mostly analyzed the influence of urbanization on the spatial distributions of species and on their functional traits and phylogenetic community structures. Here, we develop a framework that relies on the originality of a species. A species is original if it possesses rare trait values compared to all others in the community of species or if it is distantly related to the other species in the community. The most original species have the greatest contributions to the diversity of that community. We studied the influence of urbanization on species originality using Vigie-flore, the French citizen science dataset of plant occurrences in the Île-de-France region. To visualize the urbanization gradient, we calculated the percentage of urbanized land cover around the species plots. Since urbanization acts simultaneously at multiple spatial levels, we considered both the local and regional spatial scales as areas for the originality calculations. Using several trait databases and recent plant phylogeny, we measured trait-based and phylogenetic originality. Then, for each plot, we calculated the mean of the local species originalities, which represented a measure of community diversity, and the mean of the regional originalities, which represented the amount of the regional diversity expressed by species in the plot. While the trait-based originality increased along the urbanization gradient at both scales, the phylogenetic originality decreased with urbanization at only the regional scale, meaning that urban communities do not contribute to regional phylogenetic diversity. We also found that urbanization made some species original at the local scale, although they were not at the regional scale. Our conceptual framework will help to identify the species responsible for variations in diversity along environmental gradients and to determine the most relevant spatial scale for urban sustainable management.

## INTRODUCTION

The United Nations (2018) predicted that by the year 2050, 68% of the global human population will live in urban areas. This process comes with a complex mix of challenges, including land-use modifications and habitat fragmentation, which increase the pressure on local remnant species diversity (Gaston, 2010). Urbanizing processes such as residential and commercial development are listed in the International Union for the Conservation of Nature (IUCN) Unified Classification of Direct Threats scheme (version 3.2) as threats with substantial impacts on biodiversity. In addition, some the components of biodiversity have the potential to provide ecosystem services to humanity (Millennium Ecosystem Assessment (MEA), 2005; Robinson and Lundholm, 2012; Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), 2019). Therefore, it is now recognized that biodiversity needs to be fully considered in urban planning and landscape design (Ahern, 2013).

Urbanization incorporates a set of anthropogenic filters and modifies natural ecological filters that determine species diversity and composition of urban areas (Williams et al., 2009; Aronson et al., 2016). Now, it has been largely recognized that cities, representing a high spatial heterogeneity of habitats, can harbor an important amount of plant and animal species (McKinney, 2002; Kühn et al., 2004; Godefroid and Koedam, 2007; Pautasso et al., 2011) and that cities play an important role in biodiversity conservation (Kowarik, 2011; Ives et al., 2016; Soanes and Lentini, 2019). In contrast, many studies show negative impacts of urbanization on biological diversity mostly by biotic homogenization (Kühn and Klotz 2006; McKinney, 2006; Devictor et al., 2007, 2008; Morelli et al., 2016). These different viewpoints might be explained by scale-dependency of the urbanization effect on biodiversity (Pautasso, 2007). However, there is still a lack of understanding of the responses of species to urbanization at different spatial scales.

One of the most investigated taxonomic groups in urban ecology is composed of vascular plant species. These species have an important role in primary productivity (Lieth, 1975) and provide habitat for other taxa (Tews et al., 2004). Moreover, vegetation structure and vascular plant diversity

have been used as general indicators of urban environment conditions (Sukopp and Weiner, 1983; Tzoulas and James, 2010). The role of urban areas in biodiversity conservation has been increasingly recognized by policy makers (Grimm et al., 2000; Gill et al., 2007; Lundholm et al., 2014). For example, the development of green infrastructure has been encouraged in cities as a means to increase the resilience of cities to global changes and to improve human well-being (European Commission's Directorate-General Environment, 2012; Kabisch et al., 2016). Plant species are the focus of this development. To guide such developments, efforts to deepen the collective understanding of the responses of plant species to urbanization are the center of urban biodiversity research (Werner and Zahner, 2009; Palma et al., 2017), conservation (McKinney, 2006) and the socioecological sciences (Elands et al., 2015).

To guide sustainable urban development, biodiversity must be evaluated first. Measures of biodiversity should ideally represent all of its multifaceted and hierarchical complexity. In the urban context, studies have mainly focused on describing species distributions and abundances (Aronson et al., 2016; Deguines et al., 2016; Guetté et al., 2017), species trait distributions (Knapp et al., 2009, 2010; Vallet et al., 2010; Williams et al., 2015; Kalusová et al., 2017) and the phylogenetic structures of urban communities (Ricotta et al., 2012, 2015; Morelli et al., 2016; Knapp et al., 2017; Sol et al., 2017). Several measures can thus be used to evaluate urban biodiversity, including species richness (McKinney, 2002, 2008; Cervelli et al., 2013), trait-based diversity (Petchey and Gaston, 2002; Mason et al., 2005; Villéger et al., 2008) and phylodiversity (Faith, 1992; Webb, 2000; Pavoine and Bonsall, 2011). Trait-based approaches allow the measurement of community trait structure (Laliberté and Legendre, 2010), and phylogenetic approaches divulge the evolutionary history of a given community (Tucker et al., 2017, 2019). However, these measures represent the whole diversity of a community and give no information about the contributions of individual species to this diversity (Kondratyeva et al., 2019).

In urbanized areas, species constantly interact with their rapidly changing biotic environment (e.g., non-native species, Gross et al., 2013; Aronson et al., 2014; Kühn et al., 2017). Identifying the



species that are the most threatened by urbanization and estimating the consequences their local extinction could have on biodiversity could help in the prioritization of conservation actions. Thus, there is a need for a metric that evaluates the contribution of each species to the diversity of a reference set of species using pairwise differences among species based on their biological characteristics. Originality measures are such a tool (see Pavoine et al., 2017; Kondratyeva et al., 2019 for reviews on originality measures). A species is original if it has rare biological characteristics within a reference set of species (Pavoine et al., 2005). Such rare biological characteristics can be the distant position of the reference species on a phylogenetic tree (Isaac et al., 2007; Jetz et al., 2014; Redding et al., 2015a; Sol et al., 2017) or they can be trait values that few species share (Mouillot et al., 2013a; Brandl et al., 2016; Pavoine et al., 2017). Thus, in a species assemblage, the less a given species shares its trait values with the other species or the more distantly related it is to the other species, the more original it is and the more it contributes to the diversity of that species set. Inversely, a species is redundant and contributes less to community diversity when it shares its trait values with many other species and/or is closely related to the other species. Species originality and redundancy are often used as indicators of ecosystem functioning (Mouillot et al., 2013a; Brandl et al., 2016). For example, higher trait-based redundancy among species would increase interspecific competition for a limiting resource but, at the same time, would ensure higher ecosystem stability in changing environments (Walker et al., 1999; Tilman et al., 2014).

Biodiversity and originality measures are scale dependent. Moreover, urbanization acts simultaneously at various spatial scales (Aronson et al., 2016), with the response of species varying across scales, according to their evolutionary histories, trait values and the ecological processes they are involved in (Levin, 1992; Chave, 2013). Thus, studying species responses to urbanization necessitates data for local communities across multiple locations. Today, an increasingly burgeoning citizen science program, particularly in ecology, allows the assembly of large amounts of monitoring data (Kobori et al., 2016; Silvertown, 2016). In France, Vigie-Nature ([www.vigienature.fr](http://www.vigienature.fr)) is one of the largest national initiatives for citizen science in biodiversity

monitoring. It has already proven to be valuable in conservation policy-making (Couvet and Prevot, 2015). One of Vigie-Nature national programs, called Vigie-flore (<http://www.vigienature.fr/fr/vigie-flore>), has gathered information on plant occurrences across the country for every year since 2009. This long-term dataset has recently allowed the reporting of rapid shifts in plant community composition in response to climate change (Martin et al., 2019). Therefore, as citizen science data accumulate across space and time, their usefulness in ecology and conservation studies is increasingly acknowledged (Dickinson et al., 2010; McKinley et al., 2017). Here, using the Vigie-flore program, we analyzed the impact of urbanization on species originality at two spatial scales: local plant communities and the regional species pool. We used the mean of species-specific originality values, which represent a measure of local trait-based (functional) diversity and phylodiversity within a plant community. We investigated three questions as follows: 1. How do regional and local species originality contribute to variations in community diversity along an urbanization gradient? 2. How do non-native species contribute to community diversity? 3. Which species in particular drive community diversity across spatial scales? Since urban environments affect both the trait-based and phylogenetic compositions of communities (Knapp et al., 2008), we measured phylogenetic and trait-based originality for plant species. Finally, with this study, we show that originality measures are worth integrating into urban biodiversity management as metrics that are capable of integrating differences among the species of a community.

## **MATERIALS & METHODS**

### **Study Area**

We used plant community data sampled across the Île-de-France region by the participants of the national citizen science survey, Vigie-flore, which constitute a large national dataset of plant occurrences in diverse habitats. The Île-de-France region occupies a territory of 12 100 km<sup>2</sup> and is the most inhabited region of France, with 12.1 million inhabitants. It includes the city of Paris, which is the French capital and the nation's largest city with 2.2 million inhabitants in 2016, and

approximately 1300 other cities and towns with varying human populations (min = 27, max = 119 645, median = 1318, Institut d'Aménagement et d'Urbanisme (IAU), 2019). This region benefits from having the highest sampling effort within the Vigie-flore program (188 grid cells out of total of 586) as well as a detailed land cover geodatabase.

### **Biological Data**

Following the standardized protocol of the Vigie-flore program, skilled amateur botanists investigate randomly selected 1 km<sup>2</sup> grid cells. Each grid cell contains eight systematically distributed 10 m<sup>2</sup> plots. In each plot, the presence of all vascular plants is recorded. We considered the set of species in each 10 m<sup>2</sup> plot as a local community assemblage. We analyzed a subset of 2362 plots collected over the years 2009 to 2017 in all types of habitats (1231 in agricultural areas, 533 in forests, 281 in urbanized areas, 261 in artificial open areas and 56 in semi-natural areas). For computational reasons, only plots with a minimum of three species were retained. Indeed, it is meaningless to measure originality in a plot with one species where there are no other species, and in a plot with two species, each of the species would have equal originality. A total of 322 plots missing geolocalization were also excluded from the study. Out of 812 species recorded by Vigie-flore and found in Île-de-France, 593 herbaceous angiosperm species were retained. To assemble as much trait data as possible, we standardized species taxonomic nomenclature among datasets by using several synonymous names and updating them by using The Plant List (<http://www.theplantlist.org/>), the Tela Botanica website of the French botanist's network (<https://www.tela-botanica.org/>) and Flora Gallica (Tison and de Foucault, 2017). For further analysis, we kept species names as found in Vigie-flore data with a taxonomical reference TAXREF (V2.0, Gargominy, 2008). We removed some data from the initial species dataset: 1. species occurring only in plots of 1 or 2 species (see explanation above); 2. species identified as accidental or obviously misidentified (e.g., alpine species that occur only once in a dataset are more likely to be identification mistakes); 3. woody species, as they seem to be affected less strongly by urbanization than other groups (Neil and Wu, 2006) and because of their outlying trait values; 4.

Bryophyta and Pteridophyta species; and 5. species that rely on human assistance (no cultivated or ornamental species). Woody and herbaceous species were distinguished following the LEDA trait database (Kleyer et al., 2008).

### **Land Cover Data**

We used vectorized maps of land cover in Île-de-France from Institut d'Aménagement et d'Urbanisme, with a spatial resolution of 12.5 meters available for two time periods (2009 to 2011 and 2012 to 2017). We used six land cover categories (Figure 1): 1. forest (24% of Île-de-France area), 2. semi-natural area (2%), 3. agricultural area (50%), 4. artificial open area (6%), 5. water (2%) and 6. urbanized area (16%, including individual housing (8%), collective housing (2%), industrial and business areas (2%), facilities (1%), quarries, dumps and worksites (0.5%) and transportation infrastructures (2.5%)). A detailed description of each category is available online (<http://www.iau-idf.fr/>). Using the ArcGIS® software for geographical data manipulation (Environmental Systems Research Institute, Esri, 2018, ArcGIS release 10.5.1), we linked the Vigie-flore plots with land cover data and calculated the proportion of urbanized areas as a surrogate for the urbanization gradient, i.e., we calculated the percentage of urbanized area inside each 1x1 km<sup>2</sup> grid cell where Vigie-flore plots occurred.

### **Species Biological Characteristics and Phylogeny**

We collected information on the biological traits of each species from several databases (Table S1): the BiolFlor database on biological and ecological traits of the German flora (Klotz et al., 2002), LEDA database of life-history traits of the Northwest European flora (Kleyer et al., 2008), TRY global database of plant traits (Kattge et al., 2011), Ecoflora database of British Isles (<http://www.ecoflora.co.uk>, Fitter and Peat, 1994) and Catminat database of French flora (Julve, 1998) along with numerous botanical garden resources across Europe (e.g., the Royal Botanic Gardens, Kew, POWO, 2019). The data included three binary (life span, pollination vector, seed dispersal mode), one ordinal (type of reproduction), one circular (beginning of flowering) and six

quantitative traits (seed weight, leaf dry mass content, leaf size, specific leaf area, plant height and duration of flowering) related to the plant species dispersal, establishment, reproduction and persistence (Table S2). We selected traits that are widely used in functional ecology studies for their presumed roles in the context of urbanization (Knapp et al., 2009, 2010; Vallet et al., 2010; Williams et al., 2015; Kalusová et al., 2017) and that were available for the majority of our species. However, the Grime's strategies (Grime et al., 1988) or the Raunkiær life forms (Raunkiær, 1932) that were largely available for our species are synthetic traits and were not retained. Similarly, the Ellenberg indicator values were not retained either because they are not functional traits *per se* but represent the ecological preferences of species (Ellenberg et al., 1991) by characterizing a nonrandom association of plant species with particular characteristics of the environment (Garnier et al., 2017). We investigated the correlations among all retained biological traits to avoid potential collinearity. Highly correlated traits were not retained if the absolute values of Pearson correlation coefficient were superior to 0.7 (Dormann et al., 2013). In complement to the biological traits, we calculated species urbanity following Hill et al. (2002), which, like the Ellenberg indicator values, is an environmental association (*sensu* Garnier et al., 2017). We used species urbanity *a posteriori* to link species urban preferences with their trait-based and phylogenetic originality values. We calculated species urbanity as the mean percentage of urbanized area in all grid cells where the species occurred. In addition, we used the native status of the species that was extracted from the Île-de-France plant species list established by the Conservatoire Botanique National du Bassin Parisien (CBNBP, 2016) to separate the species into two groups of native and non-native species. A species was defined as non-native if it was introduced by humans into the region after the year 1500. We used DaPhnE (Durka and Michalski, 2012), a dated phylogeny that is resolved to the species level and covers the vascular flora of the British Isles, Germany, the Netherlands and Switzerland and thus all taxa from the LEDA and BiolFlor trait databases. The final tree was pruned to our species pool (Table S2) in R (R Core Team 2019).

## DATA ANALYSES

### Community and Species Metrics

To measure originality, we used the metric named  $AV$  for "average" as it represents the average dissimilarity between a focal species and all others in a set (Pavoine et al., 2017). The  $AV$  metric requires a dissimilarity matrix based on species phylogeny or the species trait values. To form a trait-based dissimilarity matrix, we first log-transformed the most asymmetrically distributed quantitative traits: canopy height, leaf size, and mean seed weight to ensure a bell-shaped distribution of values. We then used the mixed-variable coefficient of dissimilarity (Pavoine et al., 2009) associated with the most appropriate coefficients to estimate trait-based distances between pairs of species: the Jaccard metric (Jaccard, 1901) for binary traits, the Podani metric (Podani, 1999) for ordinal traits and the Manhattan metric (Farris, 1972; Nelson and Horn, 1975) for quantitative traits. To form a phylogenetic dissimilarity matrix, we calculated the patristic distances between pairs of species (sum of the branch lengths on the shortest path between two species in the phylogenetic tree).

Using the dissimilarity matrices, we calculated the phylogenetic and trait-based originality scores of each species with the *distinctDis* function of the R package *adiv* (Pavoine, 2018). Following Redding et al. (2015a), for each species, we calculated a regional originality score considering all species in our dataset ( $rAV$ ) and local originality scores considering only the species from a given plot ( $lAV$ ). Thus, for the same species, local originality varies depending on the composition of a given plot, whereas regional originality will be constant for a given species over all plots (Figure 2). We named a given plot  $P$  and the region  $R$  (union of all sampled plots).  $N_P$  is the number of species in plot  $P$ ,  $N_R$  is the number of species in region  $R$ , and  $d_{ij}$  is the functional or phylogenetic distance between species  $i$  and  $j$ . The regional originality of species  $i$  is as follows:

$$rAV_i = \frac{1}{N_R - 1} \sum_{j \in R} d_{ij}$$

If species  $i$  occurs in plot  $P$ , its local originality in plot  $P$  is as follows:

$$lAV_{iP} = \frac{1}{N_P - 1} \sum_{j \in P} d_{ij}$$

We defined the diversity of a given plot as the average of the local originalities of its species. This diversity metric gives a unique value to each set of species. Thus, in each plot  $P$ , we calculated the mean of the local originalities of the species ( $MIO_P$ ), which represents a measure of the local trait-based diversity or phylogenetic diversity within that plot, depending on whether trait-based or phylogenetic dissimilarities ( $d_{ij}$ ) were considered;  $MIO_P$  can be calculated as follows:

$$MIO_P = \frac{1}{N_P} \sum_{i \in P} lAV_{iP} = \frac{1}{N_P} \frac{1}{N_P - 1} \sum_{i,j \in P} d_{ij}$$

In each plot  $P$ , we also calculated the mean of the regional species originalities ( $MrO_P$ ), taken from regionally calculated values ( $rAV$ ) but only for species composing plot  $P$  (Figure 2).  $MrO_P$  can be viewed as the average contribution of a species from plot  $P$  to the (trait-based or phylogenetic) regional diversity and can be calculated as follows:

$$MrO_P = \frac{1}{N_P} \sum_{i \in P} rAV_i = \frac{1}{N_P} \sum_{i \in P} \frac{1}{N_R - 1} \sum_{j \in R} d_{ij}$$

To test if the observed variations in the mean trait-based and phylogenetic originality along the urbanization gradient were due to changes in the originality values of many species or were carried by a few very original (or very redundant) species, we calculated the skewness of the phylogenetic and trait-based originalities in each plot. We did so for the local and regional originality values.

Negative skewness of originality or left-skewed originality represented scenarios with many species of high originality and few redundant species. In contrast, positive skewness of originality or right-skewed originality represented many redundant species and a few highly original species.

We then created a dataset without 85 non-native species and recalculated all local and regional originality scores and all plot-level metrics (species richness and mean and skewness of the originalities) to assess the contributions of non-native species to originality. The species richness was calculated per plot for both datasets. Finally, we ranked species from the complete dataset (native and non-native species) by their regional and local originality values and selected the top scoring 5% of the species (the first 30 in the list). We also tested how many species become original at the local level although they were not at the regional level. To do so, we calculated the ratio of all local originality values to the regional originality value ( $LAV_{ip} / rAV_i$ ) for each species.

### **Statistical and Spatial Modeling**

We modeled the variation in community metrics with the complete dataset and the only native species dataset with generalized mixed linear models in the R package *glmmTMB* (Millar, 2011); the mean, skewness and ratio of trait-based and phylogenetic species originality at local and regional levels were calculated with a Gaussian distribution, while the species richness was calculated with a negative binomial distribution and a quadratic link between the mean and the variance. We included as fixed explanatory variables: the urbanization percentage in the 1 km<sup>2</sup> cell surrounding a given plot, the species richness (except for the model where species richness was a response variable) and the survey year, which was transformed by subtracting its minimum value to allow the comparison between units of variables. We also included an interaction effect between urbanization and species richness. The IDs of the grid cells and the IDs of the plots were modeled as random nested variables to account for the pseudoreplication of the plots at the same location. Species-specific local originality was tested as a response by adding to the previous mixed models the species name as a random factor to control for species that appeared in several plots. Backward selection of variables was based on Akaike information criterion (AIC) values, where the lowest



AIC value denoted the best model. We calculated the marginal and conditional  $R^2$  values as proposed in Nakagawa and Schielzeth (2013). The marginal  $R^2$  value describes the proportion of variance explained by fixed covariables only. The conditional  $R^2$  value describes the proportion of variance explained by fixed and random covariables.

As our data were distributed in space, we tested for spatial independence in the residuals of each model. We used the local Moran's I test, or the local indicator of spatial association (LISA, Anselin, 1995), and visualized it with a correlogram of 500 meter increments for neighborhood definition. The significance was assessed based on 999 randomizations with the R package *ncf* (Bjørnstad and Falck, 2001). No spatial autocorrelation was detected in the model residuals, and thus, no correction was needed (Figure S1). All statistical analyses were performed with R software version 3.6.0 (R Core Team, 2019), and statistical significance was considered to be present at  $\alpha = 0.05$ .

## RESULTS

### Community Level Metrics

All model results with the complete dataset are presented in Table 1. We observed a significant positive effect of urbanization on the mean trait-based originality, or trait-based diversity, of plots at both the regional and local scales. Species richness had a significant positive effect only on local-scale trait-based diversity. That is, the higher species richness was, the more the species differed locally in their traits. There was, however, a significant interaction effect between urbanization and species richness at both scales, such that the positive effect of urbanization on trait-based diversity (*MIO* and *MrO* with trait-based  $d_{ij}$  values) tended to decrease with increasing species richness and *vice versa* to the point of becoming negative for the highest levels of species richness (Figure 3A,B). Trait-based diversity at the local scale significantly decreased with the survey year. The mean phylogenetic originality of the plots decreased with urbanization and species richness only when evaluated at the regional scale (*MrO* with phylogenetic  $d_{ij}$  values, Figure 3C). However, there was also a significant interaction between species richness and urbanization: the negative effect of urbanization decreased with increasing species richness and *vice versa*. The mean phylogenetic

originality of the plots increased with the sample year at the regional scale only.

The skewness of trait-based originality at the local scale and phylogenetic originality at the local and regional scales did not change with urbanization. We found that the skewness of local trait-based originality significantly changed with species richness from close to zero to positive values (Figure 4A). The skewness of local phylogenetic originality changed accordingly. The interaction between urbanization and species richness was not significant, and models were simplified to the main effects only, except for the skewness of the local phylogenetic originality. A significant interaction effect on the skewness of local phylogenetic originality was such that the positive effect of species richness decreased with urbanization to the point of becoming negative for the highest levels of urbanization (Figure 4D). The skewness of regional trait-based originality changed from near-zero values to positive values with urbanization and species richness (Figure 4B). For the skewness of regional phylogenetic originality, only species richness had a significant positive effect (Figure 4C). The sampling year was not linked with trait-based and phylogenetic originality skewness at both scales. Finally, the species richness of the plots increased along the urbanization gradient and did not change significantly across the years.

### **Species Level Metrics**

For each species, we calculated its local and regional trait-based and phylogenetic originalities, which allowed the identification of the most original species. Among them, we found common species such as the annual bluegrass *Poa annua* and the dandelion *Taraxacum* sect. *Ruderalia* (Table S3). When we modeled the species-specific originalities directly, only local originality was investigated. The trait-based originality for each species increased significantly with both urbanization and species richness and decreased across the years. The significant interaction between urbanization and species richness on the species-specific trait-based originality decreased the positive urbanization effect in plots with greater species richness (Figure S2A). Species phylogenetic originality was not influenced by urbanization but increased with species richness and decreased with the sampling year. There was, however, a significant effect of the interaction

between urbanization and species richness on species phylogenetic originality; the positive effect of species richness decreased in the most urbanized plots (Figure S2B). Finally, Spearman's correlation was not significant between the per-species local phylogenetic and trait-based originalities ( $r = 0.04$ ,  $p = 0.20$ ).

The ratio of local to regional trait-based originality increased significantly with both urbanization and species richness. Due to the interaction with species richness, the positive effect of urbanization on the trait-based ratio of originality decreased in richer plots and *vice versa* (Figure S3A). The sampling year was excluded from the model after AIC backward variable selection. We found that Spearman's correlation between the mean of the local trait-based originalities of a species and its regional trait-based score was highly significant ( $r = 0.78$ ,  $p < 0.001$ ). Finally, species urbanity increased with the local trait-based originalities of native and non-native species (Figure S4A, Spearman's correlation coefficient  $r = 0.17$ ,  $p < 0.001$ ) but not with the regional originality ( $r = -0.01$ ,  $p = 0.80$ ). Species urbanity was higher in species that became more original locally than they were regionally (Figure S4A).

Concerning the ratio of local to regional phylogenetic originality, no effect of urbanization alone was detected, but there was a significant increase with species richness. Moreover, there was a significant effect of the interaction between urbanization and species richness on the ratio of local to regional phylogenetic originality; the positive effect of species richness decreased in the most urbanized plots (Figure S3B). The phylogenetic originality ratio decreased with the sampling year. We also found that the most phylogenetically original species were not the same at the local and regional scales. In addition, Spearman's correlation was low and not significant ( $r = -0.06$ ,  $p = 0.11$ ) between the average of the local phylogenetic originalities of a species and its regional phylogenetic originality score. Finally, species urbanity was higher in less phylogenetically original species at the regional scale in both native and non-native species (Spearman's correlation coefficient  $r = -0.15$ ,  $p < 0.001$ , Figure S4B).

## Dataset without Non-native Species

Removing 85 non-native species (14.6% of the total regional pool), which occurred in 709 out of 2362 plots, had a minimal effect on the results of the previous models (Table S4). The only differences we found were that 1. the year of sampling had a significant negative effect on mean regional trait-based originality (the effect was not significant with the complete dataset); 2. inversely, it had no significant effect on the mean regional phylogenetic originality (positive effect with the complete dataset); and 3. the skewness of the local phylogenetic originality was not impacted by species richness (positive effect with the complete dataset).

At the local scale, non-native species were significantly more original than native species due to their trait values (Wilcoxon paired rank test  $W=5819854$ ,  $p < 0.001$ , Figure S5A) and less original due to their phylogeny ( $W=11347166$ ,  $p < 0.001$ , Figure S5B). Twelve out of thirty species in the top 5% of trait-based original species at the local scale were non-native. At the regional level, non-native species were significantly more original than native species in their traits ( $W = 18069$ ,  $p < 0.001$ , Figure S5C) and not significantly different by their phylogenetic relations ( $W = 24375$ ,  $p = 0.62$ , Figure S5D).

## DISCUSSION

Urbanization affects ecosystems at multiple scales (Aronson et al. 2016) by changing community phylogenetic structure (Ricotta et al., 2009; Knapp et al., 2017; Sol et al., 2017) and species trait distributions (Williams et al., 2015). While urban ecosystem structure is often studied at the community scale, it is still difficult to evaluate the role and contribution of each species to local diversity. We provide a promising tool that relies on the originality values calculated for each species relative to the co-occurring species within its community. Our approach is conceptually simple: species originality is measured as an average distance from a focal species to all the others in the community. It has recently been proposed to use the originality indices for diversity evaluations (see Kondratyeva et al., 2019, section II.3.b). Species originality indices can indeed be used for evaluation of local diversity by calculating the mean of species originality scores in a local set of

species (Ricotta et al., 2016). Thus, our study is a first application of this framework to the urban context that considers the two spatial scales, which are the area over which originality is calculated.

### **Trait-based Originality of Urban Communities**

In our case study, the regional trait-based originality of a species increased the local trait-based diversity of communities with increasing urbanization. In other words, species that were highly original at the regional scale were also the most original at the local scale relative to their trait-based dissimilarities, and original species were more frequent in urbanized areas than in non-urbanized areas. Species in urbanized areas were even more original at the local scale than at the regional scale, rendering them the unique representatives of some traits. However, the most important ecosystem processes are usually ensured by common redundant species (Mouillot et al., 2013b; Brandl et al., 2016). To a certain level of species richness, plant community trait-based diversity (and species trait-based originality) continued to increase with urbanization. Beyond that level, the number of redundant species started to increase, accordingly leading to positive skewness of the originality values and making the communities more resilient to environmental changes (McLean et al., 2019). Therefore, in more-urbanized areas, a few species with high trait-based originality contributed to the increase of community trait-based diversity but did not participate much in ecosystem stability. However, throughout the years of the study, trait-based originality decreased at both scales, and fewer species became original at the local scale compared to at the regional scale. Urban land modification may intensify with time and indicate a growing biotic homogenization of local plant communities (McKinney, 2006).

The species with the highest trait-based originality occurred rarely, only appearing once or twice in the whole 10-year dataset. However, three of these species have been found in many different communities and in the top 5% of original species: *Erigeron sumatrensis*, which is a highly original non-native species with high urbanity; and species of the *Taraxacum* sect. *Ruderalia* and *Poa annua*, both of which are native and highly common in the Île-de-France region. Thus, these three widespread species greatly contributed to the trait-based diversity of local communities and of the

whole region. At the same time, the omnipresence of the same species in urban habitats can lead to biological homogenization, such shown in birds (Devictor et al., 2008) and plants (McKinney, 2006) across cities. Indeed, *T. sect. ruderalia* and *P. annua* are urbanoneutral species (Hill et al., 2002); in other words, they are less affected than other species by urbanization disturbances (CBNBP, 2016). Such highly common and dominant species could take over the majority of urban ecosystem resources, pushing other less frequent species to adapt different trait values, giving the urbanoneutral species higher trait-based originality.

### **Phylogenetic Originality of Urban Communities**

The patterns we obtained with phylogenetic originality were different from those obtained with trait data. Knapp et al. (2008) found that urbanophilic species with well-adapted traits for urban environmental conditions are phylogenetically more distinct in urban areas than maladapted species. In our dataset, a highly phylogenetically original plant species was not necessarily original in its traits, and *vice versa*. This corresponds to findings of Lososová et al. (2016) that phylogenetic diversity is only a weak proxy for the functional diversity of urban plant communities. Overall, the use of phylodiversity as a proxy for trait-based diversity (feature diversity, Faith, 1992) has been recently challenged (Mouquet et al., 2012; Cadotte et al., 2017) and has received mixed empirical evidence (e.g., Mazel et al., 2017, 2018, 2019; but see Owen et al., 2019). However, there is emerging evidence of rapid urbanization-driven trait evolution that stresses the need to include the eco-evolutionary implications of urbanization, especially for species that may play key roles in ecosystem functioning (Alberti, 2015).

Contrary to the trends seen in trait-based originality, local and regional phylogenetic originalities were not correlated in our dataset. There was no straightforward evidence for an effect of urbanization on the phylogenetic diversity calculated with local originality values. It seems that urbanization levels predominantly affect the contribution of local plots to the regional phylodiversity. This is in line with the results of Knapp et al. (2012), who found a decrease in phylodiversity on a large scale but no clear trend in phylodiversity on the small scale of household

yards across a gradient of housing density in the US. Therefore, we proposed to study the effects of urbanization on herbaceous plant community phylodiversity through regionally measured originality, where all relations in the species pool were considered. In our data, more-related species (less original) at the regional scale constituted local urban communities of the Île-de-France region, thus decreasing regional phylodiversity at the local scale. Many species became phylogenetically original in rich local communities compared to their regional originality; however, increasing urbanization slowed down this increase. A community with a few distantly related species can have a higher phylodiversity than richer communities with more closely related species (Knapp et al., 2012). Indeed, in our dataset, the mean regional phylogenetic originality represented in a single plot decreased with growing species richness. Moreover, there was a global gain in regional phylogenetic originality through the years, but a loss of phylogenetically original species at a local scale, again showing the importance of the spatial scale of a study.

Species found in the top 5% of the most phylogenetically original species at local and regional scales were mostly monocotyledons (Table S3). This is primarily an effect of the lower number of monocotyledon species than dicotyledon species in the regional pool. Because of the large phylogenetic age of the split between these two plant groups (Chaw et al., 2004), a lower number of monocotyledons makes these species more distantly related on average to all species other than those that are classified as dicotyledons.

### **Species Richness and Non-native Species Originality**

Even if our 586 species set was a sample of the whole Île-de-France region flora, where more than 2064 vascular plant species have been recorded (CBNBP, 2016), we found a broadly observable pattern of species richness increase in local plots with urbanization both in plots with and without non-native species. Non-native species are often promoted in areas disturbed by urbanization (Kühn et al., 2017). Nevertheless, non-native species did not contribute more to the community diversity, and their removal in our data hardly influenced community structure, although non-native species were more functionally original compared to native species at both the local and regional

levels. Non-native species are also filtered by urbanization and have specific environmental preferences, making them rare specialists within many urban habitat patches (McKinney, 2006; Godefroid and Ricotta, 2018). In contrast, the observed loss in mean regional phylogenetic originality in urbanized areas could not be outweighed by the presence of non-native species, which are less phylogenetically original than natives. Similarly, Knapp et al. (2017) found that changes in the native flora of Halle (Germany), which caused a loss of phylogenetic diversity, could not be compensated by the introduction of non-native species. Moreover, the presence of non-native species decreased the phylodiversity of plant communities in Ricotta et al. (2009) and Čeplová et al. (2015) and the similarly for bird communities where exotic species could not completely compensate for the loss of phylogenetic diversity resulting from the loss of native species (Sol et al., 2017). Finally, our results are also in line with the trend of non-native species promoting homogenization globally (Winter et al., 2009) and contrast the findings of Kühn and Klotz (2006), which indicated that after 1500 AD, non-native species promoted differentiation of the plant species assemblages in cities.

## **CONCLUSION AND DIRECTIONS FOR FURTHER RESEARCH**

Cities are complex urban ecosystems, and their functioning tends to be different than that of natural ecosystems because of the human element (Rebele, 1994; Grimm et al., 2000). Expanding ecological knowledge and expertise can enhance our understanding of the responses of biodiversity to urbanization. Local ecological knowledge is required for more targeted urban management (McPhearson et al., 2016). Originality measures determine which species contribute the most to local and regional diversities. Applied to the urban context, originality measures can easily evaluate local biodiversity by integrating multiscale biological species information.

Our originality framework helps to determine the most relevant spatial scale for efficient urban management (e.g., plant phylogenetic originality was influenced by urbanization on a regional scale only). Moreover, the originality framework determines which original or redundant species are responsible for variations in trait-based diversity and phylodiversity. Here, we showed notably that



urbanization causes some species to have higher trait-based originality at the local community scale than at the regional scale. Although we observed that original species were more frequent in urbanized areas, we also showed that this originality decreased over time, resulting in less diverse communities and regions. Moreover, we showed that the increase in trait-based originality in urbanized areas may be partly hampered by the presence of redundant species in the richest plots. In our case study, the phylogenetic diversity of local plots was not impacted by urbanization. However, species in local urbanized plots tended to be among the least regionally original species, implying a negative effect of urbanization on regional phylogenetic diversity.

Although our study demonstrates the potential of the originality framework, we do not pretend to affirm any generalizations about urban plant communities based on our case study. Here, the urbanization gradient was characterized by a single measure, i.e., the proportion of urbanized areas (including individual and collective housing, industrial and business areas, facilities, quarries, dumps and worksites and transportation infrastructure). Although this urbanization measure combines diverse types of urban infrastructures and impervious surfaces, it does not distinguish among the range of urban microhabitats and does not account for many urbanization-related changes, such as soil compaction, eutrophication and air, soil and water pollution (McDonnell and Hahs, 2008). Therefore, future urban ecology studies could use multidimensional composite measures of urbanization to provide a more holistic understanding of the effects of urbanization on biological diversity (Moll et al., 2019). In addition, our dataset did not allow us to account for species abundance, which can drastically change conclusions about ecosystem functioning, where a high number of individuals supporting the same trait can ensure greater stability (Walker et al., 1999; D'agata et al., 2016). The weighting of originality measure with species relative abundances could address such an issue, thus regulating the relative importance given to rare vs. abundant species (Kondratyeva et al., 2019).

To our knowledge, originality measures have usually been employed for conservation purposes on a global scale (e.g., Jetz et al., 2014; Redding et al., 2015a; Thévenin et al., 2018). Here, we applied

them at the local scale of urban plant communities. Future research could now aim to develop cross-city comparative research on species originality using urban systems networks, such as the Urban Biodiversity Research Coordination Network (UrBioNet, <http://urbionet.weebly.com/>). Citizen science programs, which are deployed in urban areas, are also a good starting base for exchange in sustainable city development, which include scientists, policymakers and citizens. To complement these programs, we showed that our conceptual framework of originality measures deserves more attention by city planners working tightly with urban ecologists.

### **CONFLICT OF INTEREST STATEMENT**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

### **AUTHOR CONTRIBUTIONS**

AK and SP conceived the ideas of the research presented herein. AK, SP, and IK designed the analytical methods. WD and AK assembled the project phylogeny. AK, WD and IK assembled the trait data. NM, GM, EM, and JV supervised the Vigie-flore data collection and management. AK analyzed the data and wrote the first version of the manuscript. All authors contributed to the final manuscript edition and approved the submitted version.

### **FUNDING**

This work was funded by a grant from the Agence Nationale de la Recherche under the LabEx BCDiv project reference 10-LABX-0003 within the framework of the program 'Investing for the future' (11-IDEX-0004).

### **ACKNOWLEDGEMENTS**

We thank Emmanuelle Porcher for the valuable comments on previous versions of the manuscript and all Vigie-flore volunteers for data collection. The study has been supported by the TRY initiative on Plant Traits (<http://www.try-db.org>). The TRY initiative and database is hosted, developed and maintained by J. Kattge and G. Boenisch (Max Planck Institute for

Biogeochemistry, Jena, Germany). TRY is currently supported by Future Earth/bioDISCOVERY and the German Center for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig.

## **DATA AVAILABILITY**

This study used plant trait values from several datasets that can be found online:

BioFlor: data available under registration at [www.biolflor.de](http://www.biolflor.de)

LEDA: available at <http://www.leda-traitbase.org>

Ecoflora: <http://www.ecoflora.co.uk>

Catminat: <http://philippe.julve.pagesperso-orange.fr/catminat.htm>.

TRY initiative and database: <http://www.try-db.org>

Seed Information Database. <http://data.kew.org/sid/> (September 2019)

DaPhnE phylogeny: <http://esapubs.org/archive/ecol/E093/214/>

CBNBP: <http://cbnbp.mnhn.fr/cbnbp/>

The following datasets are not publicly available. Requests to access these datasets should be directed to [assad.alicherif@iau-idf.fr](mailto:assad.alicherif@iau-idf.fr) for land cover data of the Île-de-France region and to [vigie-flore@mnhn.fr](mailto:vigie-flore@mnhn.fr) for Vigie-flore plant occurrence data.

## REFERENCES

- Ahern, J. (2013). Urban landscape sustainability and resilience: The promise and challenges of integrating ecology with urban planning and design. *Landsc. Ecol.* 28, 1203–1212. doi:10.1007/s10980-012-9799-z
- Alberti, M. (2015). Eco-evolutionary dynamics in an urbanizing planet. *Trends Ecol. Evol.* 30, 114–126. doi:10.1016/j.tree.2014.11.007
- Anselin, L. (1995). Local Indicators of Spatial Association-LISA. *Geogr. Anal.* 27, 93–116.
- Aronson, M. F. J., Nilon, C. H., Lepczyk, C. A., Parker, T. S., Warren, P. S., Cilliers, S. S. et al. (2016). Hierarchical filters determine community assembly of urban species pools. *Ecology* 97, 2952–2963. doi: 10.1002/ecy.1535.
- Aronson, M. F. J., La Sorte, F. A., Nilon, C. H., Katti, M., Goddard, M. A., Lepczyk, C. A, et al. (2014). A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proc. R. Soc. B* 281, 20133330. doi:10.1098/rspb.2013.3330
- Bjørnstad, O. N., and Falck, W. (2001). Nonparametric spatial covariance functions: Estimation and testing. *Environ. Ecol. Stat.* 8, 53–70. doi:10.1023/A:1009601932481.
- Brandl, S. J., Emslie, M. J., and Ceccarelli, D. M. (2016). Habitat degradation increases functional originality in highly diverse coral reef fish assemblages. *Ecosphere* 7, e01557. doi:10.1002/ecs2.1557
- Cadotte, M. W., Davies, T. J., and Peres-Neto, P. R. (2017). Why phylogenies do not always predict ecological differences. *Ecol. Monogr.* 87, 535–551. doi:10.1002/ecm.1267.
- Čeplová, N., Lososová, Z., Zelený, D., Chytrý, M., Danihelka, J., Fajmon, K., et al. (2015). Phylogenetic diversity of central-European urban plant communities: effects of alien species and habitat types. *Preslia* 87, 1–16.
- Cervelli, E. W., Lundholm, J. T., and Du, X. (2013). Spontaneous urban vegetation and habitat heterogeneity in Xi'an, China. *Landsc. Urban Plan.* 120, 25–33. doi:10.1016/j.landurbplan.2013.08.001
- Chave, J. (2013). The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecol. Lett.* 16, 4–16. doi:10.1111/ele.12048
- Chaw, S. M., Chang, C. C., Chen, H. L., Li, W. H. (2004) Dating the Monocot–Dicot Divergence and the Origin of Core Eudicots Using Whole Chloroplast Genomes. *J. Mol. Evol.* 58, 424–441. doi : 10.1007/s00239-003-2564-9
- CBNBP (2016). Catalogue de la flore de l'Île-de-France, version mai 2016. <http://cbnbp.mnhn.fr/cbnbp/ressources/ressources.jsp>

- Couvet, D., and Prevot, A. C. (2015). Citizen-science programs: Towards transformative biodiversity governance. *Environ. Dev.* 13, 39–45. doi:10.1016/j.envdev.2014.11.003
- D'agata, S., Vigliola, L., Graham, N. A. J., Wantiez, L., Parravicini, V., Villéger, S., et al. (2016). Unexpected high vulnerability of functions in wilderness areas: evidence from coral reef fishes. *Proc. R. Soc. B Biol. Sci.* 283, 20160128. doi:10.1098/rspb.2016.0128
- Deguines, N., Julliard, R., de Flores, M., and Fontaine, C. (2016). Functional homogenization of flower visitor communities with urbanization. *Ecol. Evol.* 6, 1967–1976. doi:10.1002/ece3.2009
- Devictor, V., Julliard, R., Clavel, J., Jiguet, F., Lee, A., and Couvet, D. (2008). Functional biotic homogenization of bird communities in disturbed landscapes. *Glob. Ecol. Biogeogr.* 17, 252–261. doi:10.1111/j.1466-8238.2007.00364.x
- Devictor, V., Julliard, R., Couvet, D., Lee, A., and Jiguet, F. (2007). Functional homogenization effect of urbanization on bird communities. *Conserv. Biol.* 21, 741–751. doi:10.1111/j.1523-1739.2007.00671.x
- Dickinson, J. L., Zuckerberg, B., and Bonter, D. N. (2010). Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annu. Rev. Ecol. Evol. Syst.* 41, 149–172. doi:10.1146/annurev-ecolsys-102209-144636
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., et al. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36, 027–046. doi:10.1111/j.1600-0587.2012.07348.x
- Durka, W., and Michalski, S. G. (2012). Daphne: a dated phylogeny of a large European flora for phylogenetically informed ecological analyses. *Ecology* 93, 2297–2297. doi:10.1890/12-0743.1
- Elands, B. H. M., Wiersum, K. F., Buijs, A. E., and Vierikko, K. (2015). Policy interpretations and manifestation of biocultural diversity in urbanized Europe: conservation of lived biodiversity. *Biodivers. Conserv.* 24, 3347–3366. doi:10.1007/s10531-015-0985-6
- Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W., Paulissen, D. (1991). Zeigerwerte von Pflanzen in Mitteleuropa. *Scr. Geobot.* 188, 1–248.
- European Commission's Directorate-General Environment (2012). "The Multifunctionality of Green Infrastructure" in Science for Environment Policy, In-depth Reports. Bristol, 1–37.
- Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61, 1–10. doi: 10.1016/0006-3207(92)91201-3
- Farris, J. S. (1972). Estimating phylogenetic trees from distance matrices. *Amer. Nat.* 106, 645–

668.

Fitter, A. H., and Peat, H. J. (1994). The Ecological Flora Database. *J. Ecol.* 82, 415–425.  
Gargominy, O. 2008. *TAXREF v2.0*. Muséum national d'Histoire naturelle, Paris. Archive de téléchargement contenant 6 fichiers.

Garnier, E., Stahl, U., Laporte, M. A., Kattge, J., Mougenot, I., Kühn, I., et al. (2017). Towards a thesaurus of plant characteristics: an ecological contribution. *J. Ecol.* 105, 298–309. doi:10.1111/1365-2745.12698

Gaston, K. (Ed.). (2010). *Urban Ecology (Ecological Reviews)*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511778483

Gill, S. E., Handley, J. F., Ennos, A. R., and Pauleit, S. (2007). Adapting cities for climate change: The role of the green infrastructure. *Built Environ.* 33, 115–133. doi:10.2148/benv.33.1.115

Godefroid, S., and Ricotta, C. (2018). Alien plant species do have a clear preference for different land uses within urban environments. *Urban Ecosyst.* 21, 1189–1198. doi:10.1007/s11252-018-0792-4

Godefroid, S., and Koedam, N. (2007). Urban plant species patterns are highly driven by density and function of built-up areas. *Landsc. Ecol.* 22, 1227–1239. doi:10.1007/s10980-007-9102-x

Grime, J. P., Hodgson, J. G., and Hunt, R. (1988). *Comparative plant ecology: a functional approach to common British species*. London, UK: Chapman & Hall.

Grimm, N. B., Grove, J. M., Pickett, S. T. A., and Redman, C. A. (2000). Integrated approaches to long-term studies of urban ecological systems. *BioScience* 50: 571–584. doi:10.1641/0006-3568(2000)050[0571:IA TLTO]2.0.CO;2

Gross, N., Börger, L., Duncan, R. P., and Hulme, P. E. (2013). Functional differences between alien and native species: Do biotic interactions determine the functional structure of highly invaded grasslands? *Funct. Ecol.* 27, 1262–1272. doi:10.1111/1365-2435.12120

Guetté, A., Gaüzère, P., Devictor, V., Jiguet, F., and Godet, L. (2017). Measuring the synanthropy of species and communities to monitor the effects of urbanization on biodiversity. *Ecol. Indic.* 79, 139–154. doi:10.1016/j.ecolind.2017.04.018

Hill, M. O., Roy, D. B., and Thompson, K. (2002). Hemeroby, urbanity and ruderality: Bioindicators of disturbance and human impact. *J. Appl. Ecol.* 39, 708–720. doi:10.1046/j.1365-2664.2002.00746.x

IAU. (2019). Population des communes d'Île-de-France (Donnée INSEE). Available online at: <http://data.iau-idf.fr/>

Isaac, N. J. B., Turvey, S. T., Collen, B., Waterman, C., and Baillie, J. E. M. (2007). Mammals on the EDGE: Conservation priorities based on threat and phylogeny. PLoS One 2, e296. doi:10.1371/journal.pone.0000296

Ives, C. D., Lentini, P. E., Threlfall, C. G., Ikin, K., Shanahan, D. F., Garrard, G. E., et al. (2016). Cities are hotspots for threatened species. Glob. Ecol. Biogeogr. 25, 117–126. doi:10.1111/geb.12404

IPBES. (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. eds. Díaz, S., Settele, J., Brondizio E.S., Ngo, H. T., Guèze, M., Agard, J. et al. IPBES secretariat, Bonn, Germany.

IUCN. (2012). Threats Classification Scheme, version 3.2. Available online: <https://www.iucnredlist.org/resources/threat-classification-scheme>

Jaccard, P. (1901). Etude comparative de la distribution florale dans une portion des Alpes et des Jura. Bulletin de la Société Vaudoise des Sciences Naturelles 37, 547–579.

Jetz, W., Thomas, G. H., Joy, J. B., Redding, D. W., Hartmann, K., and Mooers, A. O. (2014). Global Distribution and Conservation of Evolutionary Distinctness in Birds. Curr. Biol. 24, 919–930. doi:10.1016/j.cub.2014.03.011

Julve, Ph. (1998). ff - Baseflor. Index botanique, écologique et chorologique de la flore de France. Version : « 2019 ». <http://perso.wanadoo.fr/philippe.julve/catminat.htm>

Kabisch, N., Frantzeskaki, N., Pauleit, S., Naumann, S., Davis, M., Artmann, M. et al. (2016). Nature-based solutions to climate change mitigation and adaptation in urban areas: perspectives on indicators, knowledge gaps, barriers, and opportunities for action. Ecol. Soc. 21, 39. doi:10.5751/es-08373-210239

Kalusová, V., Čeplová, N., and Lososová, Z. (2017). Which traits influence the frequency of plant species occurrence in urban habitat types? Urban Ecosyst. 20, 65–75. doi:10.1007/s11252-016-0588-3

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., and Bönsch, G. (2011). TRY – a global database of plant traits. Global Change Biology, 17:2905–2935. doi : 10.1111/j.1365-2486.2011.02451.x

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

Klotz, S., Kühn, I., Durka, W., Eds (2002). *BIOLFLOR - Eine Datenbank zu biologisch-ökologischen*

*Merkmale der Gefäßpflanzen in Deutschland*. Bonn: Bundesamt für Naturschutz.

Knapp, S., Dinsmore, L., Fissore, C., Hobbie, S. E., Jakobsdottir, I., Kattge, J., et al. (2012). Phylogenetic and functional characteristics of household yard floras and their changes along an urbanization gradient. *Ecology* 93, 83–98. doi:10.1890/11-0392.1

Knapp, S., Kühn, I., Stolle, J., and Klotz, S. (2010). Changes in the functional composition of a Central European urban flora over three centuries. *Perspect. Plant Ecol. Evol. Syst.* 12, 235–244. doi:10.1016/j.ppees.2009.11.001

Knapp, S., Kühn, I., Bakker, J. P., Kleyer, M., Klotz, S., Ozinga, W. A., et al. (2009). How species traits and affinity to urban land use control large-scale species frequency. *Divers. Distrib.* 15, 533–546. doi:10.1111/j.1472-4642.2009.00561.x

Knapp, S., Kühn, I., Schweiger, O., and Klotz, S. (2008). Challenging urban species diversity: Contrasting phylogenetic patterns across plant functional groups in Germany. *Ecol. Lett.* 11, 1054–1064. doi:10.1111/j.1461-0248.2008.01217.x

Knapp, S., Winter, M., and Klotz, S. (2017). Increasing species richness but decreasing phylogenetic richness and divergence over a 320-year period of urbanization. *J. Appl. Ecol.* 54, 1152–1160. doi:10.1111/1365-2664.12826

Kobori, H., Dickinson, J. L., Washitani, I., Sakurai, R., Amano, T., Komatsu, N., et al. (2016). Citizen science: a new approach to advance ecology, education, and conservation. *Ecol. Res.* 31, 1–19. doi:10.1007/s11284-015-1314-y

Kondratyeva, A., Grandcolas, P., and Pavoine, S. (2019). Reconciling the concepts and measures of diversity, rarity and originality in ecology and evolution. *Biol. Rev.* 94, 1317–1337. doi:10.1111/bry.12504

Kowarik, I. (2011). Novel urban ecosystems, biodiversity, and conservation. *Environ. Pollut.* 159, 1974–1983. doi:10.1016/j.envpol.2011.02.022

Kühn, I., Brandl, R., and Klotz, S. (2004). The flora of German cities is naturally species rich. *Evol. Ecol. Research* 6, 749–764.

Kühn, I., and Klotz, S. (2006). Urbanization and homogenization - Comparing the floras of urban and rural areas in Germany. *Biol. Conserv.* 127, 292–300. doi:10.1016/j.biocon.2005.06.033

Kühn, I., Wolf, J., and Schneider, A. (2017). Is there an urban effect in alien plant invasions? *Biol. Invasions* 19, 3505–3513. doi:10.1007/s10530-017-1591-1

Laliberté, E., and Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits A distance-based framework for measuring from multiple traits



functional diversity. *Ecology* 91, 299–305. doi:10.1890/08-2244.1

Levin, S. A. (1992). The problem of pattern and scale of ecology. *Ecology* 73, 1943-1967. <https://doi.org/10.2307/1941447>

Lieth, H. (1975). "Primary production of the major vegetation units of the world" in *Primary Productivity of the Biosphere. Ecological Studies (Analysis and Synthesis)*, vol 14, eds. H. Lieth and R. H. Whittaker (Berlin, Heidelberg: Springer), 203-215. doi: [https://doi.org/10.1007/978-3-642-80913-2\\_10](https://doi.org/10.1007/978-3-642-80913-2_10)

Lososová, Z., Čeplová, N., Chytrý, M., Tichý, L., Danihelka, J., Fajmon, K., et al. (2016). Is phylogenetic diversity a good proxy for functional diversity of plant communities? A case study from urban habitats. *J. Veg. Sci.* 27, 1036–1046. doi:10.1111/jvs.12414

Lundholm, J. T., Weddle, B. M., and Macivor, J. S. (2014). Snow depth and vegetation type affect green roof thermal performance in winter. *Energy Build.* 84, 299–307. doi:10.1016/j.enbuild.2014.07.093

Martin, G., Devictor, V., Motard, E., Machon, N., and Porcher, E. (2019). Short-term climate-induced change in French plant communities. *Biol. Lett.* 15, 20190280. doi:10.1098/rsbl.2019.0280

Mason, N. W. H., Mouillot, D., Lee, W. G., and Wilson, J. B. (2005). Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos* 111, 112–118. doi: 10.1111/j.0030-1299.2005.13886.x

Mazel, F., Pennell, M. W., Cadotte, M. W., Diaz, S., Dalla Riva, G. V., Grenyer, R., et al. (2018). Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nat. Commun.* 9, 2888. doi:10.1038/s41467-018-05126-3

Mazel, F., Mooers, A., Riva, G. V. D., and Pennell, M. W. (2017). Conserving phylogenetic diversity can be a poor strategy for conserving functional diversity. *Syst. Biol.* 66, 1019–1027. doi:10.1093/sysbio/syx054

McDonnell, M. J., and Hahs, A. K. (2008). The use of gradient analysis studies in advancing our understanding of the ecology of urbanizing landscapes: Current status and future directions. *Landsc. Ecol.* 23, 1143–1155. doi:10.1007/s10980-008-9253-4

McKinley, D. C., Miller-Rushing, A. J., Ballard, H. L., Bonney, R., Brown, H., Cook-Patton, S. C., et al. (2017). Citizen science can improve conservation science, natural resource management, and environmental protection. *Biol. Conserv.* 208, 15–28. doi:10.1016/j.biocon.2016.05.015

McKinney, M. L. (2008). Effects of urbanization on species richness: A review of plants and animals. *Urban Ecosyst.* 11, 161–176. doi:10.1007/s11252-007-0045-4

McKinney, M. L. (2006). Urbanization as a major cause of biotic homogenization. *Biol. Conserv.* 127, 247–260. doi:10.1016/j.biocon.2005.09.005

McKinney, M. L. (2002) Urbanization, Biodiversity, and Conservation: The impacts of urbanization on native species are poorly studied, but educating a highly urbanized human population about these impacts can greatly improve species conservation in all ecosystems. *BioScience* 52, 883–890, doi:10.1641/0006-3568(2002)052[0883:UBAC]2.0.CO;2

McLean, M., Auber, A., Graham, N. A. J., Houk, P., Villéger, S., Violle, C., et al. (2019). Trait structure and redundancy determine sensitivity to disturbance in marine fish communities. *Glob. Chang. Biol.*, In press. doi:10.1111/gcb.14662.

McPhearson, T., Pickett, S. T. A., Grimm, N. B., Niemelä, J., Alberti, M., Elmqvist, T., et al. (2016). Advancing Urban Ecology toward a Science of Cities. *Bioscience* 66, 198–212. doi:10.1093/biosci/biw002

Millar, R. B. (2011). *Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB*. John Wiley & Sons.

MEA (2005). *Ecosystems and Human Wellbeing: Current State and Trends: Findings of the Condition and Trends Working Group*. Washington: Island Press. <http://www.millenniumassessment.org>.

Moll, R. J., Cepek, J. D., Lorch, P. D., Dennis, P. M., Tans, E., Robison, T., et al. (2019). What does urbanization actually mean? A framework for urban metrics in wildlife research. *J. Appl. Ecol.* 56, 1289–1300. doi:10.1111/1365-2664.13358.

Morelli, F., Benedetti, Y., Ibáñez-Álamo, J. D., Jokimäki, J., Mänd, R., Tryjanowski, P., et al. (2016). Evidence of evolutionary homogenization of bird communities in urban environments across Europe. *Glob. Ecol. Biogeogr.* 25, 1284–1293. doi:10.1111/geb.12486

Mouillot, D., Bellwood, D. R., Baraloto, C., Chave, J., Galzin, R., Harmelin-Vivien, M., et al. (2013a). Rare Species Support Vulnerable Functions in High-Diversity Ecosystems. *PLoS Biol.* 11. doi:10.1371/journal.pbio.1001569

Mouillot, D., Graham, N. A. J., Villéger, S., Mason, N. W. H., and Bellwood, D. R. (2013b). A functional approach reveals community responses to disturbances. *Trends Ecol. Evol.* 28, 167–177. doi:10.1016/j.tree.2012.10.004

Mouquet, N., Devictor, V., Meynard, C. N., Munoz, F., Bersier, L. F., Chave, J., et al. (2012). Ecophylogenetics: Advances and perspectives. *Biol. Rev.* doi:10.1111/j.1469-185X.2012.00224.x

Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R<sup>2</sup> from

generalized linear mixed-effects models. *Methods Ecol. Evol.* 4, 133–142. doi:10.1111/j.2041-210x.2012.00261.x

Neil, K., and Wu, J. (2006). Effects of urbanization on plant flowering phenology: A review. *Urban Ecosyst.* 9, 243–257. doi:10.1007/s11252-006-9354-2

Nelson, C.H., and Van Horn, G.S. (1975). A new simplified method for constructing Wagner networks and the cladistics of *Pentachaeta* (Compositae, Astereae). *Brittonia* 27, 362–372. doi:10.2307/2805515

Owen, N.R., Gumbs, R., Gray, C.L., and Faith, D.P. (2019). Global conservation of phylogenetic diversity captures more than just functional diversity. *Nat. Commun.* 10, 859. doi:10.1038/s41467-019-08600-8

Palma, E., Catford, J. A., Corlett, R. T., Duncan, R. P., Hahs, A. K., McCarthy, M. A., et al. (2017). Functional trait changes in the floras of 11 cities across the globe in response to urbanization. *Ecography* 40, 875–886. doi:10.1111/ecog.02516

Pautasso, M., Böhning-Gaese, K., Clergeau, P., Cueto, V. R., Dinetti, M., Fernández-Juricic, E., et al. (2011). Global macroecology of bird assemblages in urbanized and semi-natural ecosystems. *Glob. Ecol. Biogeogr.* 20, 426–436. doi:10.1111/j.1466-8238.2010.00616.x

Pautasso, M. (2007) Scale dependence of the correlation between human population presence and vertebrate and plant species richness. *Ecol Letters.* 10, 16-24.

Pavoine, S. (2018). *adiv: Analysis of Diversity*. R package version 1.2. <https://CRAN.R-project.org/package=adiv>

Pavoine, S., Bonsall, M. B., Dupaix, A., Jacob, U., and Ricotta, C. (2017). From phylogenetic to functional originality: Guide through indices and new developments. *Ecol. Indic.* 82, 196–205. doi:10.1016/j.ecolind.2017.06.056.

Pavoine, S., and Bonsall, M. B. (2011). Measuring biodiversity to explain community assembly: A unified approach. *Biol. Rev.* 86, 792–812. doi:10.1111/j.1469-185X.2010.00171.x

Pavoine, S., Vallet, J., Dufour, A.-B., Gachet, S., and Daniel, H. (2009). On the challenge of treating various types of variables: application for improving the measurement of functional diversity. *Oikos* 118: 391–402. doi:10.1111/j.1600-0706.2009.16668.x

Pavoine, S., Ollier, S. and Dufour, A.B. (2005). Is the originality of a species measurable? *Ecol. Lett.* 8, 579-586. doi: 10.1111/j.1461-0248.2005.00752.x

Petchey, O. L., and Gaston, K. J. (2002). Extinction and the loss of functional diversity. *Proc. R. Soc. B Biol. Sci.* 269, 1721–1727. doi:10.1098/rspb.2002.2073

Podani, J. (1999) Extending Gower's general coefficient of similarity to ordinal characters.

Taxon 48, 331–340.

Popescu, A.-A., Huber, K.T., and Paradis, E. (2012). ape 3.0: new tools for distance based phylogenetics and evolutionary analysis in R. *Bioinformatics* 28, 1536–1537. doi:10.1093/bioinformatics/bts184

POWO (2019). Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; <http://www.plantsoftheworldonline.org/> Retrieved January 2019.  
R Development Core Team. 2008 R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. See <http://www.r-project.org/>.

Royal Botanic Gardens Kew. (2019) Seed Information Database (SID). Version 7.1. Available from: <http://data.kew.org/sid/> (September 2019).

Raunkiaer, C. (1934). The life forms of plants and statistical plant geography. Being the collected papers of C. Raunkiaer. Oxford: Clarendon Press.

Rebele, F. (1994). Urban ecology and special features of urban ecosystems. *Global Ecol. And Biogeogr.* 4, 173-187. doi: 10.2307/2997649

Redding, D. W., Mooers; A. O., Şekercioğlu, Ç. H., Collen, B. (2015a). Global evolutionary isolation measures can capture key local conservation species in Nearctic and Neotropical bird communities. *Phil. Trans. R. Soc. B* 370, 20140013. doi:10.1098/rstb.2014.0013

Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Meth. Ecol. Evol.* 3, 217–223. doi:10.1111/j.2041-210X.2011.00169.x

Ricotta, C., de Bello, F., Moretti, M., Caccianiga, M., Ceranolini, B., and Pavoine, S. (2016). Measuring the functional redundancy of biological communities: a quantitative guide. *Methods Ecol. Evol.* 7, 1386–1395. doi:10.1111/2041-210X.12604

Ricotta, C., Godefroid, S., Heathfield, D., and Mazzoleni, S. (2015). Limited evidence of local phylogenetic clustering in the urban flora of Brussels. *Plant Biosyst.* 149, 31–37. doi:10.1080/11263504.2013.809029

Ricotta, C., Heathfield, D., Godefroid, S., and Mazzoleni, S. (2012). The effects of habitat filtering on the phylogenetic structure of the urban flora of Brussels (Belgium). *Community Ecol.* 13, 97–101. doi:10.1556/comec.13.2012.1.12

Ricotta, C., La Sorte, F. A., Pyšek, P., Rapson, G. L., Celesti-Grapow, L., and Thompson, K. (2009). Phyloecology of urban alien floras. *J. Ecol.* 97, 1243–1251. doi:10.1111/j.1365-2745.2009.01548.x.

Robinson, S. L., and Lundholm, J. T. (2012). Ecosystem services provided by urban spontaneous vegetation. *Urban Ecosyst.* 15, 545–557. doi:10.1007/s11252-012-0225-8.

- Silvertown, J. (2016). A new dawn for citizen science. *Trends Ecol. Evol.* 24, 467–471.
- Soanes, K., and Lentini, P. E. (2019). When cities are the last chance for saving species. *Front. Ecol. Environ.* 17, 225–231. doi:10.1002/fee.2032
- Sol, D., Bartomeus, I., González-Lagos, C., and Pavoine, S. (2017). Urbanisation and the loss of phylogenetic diversity in birds. *Ecol. Lett.* 20, 721–729. doi:10.1111/ele.12769
- Sukopp, H., Werner, P. (1983). “Urban environments and vegetation,” in *Man’s Impact on Vegetation*, eds W. Holzner, M.J.A.Werger, I. Ikusima. The Hague: Junk Publishers, 247–260.
- Tews, J., Brose, U., Grimm, V., Tielborger, K., Wichmann, M. C., Schwager, M. et al. (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. *J. Biogeogr.* 31, 79–92. <https://doi.org/10.1046/j.0305-0270.2003.00994.x>
- Thévenin, C., Mouchet, M., Robert, A., Kerbiriou, C., and Sarrazin, F. (2018). Reintroductions of birds and mammals involve evolutionarily distinct species at the regional scale. *Proc. Natl. Acad. Sci.* 115, 3404–3409. doi:10.1073/pnas.1714599115
- Tilman, D., Forest, I., and Cowles, J. M. (2014). Biodiversity and ecosystem functioning. *Annu. Rev. Ecol. Evol. Syst.* 45, 471–493. doi:10.1146/annurev-ecolsys-120213-091917
- Tucker, C. M., Aze, T., Cadotte, M. W., Cantalapiedra, J. L., Chisholm, C., Díaz, S., et al. (2019). Assessing the utility of conserving evolutionary history. *Biol. Rev.* In press. doi:10.1111/brv.12526
- Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., et al. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biol. Rev.* 92, 698–715. doi:10.1111/brv.12252
- Tzoulas, K., and James, P. (2010). Making biodiversity measures accessible to non-specialists: an innovative method for rapid assessment of urban biodiversity. *Urban Ecosyst.* 13, 113–127. doi:10.1007/s11252-009-0107-x
- Vallet, J., Beaujouan, V., Pithon, J., Rozé, F., and Daniel, H. (2010). The effects of urban or rural landscape context and distance from the edge on native woodland plant communities. *Biodivers. Conserv.* 19, 3375–3392. doi:10.1007/s10531-010-9901-2
- Villéger, S., Mason, N. W. H., and Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* 89, 2290–2301. doi:10.1890/07-1206.1
- Walker, B., Kinzig, A., and Langridge, J. (1999). Plant attribute diversity, resilience, and ecosystem function: The nature and significance of dominant and minor species. *Ecosystems* 2, 95–113. doi:10.1007/s100219900062

Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: An example for rain forest. *Am. Nat.* 156, 145–155. doi:10.1086/303378.

Williams, N. S. G., Hahs, A. K., and Vesk, P. A. (2015). Urbanisation, plant traits and the composition of urban floras. *Perspect. Plant Ecol. Evol. Syst.* 17, 78–86. doi:10.1016/j.ppees.2014.10.002.

Williams, N. S. G., Schwartz, M. W., Vesk, P. A., McCarthy, M. A., Hahs, A. K., Clemants, S. E., et al. (2009). A conceptual framework for predicting the effects of urban environments on floras. *J. Ecol.* 97, 4–9. doi:10.1111/j.1365-2745.2008.01460.x

Werner, P., and Zahner, R. (2009) Biological diversity and cities. A review and bibliography. BfN-Skripten 245, 52–129.

Winter, M., Schweiger, O., Klotz, S., Nentwig, W., Andriopoulos, P., Arianoutsou, M., et al. (2009). Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21721–21725. doi:10.1073/pnas.0907088106

United Nations, (2018). World urbanization prospects. The 2018 Revision. Highlights. Available online: <https://population.un.org/wup/Publications/>

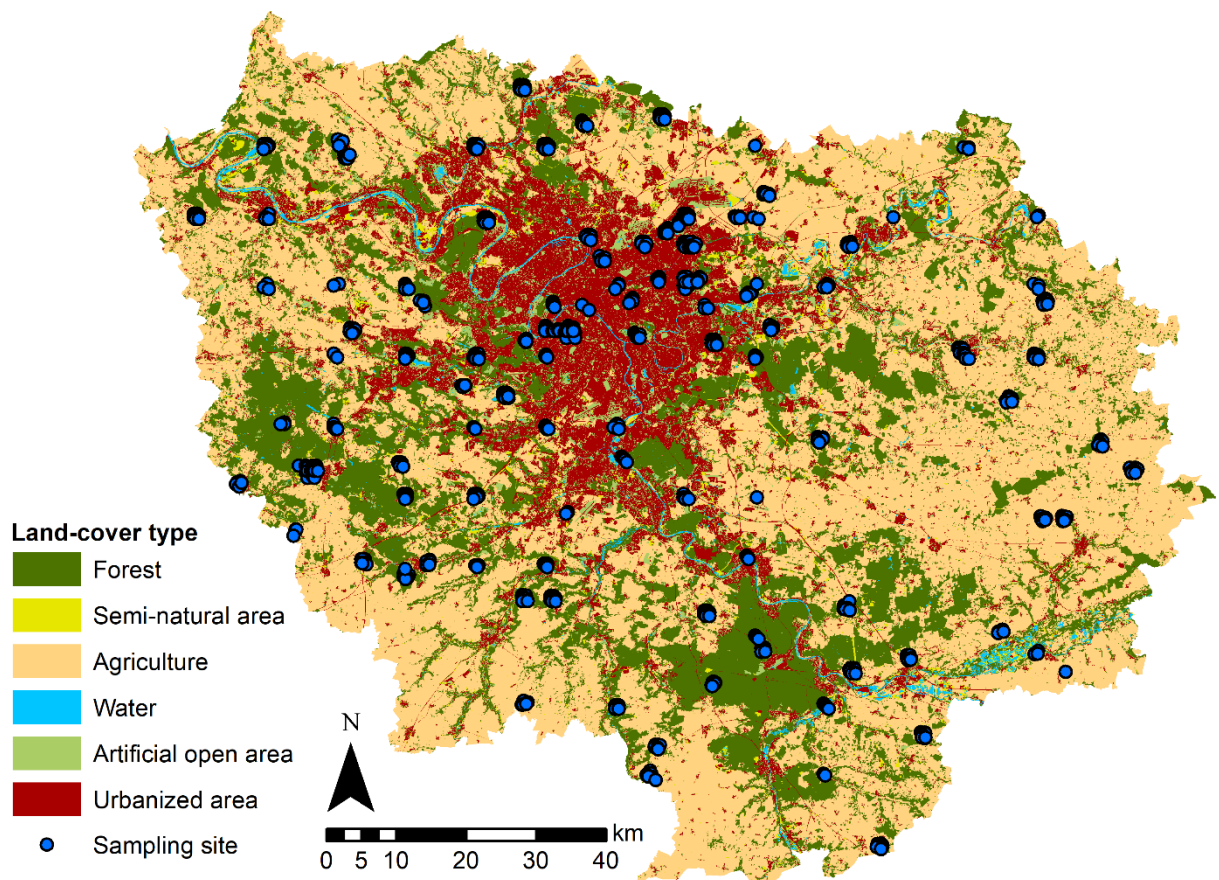
**Table 1** Results for the generalized linear mixed models (GLMM) that were run on the complete dataset (native + non-native species), including 9 community level metrics (means and skewness of local and regional trait-based and phylogenetic originalities and species richness) and 4 species-specific measures (trait-based and phylogenetic originality and local to regional ratios).

The Estimated coefficients and their significance levels are shown for the explanatory variables that remained after backward selection of variables based on Akaike information criterion (AIC) values, including urbanization percentage, species richness, the interaction (U×SR) between urbanization and species richness and the sampling year. The value "excluded" means that the explanatory variable was not retained in the final model according to the Akaike criterion. The conditional and marginal  $R^2$  values describe the proportion of variance explained by the fixed ( $R^2_c$ ) and random+fixed ( $R^2_m$ ) covariables. P-values: <sup>NS</sup> = nonsignificant ( $P > 0.050$ ); \* =  $0.010 < P \leq 0.050$ ; \*\* =  $0.001 < P \leq 0.010$ ; \*\*\* =  $P \leq 0.001$

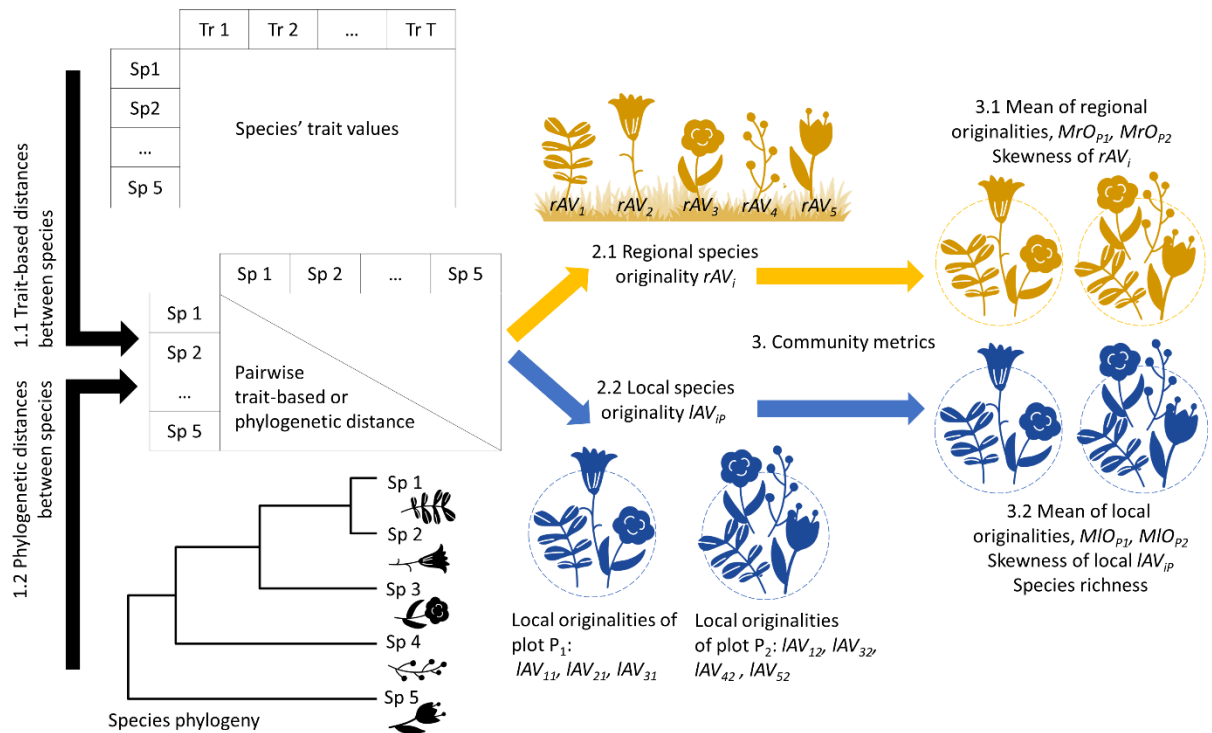
Dependent variable	Intercept	Urbanization (U)	Species richness (SR)	Sampling year	U×SR	R <sup>2</sup> <sub>c</sub>	R <sup>2</sup> <sub>m</sub>
Mean local trait-based originality	4.7 10 <sup>-1</sup> ***	7.7 10 <sup>-4</sup> ***	9.9 10 <sup>-4</sup> ***	-8.3 10 <sup>-4</sup> **	-2.4 10 <sup>-5</sup> ***	0.44	0.14
Mean regional trait-based originality	5.2 10 <sup>-1</sup> ***	2.8 10 <sup>-4</sup> ***	-2.1 10 <sup>-5</sup> NS	-1.3 10 <sup>-4</sup> NS	-9.8 10 <sup>-6</sup> ***	0.64	0.10
Mean local phylogenetic originality	2.3 10 <sup>2</sup> ***	-3.4 10 <sup>-2</sup> NS	9.6 10 <sup>-2</sup> NS	-3.9 10 <sup>-1</sup> NS	-4.9 10 <sup>-3</sup> NS	0.27	0.01
Mean regional phylogenetic originality	2.4 10 <sup>2</sup> ***	-8.6 10 <sup>-2</sup> ***	-2.4 10 <sup>-1</sup> ***	1.3 10 <sup>-1</sup> **	2.2 10 <sup>-3</sup> *	0.63	0.10
Skewness of local trait-based originality	9.2 10 <sup>-2</sup> ***	7.5 10 <sup>-4</sup> NS	2.6 10 <sup>-2</sup> ***	-2.3 10 <sup>-3</sup> NS	excluded	0.22	0.14
Skewness of regional trait-based originality	-2.1 10 <sup>-2</sup> NS	1.5 10 <sup>-3</sup> ***	3.4 10 <sup>-2</sup> ***	-2.7 10 <sup>-3</sup> NS	excluded	0.34	0.20
Skewness of local phylogenetic originality	4.0 10 <sup>-1</sup> ***	6.8 10 <sup>-4</sup> NS	6.5 10 <sup>-3</sup> **	-4.9 10 <sup>-3</sup> NS	-2.3 10 <sup>-4</sup> ***	0.15	0.01
Skewness of regional phylogenetic originality	-4.5 10 <sup>-2</sup> NS	9.2 10 <sup>-4</sup> NS	1.5 10 <sup>-2</sup> ***	2.2 10 <sup>-3</sup> NS	excluded	0.28	0.05
Species richness	1.0 10 <sup>NS</sup>	4.5 10 <sup>-3</sup> ***	omitted	-4.1 10 <sup>-1</sup> NS	omitted	3.7 10 <sup>-2</sup>	2.3 10 <sup>-3</sup>
Trait-based originality	4.8 10 <sup>-1</sup> ***	3.8 10 <sup>-4</sup> ***	9.1 10 <sup>-4</sup> ***	-5.6 10 <sup>-4</sup> ***	-1.1 10 <sup>-5</sup> ***	0.68	0.03
Phylogenetic originality	9.4 10 <sup>-3</sup> NS	8.0 10 <sup>-4</sup> NS	8.4 10 <sup>-3</sup> ***	-6.9 10 <sup>-3</sup> **	-1.6 10 <sup>-4</sup> **	0.46	3.8 10 <sup>-2</sup>
Local/regional trait-based originality	9.1 10 <sup>-1</sup> ***	7.4 10 <sup>-4</sup> ***	1.8 10 <sup>-3</sup> ***	excluded	-2.0 10 <sup>-5</sup> ***	0.46	0.06
Local/regional phylogenetic originality	9.5 10 <sup>-1</sup> ***	1.4 10 <sup>-4</sup> NS	1.2 10 <sup>-3</sup> ***	-9.5 10 <sup>-4</sup> *	-2.4 10 <sup>-5</sup> ***	0.54	2.9 10 <sup>-2</sup>



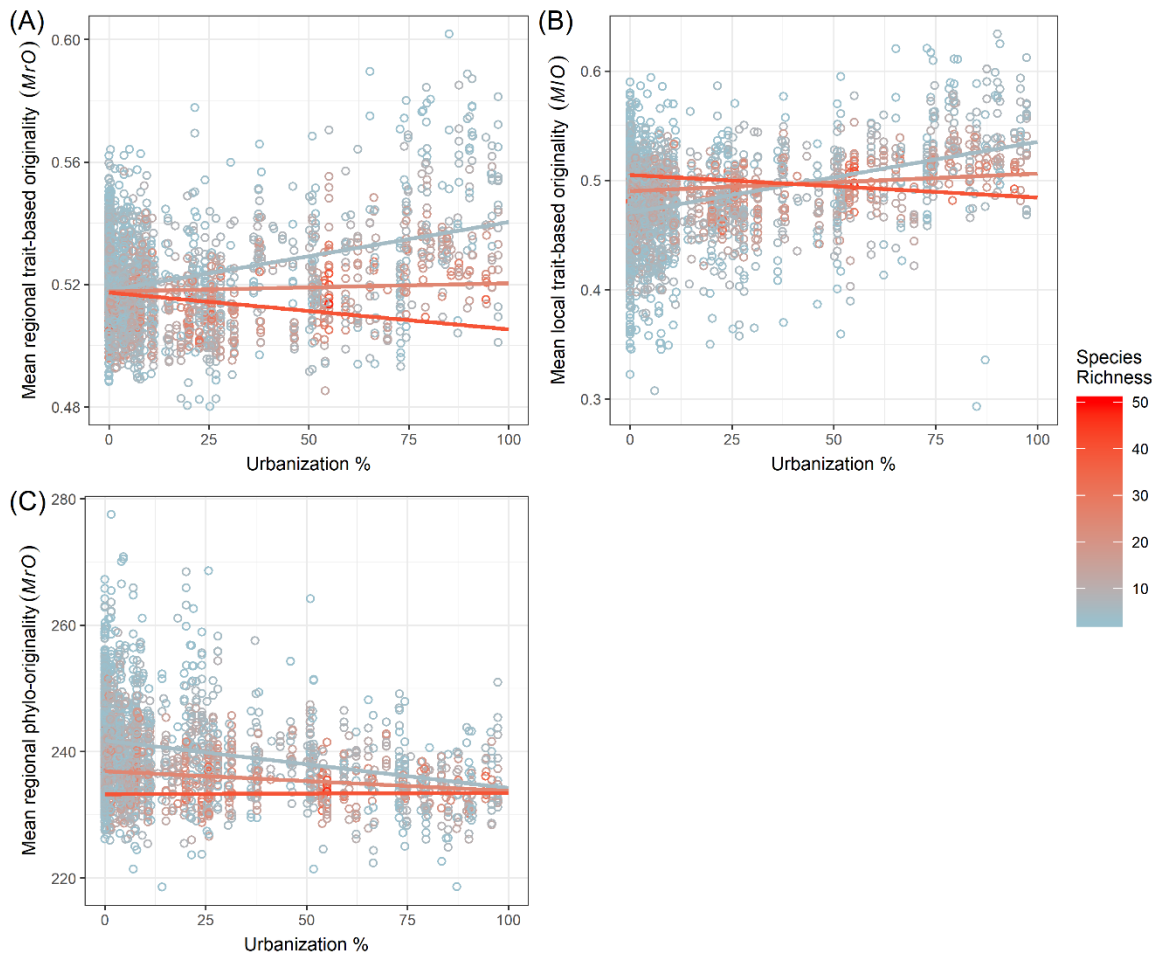
## Figures



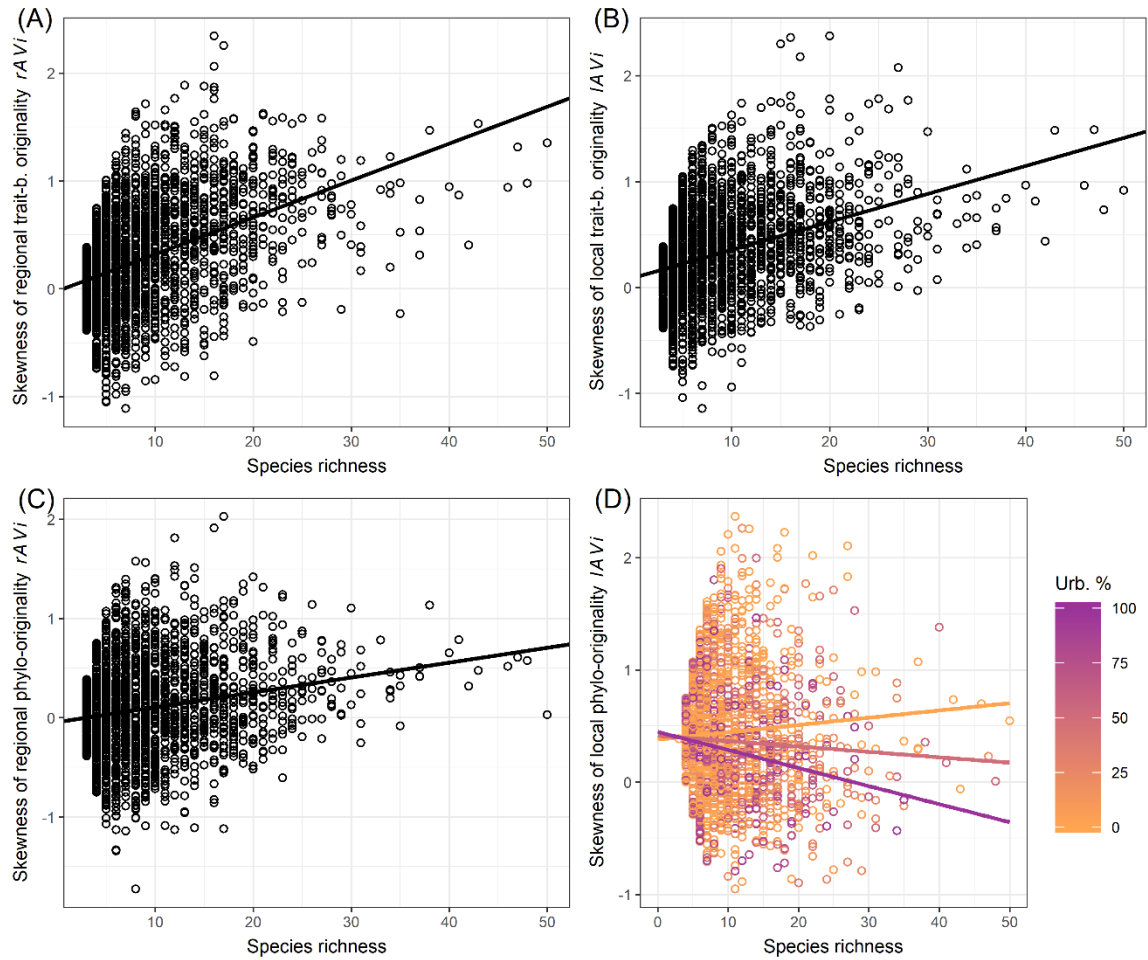
**Figure 1** Land cover and spatial distribution of the sampling plots. The map shows the Île-de-France region, including its largest city Paris and approximately 1300 other cities and towns, which is represented by urbanized areas in red (including individual housing (8%), collective housing (2%), industrial and business areas (2%), facilities (1%), quarries, dumps and worksites (0.5%) and transportation infrastructure (2.5%)). The urbanization cover and other land cover types (spatial resolution of 12.5 meters) that correspond to forests (dark green; 24% of total area), seminatural areas (bright green, 2%), agricultural areas (beige, 50%), artificial open areas (violet, 6%) and rivers and water points (blue, %) were extracted from data from the Institut d'Aménagement et d'Urbanisme (IAU, 2017). The blue dots represent the 2362 sampling plots of Vigie-flore from 2009 to 2017 (many of them are superimposed).



**Figure 2** Conceptual framework for measuring species trait-based and phylogenetic originality and diversity at multiple spatial scales. We illustrate this framework with a theoretical set of five plant species (Sp). First, the pairwise distances among species are calculated to form a dissimilarity matrix based on a set of species trait values using Gower’s metric for several types of traits (1.1, where Tr T is a species trait) or using species phylogenetic distances represented by branch lengths in millions of years (1.2). Second, species originality (trait-based or phylogenetic) is calculated at the regional scale considering all species from the regional pool (2.1,  $rAV_i$ ) and at the local scale considering only species from a local plot P (2.1,  $IAV_{ip}$ ). The fewer trait values that species Sp shares with the other species (in a region or in a local plot), or the fewer species Sp are related to the other species, the more original it is. Finally, community metrics are measured for each plot P (3): the mean of regionally measured originality values ( $MrO_p$ ) and their skewness, the mean of locally measured originality values ( $MIO_p$ ) and their skewness, and species richness.  $MrO_p$  represents the regional originality of the local plot, and  $MIO_p$  represents a measure of local species diversity.



**Figure 3** Variation in (A) mean regional trait-based originality, (B) mean local trait-based originality (local trait-based diversity) and (C) mean regional phylogenetic originality according to the percentage of urbanization within a 1 km<sup>2</sup> grid cell (grid cell corresponds to the Vigie-flore protocol) and species richness with a complete species dataset (natives and non-natives). One data point represents one species plot. The colors represent a scale of species richness (SR) going from blue for low SR (minimum of 3 species) to red for high SR (maximum of 50 species). Colored curves represent the estimated trends of an interaction between urbanization and SR for 5 (blue), 25 (salmon) and 40 (red) species retrieved from the mean originality GLMM models.



**Figure 4** The skewness of the (A) regional and (B) local trait-based originality and the (C) regional and (D) local phylogenetic originality according to the species richness (SR) in each species plot of the complete dataset. One data point represents one plot. In (A, B, C), the black curve represents the estimated trend retrieved from the skewness GLMM models. In (D), the colors represent a scale of urbanization gradient calculated as the percentage of urbanized area cover; the colored lines represent the estimated trends in an interaction between SR and urbanization for 0%, 50% and 100%, respectively, within a 1 km<sup>2</sup> grid cell, retrieved from the skewness GLMM model.

**Supplementary Table 1.** A list of traits and their states (for traits with multiple categories), trait description according to LEDA (Kleyer et al., 2008) and BiolFlor (Klotz et al., 2002) databases, sources data bases, type of trait variable and the proportion of total species number with missing data for each trait.

Trait	Trait states (species number)*	Description	Source data base	Variable type	Proportion (%) of total species number with missing data
Begin of flowering		Start of flowering period (month)	BiolFlor; Catminat; Multiple online sources	circular	0
Duration of flowering		Duration of flowering period (month)	BiolFlor; Catminat; Multiple online sources	quantitative	0
Life span	annual (228)	The individual cycle lasts for a maximum of one year (12 months)	BiolFlor; Ecoflora	binary	0
	biannual (73)	The plant has a vegetative growth for approx. one year before reaching the generative phase after which it completes its life cycle			
	perennial (350)	The plant has more than one generative phase in its life			

	pluriannual (33)	The plant has a vegetative growth longer than one year (up to 5 years) before completing its life cycle after the first and only generative reproduction stage			
Pollination vector	Apomixis (1)	Asexual reproduction, no fertilization			
	Geitonogamy (13)	Pollination by a neighboring flower			
	Insects (442)	Pollination by insects			
	Self spontaneous (366)	Spontaneous pollination within a flower	BiolFlor; Ecoflora; Catminat	binary	0
	Water (1)	Pollination on or below water			
	Wind (156)	Pollination by wind			
Seed dispersal mode	Self dispersal (227)	Seed dispersal by an explosive mechanism, by gravity, by autonomous placement of seeds or daughter plant away from mother plant			
	Dispersal by wind (338)	Seed dispersal by wind and by rolling over soil with wind			
	Dispersal by animals (505)	Seed dispersal by hoarding by animals, dispersed after digestion or by adhesion on animals	BiolFlor; Catminat	binary	0.1
	Dispersal by human activity (357)	Seed dispersal by trading of seeds/plants and with agricultural seeds			
	Dispersal by water flows (383)	Seed dispersal by surface currents and by raindrops			

Method of reproduction	s (297)	Reproduction only by seed or by spore	BiolFlor; Ecoflora	ordinal	0.1
	ssv (39)	Reproduction is mostly by seed, rarely vegetative			
	sv (205)	Reproduction can be by seed and vegetative equally			
	vvs (22)	Reproduction is mostly vegetative, rarely by seed			
	v (2)	Reproduction is only vegetative			
Seed weight	Mean of seed weight (mg)	LEDA; TRY; Multiple online sources	quantitative	2.29	
LDMC (mg/g)	Leaf Dry Mass Content is a measure of tissue density, is the ratio dry leaf mass to fresh leaf mass	LEDA; TRY	quantitative	14	
Leaf size	Leaf size is the one-sided projected surface area of an individual leaf or lamina (mm <sup>2</sup> )	LEDA; TRY; Multiple online sources	quantitative	19.8	
SLA	Specific Leaf Area is a ratio of fresh leaf area to leaf dry mass (mm <sup>2</sup> /mg)	LEDA; TRY	quantitative	9.65	
Plant height	Mean distance between ground and the highest photosynthetic tissue of plant (meters)	BiolFlor; Multiple online sources	quantitative	0	

\* multiple states are possible for a single species.

## Supplementary material for DaPhnE phylogeny manipulation

**Supplementary Table 2.** Species names replaced in DaPhnE by their synonyms as found in Vigie-flore database:

Species name available in DaPhnE	Synonym from Vigie-flore (*name available in Daphne but changed to match with trait databases)
<i>Amaranthus chlorostachys</i>	<i>Amaranthus hybridus*</i>
<i>Centunculus minimus</i>	<i>Anagallis minima</i>
<i>Halimione pedunculata</i>	<i>Atriplex pedunculata</i>
<i>Helictotrichon pubescens</i>	<i>Avenula pubescens</i>
<i>Carex otrubae</i>	<i>Carex cuprina</i>
<i>Carex ovalis</i>	<i>Carex leporina</i>
<i>Microrrhinum minus</i>	<i>Chaenorrhinum minus</i>
<i>Calamintha menthifolia</i>	<i>Clinopodium menthifolium</i>
<i>Virga pilosa</i>	<i>Dipsacus pilosus</i>
<i>Roegneria canina</i>	<i>Elymus caninus</i>
<i>Elytrigia repens</i>	<i>Elymus repens</i>
<i>Conyza sumatrensis</i>	<i>Erigeron sumatrensis</i>
<i>Ranunculus ficaria</i>	<i>Ficaria verna</i>
<i>Picris echioides</i>	<i>Helminthotheca echioides</i>
<i>Inula conyzae</i>	<i>Inula conyza</i>
<i>Galeobdolon luteum</i>	<i>Lamiastrum galeobdolon</i>
<i>Coronopus didymus</i>	<i>Lepidium didymum</i>
<i>Cardaria draba</i>	<i>Lepidium draba*</i>
<i>Peplis portula</i>	<i>Lythrum portula</i>



<i>Medicago x varia</i>	<i>Medicago sativa*</i>
<i>Listera ovata</i>	<i>Neottia ovata</i>
<i>Oxalis fontana</i>	<i>Oxalis stricta</i>
<i>Hieracium pilosella</i>	<i>Pilosella officinarum</i>
<i>Persicaria dubia</i>	<i>Polygonum mite</i>
<i>Fallopia japonica</i>	<i>Reynoutria japonica*</i>
<i>Tripleurospermum perforatum</i>	<i>Tripleurospermum inodorum</i>
<i>Pseudolysimachion spicatum</i>	<i>Veronica spicata</i>
<i>Taraxacum sect. Ruderalia</i>	<i>Taraxacum ruderale</i>
<i>Arabis glabra</i>	<i>Turritis glabra*</i>
<i>Elytrigia x laxa</i>	<i>Elymus pungens</i>

### **Vigie-flore species not found in Daphne phylogeny and added manually:**

Almost all species from Vigie-flore were present in Daphne, except 15 species which were placed to the phylogeny by hand, using existing trees from the literature. We used published molecular phylogenies (see below) to place the missing species on the DaPhnE phylogeny according to their closest relatives using functions of packages phytools (Revell 2012) and ape (Popescu) in R (R Core Team 2019). Taxa were either placed as polytomies or halfway between existing nodes, depending on tree topologies.

### **Species added with the reference used:**

*Antinoria agrostidea* (Inda et al., 2008), *Bellis annua* (Brouillet et al., 2009), *Bromus rubens* (Fortune et al., 2008), *Bryonia cretica* (Volz and Renner, 2008), *Carex mairei* (Escudero et al., 2010), *Centaurea debeauxii* (López-Alvarado et al., 2014), *Crepis pурсifolia* (Enke and Gemeinholzer, 2008), *Helictotrichon sempervirens* (Rodrigues et al., 2017), *Hypericum perforatum* (Meseguer et al., 2013), *Impatiens balsamina* (Janssens et al., 2006), *Potentilla montana* (Kechaykin and Shmakov, 2016), *Silene laeta* (Frajman et al., 2009), *Trifolium pallidum* (Ellison et al., 2006), *Verbena bonariensis* (Marx et al., 2010), *Vulpia muralis*

(Inda et al., 2008).

**References for phylogeny manipulations:**

Brouillet, L., Lowrey, T. K., Urbatsch, L., Karaman-Castro, V., Sancho, G., Wagstaff, S., et al. (2009). “Phylogeny and evolution of the Astereae (Compositae or Asteraceae)”. In: Systematics, evolution and biogeography of the Compositae. ed. Funk, V.A, Susanna, A., Stuessy, T., and Bayer, R. (Vienna: IAPT), p. 589 – 629.

Ellison, N. W., Liston, A., Steiner, J. J., Williams, W. M., and Taylor, N. L. (2006). Molecular phylogenetics of the clover genus (*Trifolium*-Leguminosae). *Mol. Phylogenet. Evol.* 39, 688–705. doi:10.1016/j.ympev.2006.01.004.

Enke, N., and Gemeinholzer, B. (2008). Babcock revisited: new insights into generic delimitation and character evolution in *Crepis* L. (Compositae: Cichorieae) from ITS and matK sequence data. *Taxon* 57, 756–768.

Escudero, M., Hipp, A. L., and Luceño, M. (2010). Karyotype stability and predictors of chromosome number variation in sedges: A study in *Carex* section *Spirostachyae* (Cyperaceae). *Mol. Phylogenet. Evol.* 57, 353–363. doi:10.1016/j.ympev.2010.07.009.

Fortune, P. M., Pourtau, N., Viron, N., and Ainouche, M. L. (2008). Molecular phylogeny and reticulate origins of the polyploid *Bromus* species from section *Genea* (Poaceae). *Am. J. Bot.* 95, 454–464. doi:10.3732/ajb.95.4.454.

Frajman, B., Heidari, N., and Oxelman, B. (2009). Phylogenetic relationships of *Atocion* and *Viscaria* (Sileneae, Caryophyllaceae) inferred from chloroplast, nuclear ribosomal, and low-copy gene DNA sequences. *Taxon* 58, 811–824.

Inda, L. A., Segarra-Moragues, J. G., Müller, J., Peterson, P. M., and Catalán, P. (2008). Dated historical biogeography of the temperate *Loliinae* (Poaceae, Pooideae) grasses in the northern and southern hemispheres. *Mol. Phylogenet. Evol.* 46, 932–957. doi:10.1016/j.ympev.2007.11.022.

Janssens, S., Geuten, K., Yuan, Y.-M., Song, Y., Küpfer, P., and Smets, E. (2006). Phylogenetics of *Impatiens* and *Hydrocera* (Balsaminaceae) Using Chloroplast *atpB-rbcL* Spacer Sequences. *Syst.*

Bot. 31, 171–180. doi:10.1600/036364406775971796.

Kechaykin, A. A., and Shmakov, A. I. (2016). A system of subtribe Potentillinae J. Presl (Rosaceae Juss.). *Turczaninowia* 19, 114–128. doi:10.14258/turczaninowia.19.4.16.

López-Alvarado, J., Sáez, L., Filigheddu, R., Garcia-Jacas, N., and Susanna, A. (2014). The limitations of molecular markers in phylogenetic reconstruction: The case of *Centaurea* sect. *Phrygia* (Compositae). *Taxon* 63, 1079–1091. doi:10.12705/635.6.

Marx, H. E., O'Leary, N., Yuan, Y. W., Lu-Irving, P., Tank, D. C., Múlgura, M. E., et al. (2010). A molecular phylogeny and classification of Verbenaceae. *Am. J. Bot.* 97, 1647–1663. doi:10.3732/ajb.1000144.

Meseguer, A. S., Aldasoro, J. J., and Sanmartín, I. (2013). Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St. John's wort (*Hypericum*). *Mol. Phylogenet. Evol.* 67, 379–403. doi:10.1016/j.ympev.2013.02.007

Popescu, A.-A., K. T. Huber and E. Paradis (2012). "ape 3.0: new tools for distance based phylogenetics and evolutionary analysis in R." *Bioinformatics* 28, 1536–1537, DOI: 10.1093/bioinformatics/bts184.

Revell, L. J. (2012). "phytools: an R package for phylogenetic comparative biology (and other things)." *Methods Ecol. Evol.* 3(2): 217-223 DOI: 10.1111/j.2041-210X.2011.00169.x.

Rodrigues, J., Viegas, W., and Silva, M. (2017). 45S rDNA external transcribed spacer organization reveals new phylogenetic relationships in *Avena* genus. *PLoS One* 12, 1–17. doi:10.1371/journal.pone.0176170.

Volz, S. M., and Renner, S. S. (2008). Hybridization, polyploidy, and evolutionary transitions between monoecy and dioecy in *Bryonia* (Cucurbitaceae). *Am. J. Bot.* 95, 1297–1306. doi:10.3732/ajb.0800187.

**Supplementary Table 3.** Lists of top 5% (30 out of 586 species) original species using regional scores and mean local scores for each species across all the plots it occurred in. Trait-based originality unit relies on Gower distances calculated for species' pairwise trait differences, bounded between 0 and 1, and phylogenetic originality unit represents Million years of evolution.

Regional trait-based originality (min=0.44, median =0.52, max=0.67)		Mean of all local trait-based originalities (min=0.36, median=0.48, max=0.71)		Regional phylogenetic originality (min=218.51, median=237.20, max=287.28)		Mean of all local phylogenetic originalities (min=90.41, median=233.56, max=297.20)	
<i>Exaculum pusillum</i>	0.67	<i>Helictotrichon sempervirens*</i> (1)	0.71	<i>Lemna minor</i>	287.28	<i>Vulpia muralis</i> (1)	297.20
<i>Antinoria agrostidea</i>	0.67	<i>Vulpia muralis</i> (1)	0.67	<i>Alisma lanceolatum</i>	287.02	<i>Antinoria agrostidea</i> (1)	297.20
<i>Minuartia viscosa</i>	0.65	<i>Exaculum pusillum</i> (1)	0.67	<i>Alisma plantago-aquatica</i>	287.02	<i>Triticum monococcum</i> (1)	295.60
<i>Bellis annua*</i>	0.65	<i>Trifolium pallidum*</i> (1)	0.67	<i>Arum italicum</i>	286.97	<i>Carex pseudocyperus</i> (1)	295.60
<i>Erigeron sumatrensis*</i>	0.64	<i>Arctium nemorosum</i> (1)	0.65	<i>Arum maculatum</i>	286.97	<i>Platanthera bifolia</i> (2)	290.44
<i>Potentilla montana</i>	0.64	<i>Antinoria agrostidea</i> (1)	0.65	<i>Tamus communis</i>	284.20	<i>Muscari neglectum</i> (2)	290.03
<i>Phragmites australis</i>	0.63	<i>Veronica acinifolia</i> (1)	0.65	<i>Colchicum autumnale</i>	283.74	<i>Arum italicum</i> (1)	288.87
<i>Carex mairei</i>	0.63	<i>Minuartia viscosa</i> (1)	0.64	<i>Iris pseudacorus</i>	282.28	<i>Dactylorhiza maculata</i> (1)	288.23
<i>Arctium nemorosum</i>	0.63	<i>Erigeron sumatrensis*</i> (76)	0.64	<i>Allium cepa*</i>	281.44	<i>Alisma plantago-aquatica</i> (2)	286.27
<i>Anagallis minima</i>	0.63	<i>Bellis annua</i> (2)	0.64	<i>Allium sphaerocephalon</i>	281.34	<i>Arum maculatum</i> (24)	285.85
<i>Stellaria media</i>	0.63	<i>Triticum monococcum*</i> (1)	0.62	<i>Allium vineale</i>	281.34	<i>Allium cepa*</i> (1)	285.28
<i>Utricularia australis</i>	0.63	<i>Orobanche picridis</i> (11)	0.62	<i>Goodyera repens*</i>	281.12	<i>Sorghum halepense*</i> (1)	285.28
<i>Trifolium pallidum*</i>	0.62	<i>Heracleum mantegazzianum*</i> (1)	0.62	<i>Cephalanthera damasonium</i>	281.01	<i>Anthericum ramosum</i> (1)	285.28
<i>Triticum monococcum*</i>	0.62	<i>Petroselinum crispum*</i> (1)	0.62	<i>Neottia ovata</i>	280.96	<i>Setaria viridis</i> (3)	285.24
<i>Helictotrichon sempervirens*</i>	0.62	<i>Centaurea debeauxcii</i> (1)	0.61	<i>Epipactis atrorubens</i>	280.92	<i>Alisma lanceolatum</i> (1)	285.15
<i>Amaranthus hybridus*</i>	0.62	<i>Artemisia verlotiorum*</i> (2)	0.61	<i>Epipactis belleborine</i>	280.92	<i>Carex pendula</i> (1)	283.45
<i>Vulpia muralis</i>	0.62	<i>Anagallis minima</i> (1)	0.61	<i>Himantoglossum hircinum</i>	280.63	<i>Iris pseudacorus</i> (2)	282.95
<i>Poa annua</i>	0.62	<i>Impatiens balfourii*</i> (2)	0.61	<i>Ophrys apifera</i>	280.63	<i>Anemone sylvestris</i> (1)	282.80
<i>Veronica acinifolia</i>	0.62	<i>Potentilla montana</i> (5)	0.61	<i>Dactylorhiza maculata</i>	280.61	<i>Scilla autumnalis</i> (1)	282.70
<i>Centaurea debeauxcii</i>	0.61	<i>Beta vulgaris*</i> (2)	0.61	<i>Anthericum ramosum</i>	280.60	<i>Goodyera repens*</i> (1)	282.70
<i>Veronica agrestis</i>	0.61	<i>Catapodium rigidum</i> (8)	0.60	<i>Platanthera bifolia</i>	280.59	<i>Himantoglossum hircinum</i> (7)	282.24
<i>Heracleum mantegazzianum*</i>	0.61	<i>Sison amomum</i> (2)	0.60	<i>Platanthera chlorantha</i>	280.59	<i>Allium vineale</i> (5)	282.08

<i>Orobanche picridis</i>	0.61	<i>Bromus catharticus</i> * (4)	0.60	<i>Convallaria majalis</i>	280.46	<i>Tamus communis</i> (29)	281.52
<i>Bromus diandrus</i>	0.61	<i>Crepis bursifolia</i> * (3)	0.59	<i>Polygonatum multiflorum</i>	280.41	<i>Anemone nemorosa</i> (40)	280.84
<i>Arum italicum</i>	0.60	<i>Arum italicum</i> * (1)	0.59	<i>Polygonatum odoratum</i>	280.41	<i>Colchicum autumnale</i> (1)	279.06
<i>Lemna minor</i>	0.60	<i>Centaurea microptilon</i> (2)	0.59	<i>Ornithogalum pyrenaicum</i>	280.36	<i>Helictotrichon sempervirens</i> (1)	278.54
<i>Amaranthus blitoides</i> *	0.60	<i>Persicaria amphibia</i> (6)	0.58	<i>Hyacinthoides non-scripta</i>	280.28	<i>Ranunculus flammula</i> (1)	278.53
<i>Sison amomum</i>	0.60	<i>Phragmites australis</i> (2)	0.58	<i>Scilla autumnalis</i>	280.26	<i>Ranunculus bulbosus</i> (10)	278.33
<i>Poa bulbosa</i>	0.60	<i>Poa annua</i> (319)	0.58	<i>Muscari comosum</i>	280.24	<i>Lemna minor</i> (1)	275.44
<i>Taraxacum</i> sect. <i>ruderalia</i>	0.60	<i>Clinopodium menthifolium</i> (5)	0.58	<i>Muscari neglectum</i>	280.24	<i>Papaver rhoeas</i> (184)	275.31

\*non-native species

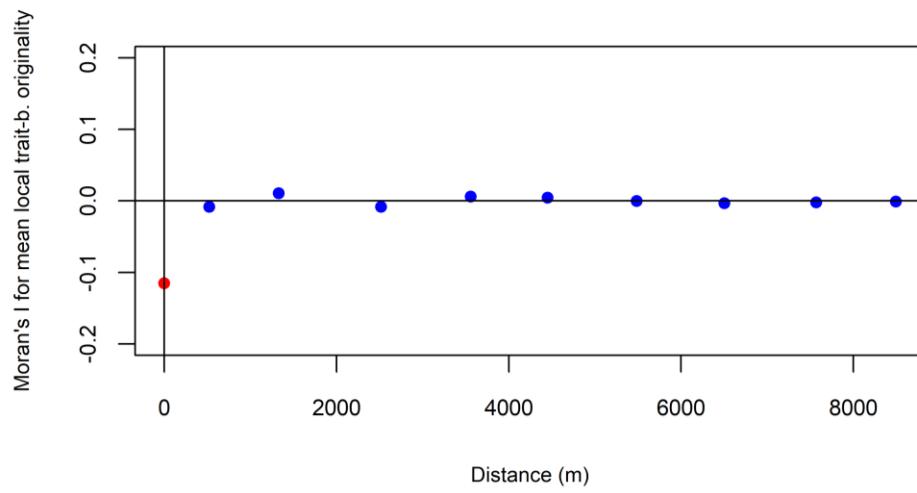
(n) species frequency across plots where it occurred in 10 years of sampling.

**Supplementary Table 4.** Results for the generalized linear mixed models (GLMM) ran on the dataset without non-native species including nine community level metrics (means and skewness of local and regional trait-based and phylogenetic originalities and species richness) and four species-specific measures (trait-based and phylogenetic originality and local to regional ratios)

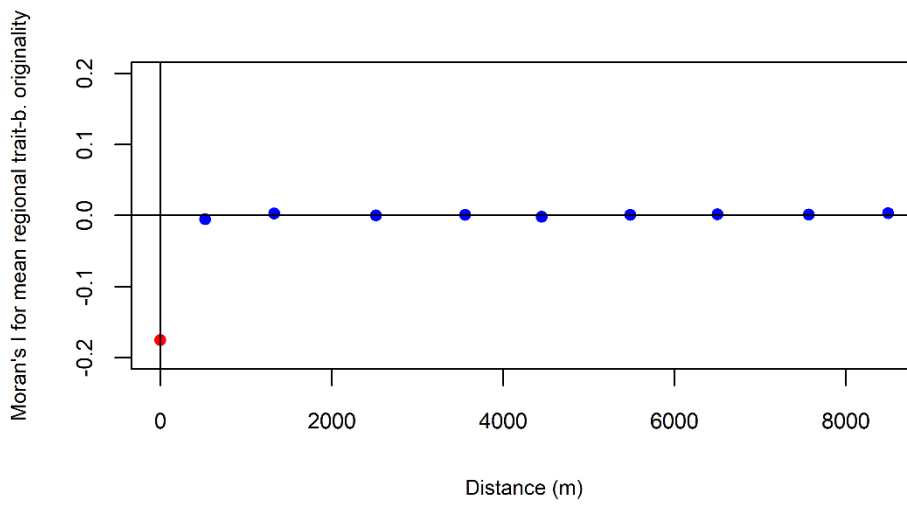
Estimated coefficients and their significance are shown for the explanatory variables that remained after backward selection of variables based on Akaike information criterion (AIC) values including urbanization percentage, species richness, interaction ( $\times$ ) between urbanization and species richness and year of sampling. The mention "excluded" means that the explanatory variable was not retained in the final model according to the Akaike criterion. Conditional and marginal  $R^2$  describes the proportion of variance explained by fixed ( $R^2_c$ ) and random+fixed ( $R^2_m$ ) covariables. P-values: <sup>NS</sup> = non-significant ( $P > 0.050$ ); \* =  $0.010 < P \leq 0.050$ ; \*\* =  $0.001 < P \leq 0.010$ ; \*\*\* =  $P \leq 0.001$

Dependent variable	Intercept	Urbanization (U)	Species richness (SR)	Year of sampling	U×SR	R <sup>2</sup> <sub>c</sub>	R <sup>2</sup> <sub>m</sub>
Mean local trait-based originality	4.693 10 <sup>-1</sup> ***	6.894 10 <sup>-4</sup> ***	9.077 10 <sup>-4</sup> ***	-5.827 10 <sup>-4</sup> *	-2.341 10 <sup>-5</sup> ***	0.435	0.119
Mean regional trait-based originality	5.150 10 <sup>-1</sup> ***	2.430 10 <sup>-4</sup> ***	-5.223 10 <sup>-7</sup> NS	-1.708 10 <sup>-4</sup> *	-1.042 10 <sup>-5</sup> ***	0.644	0.083
Mean local phylogenetic originality	2.322 10 <sup>2</sup> ***	-3.589 10 <sup>-3</sup> NS	-1.028 10 <sup>-1</sup> NS	-1.514 10 <sup>-1</sup> NS	-4.291 10 <sup>-3</sup> NS	0.242	0.004
Mean regional phylogenetic originality	2.428 10 <sup>2</sup> ***	-6.832 10 <sup>-2</sup> ***	-2.252 10 <sup>-1</sup> ***	excluded	1.541 10 <sup>-3</sup> NS	0.619	0.077
Skewness of local trait-based originality	1.110 10 <sup>-2</sup> ***	3.706 10 <sup>-5</sup> NS	2.525 10 <sup>-2</sup> ***	-1.522 10 <sup>-3</sup> NS	excluded	0.194	0.120
Skewness of regional trait-based originality	-3.719 10 <sup>-2</sup> NS	1.566 10 <sup>-3</sup> ***	3.141 10 <sup>-2</sup> ***	-9.724 10 <sup>-4</sup> NS	excluded	0.332	0.175
Skewness of local phylogenetic originality	4.187 10 <sup>-1</sup> ***	1.691 10 <sup>-4</sup> NS	3.903 10 <sup>-3</sup> NS	-4.360 10 <sup>-3</sup> NS	-1.894 10 <sup>-4</sup> **	0.149	0.012
Skewness of regional phylogenetic originality	-5.189 10 <sup>-2</sup> NS	6.392 10 <sup>-5</sup> NS	1.583 10 <sup>-2</sup> ***	3.086 10 <sup>-3</sup> NS	excluded	0.28	0.05
Species Richness	1.929 ***	3.959 10 <sup>-3</sup> ***	omitted	-4.032 10 <sup>-3</sup> NS	omitted	3.456 10 <sup>-2</sup>	1.659 10 <sup>-3</sup>
Trait-based originality	4.755 10 <sup>-1</sup> ***	3.826 10 <sup>-4</sup> ***	8.754 10 <sup>-4</sup> ***	-5.202 10 <sup>-4</sup> ***	-1.066 10 <sup>-5</sup> ***	0.655	0.034
Phylogenetic originality	1.818 10 <sup>-2</sup> NS	6.550 10 <sup>-4</sup> NS	7.694 10 <sup>-3</sup> ***	-6.562 10 <sup>-3</sup> *	-1.474 10 <sup>-4</sup> **	0.428	0.003
Local/regional trait-based originality	9.091 10 <sup>-1</sup> ***	7.388 10 <sup>-4</sup> ***	1.733 10 <sup>-3</sup> ***	excluded	-1.992 10 <sup>-5</sup> ***	0.450	0.056
Local/regional phylogenetic originality	9.564 10 <sup>-1</sup> ***	1.125 10 <sup>-4</sup> NS	1.064 10 <sup>-3</sup> ***	-8.562 10 <sup>-4</sup> *	-2.076 10 <sup>-5</sup> **	0.512	0.002

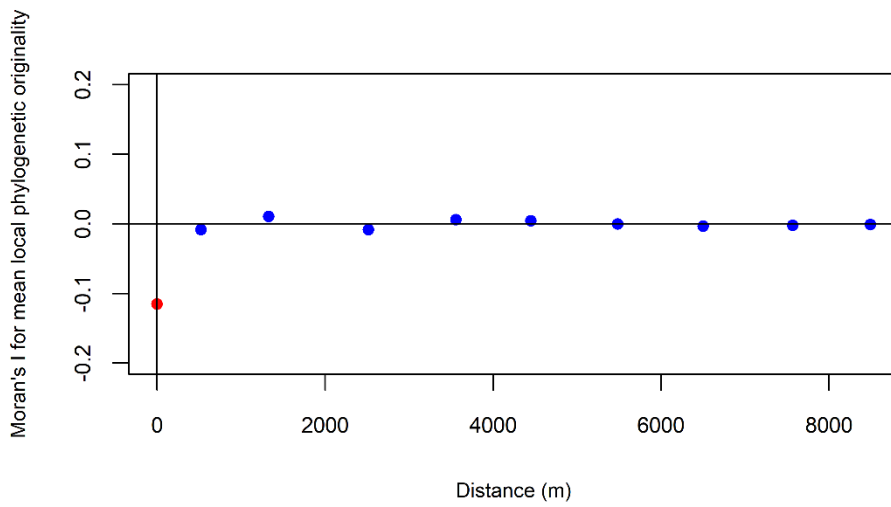
A)



B)

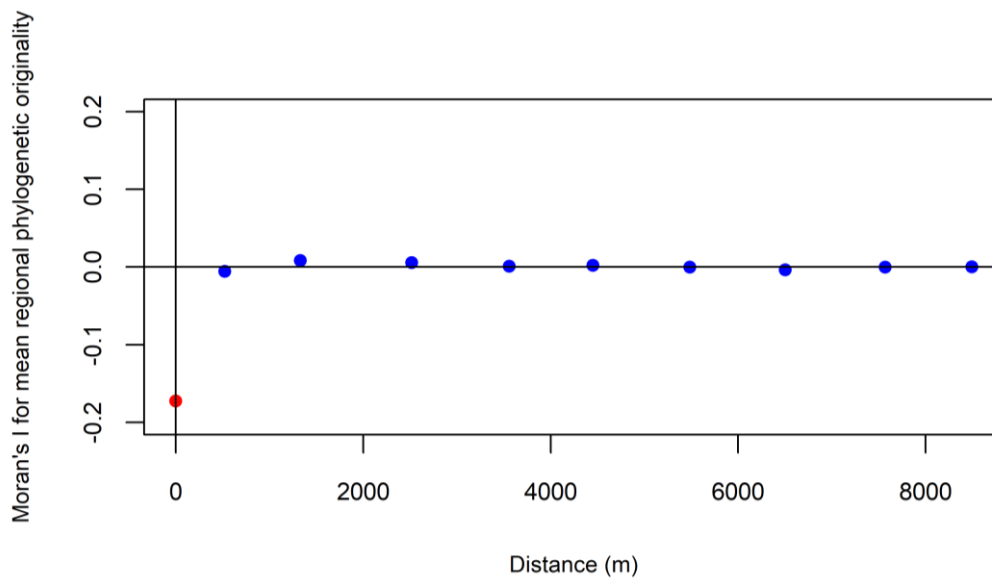


C)

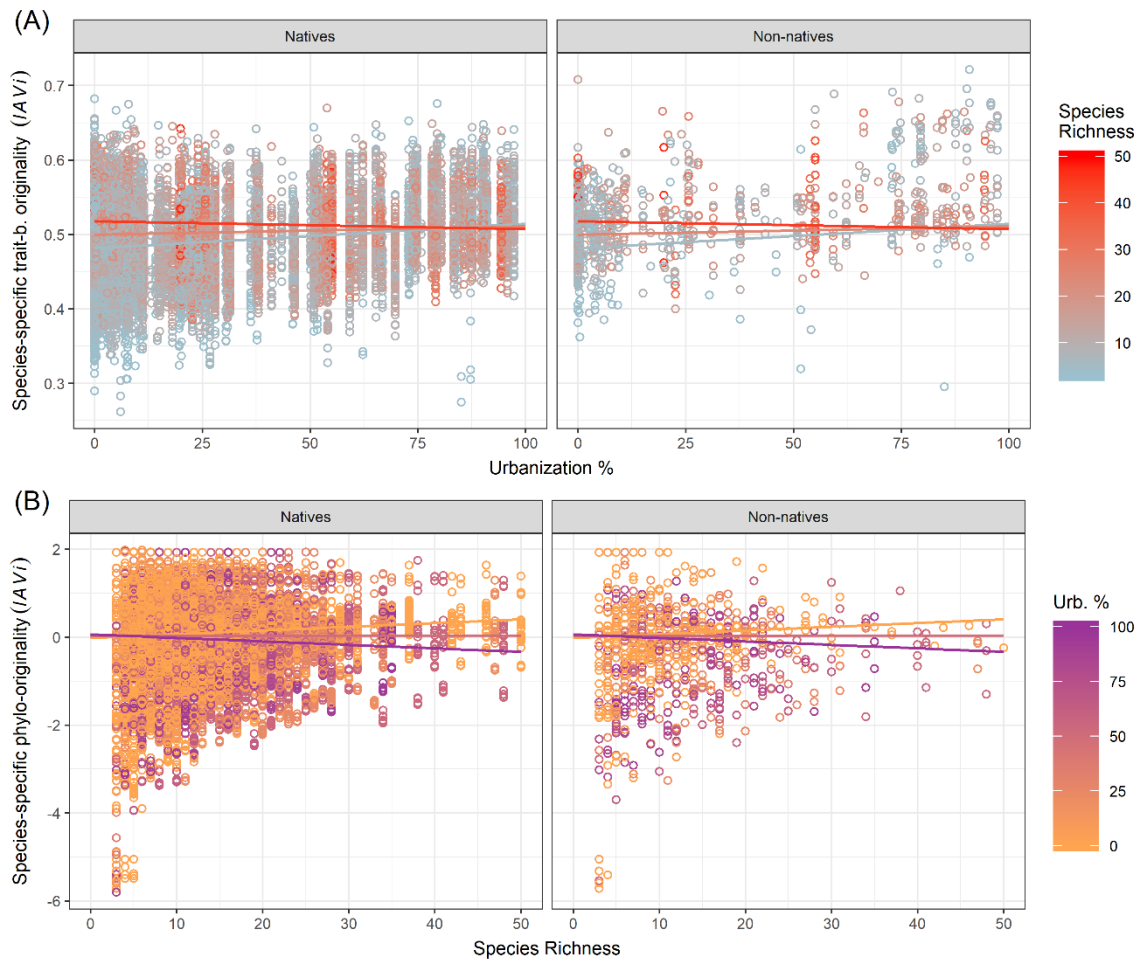




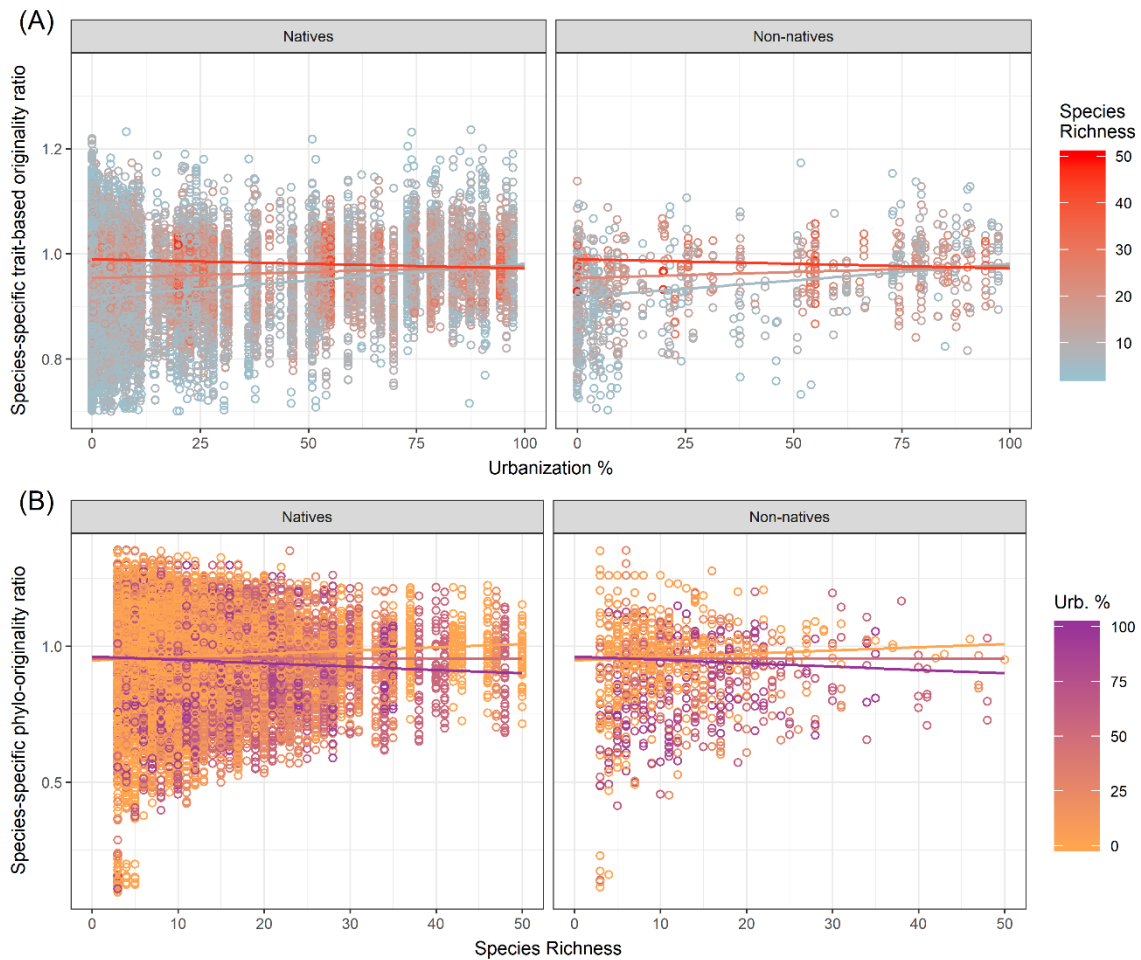
D)



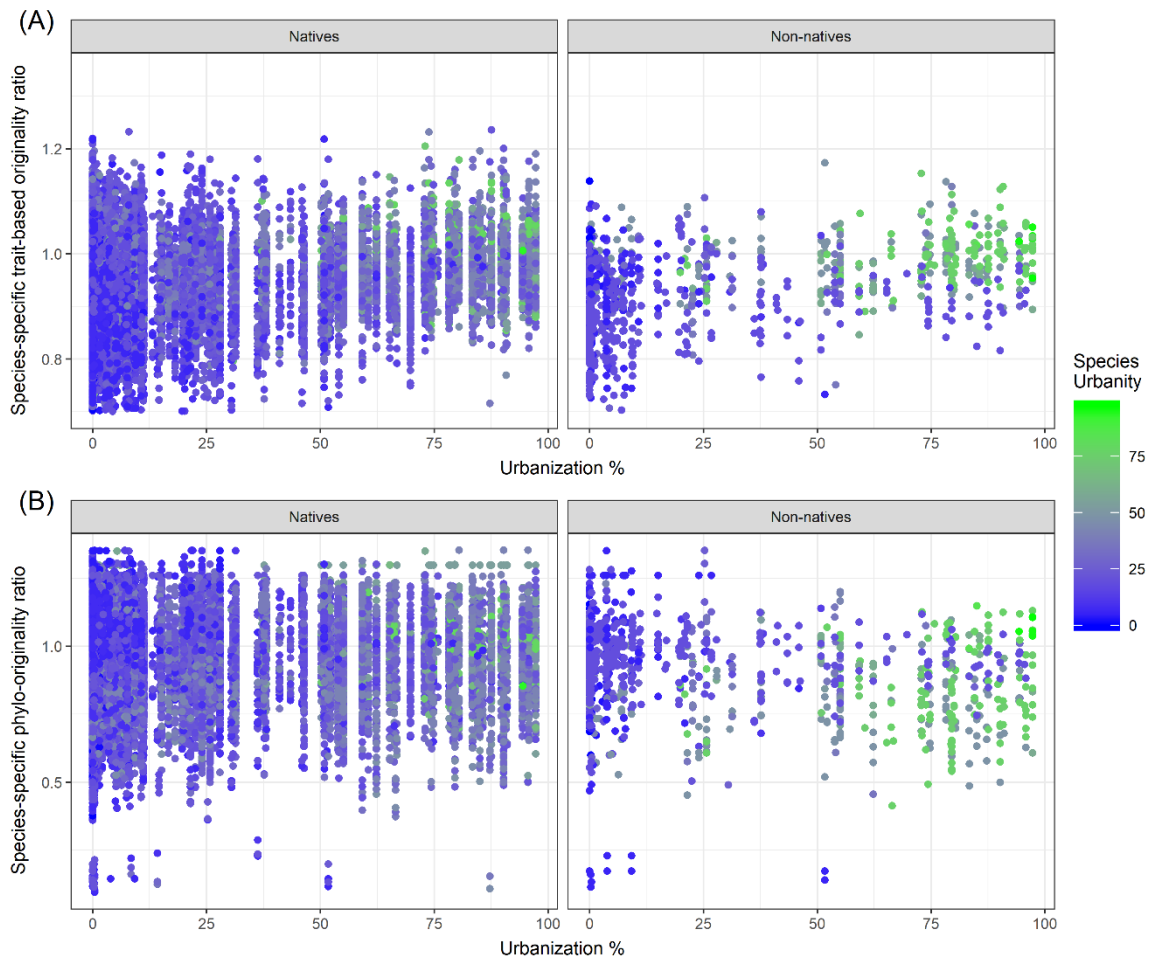
**Supplementary Figure 1.** Correlograms of the residuals of the GLMM models with complete dataset (natives and non-natives): (A) mean local and (B) regional trait-based originality models and (C) mean local and (D) regional phylogenetic originality model. Local Moran's correlation was tested on the residuals of the GLMM models. Significance of the correlation was assessed based on 999 randomizations at the 500 meters increment for neighborhood definition. First ten increment points are visualized, with blue points representing non-significance and the red points a significant spatial autocorrelation. Only the points at zero meters of distance were significant, reflecting the data structure with a systematic sampling performed each year at the same place for many species plots. No spatial autocorrelation was detected in models' residuals at any other distance. Thus, no correction for spatial autocorrelation was needed.



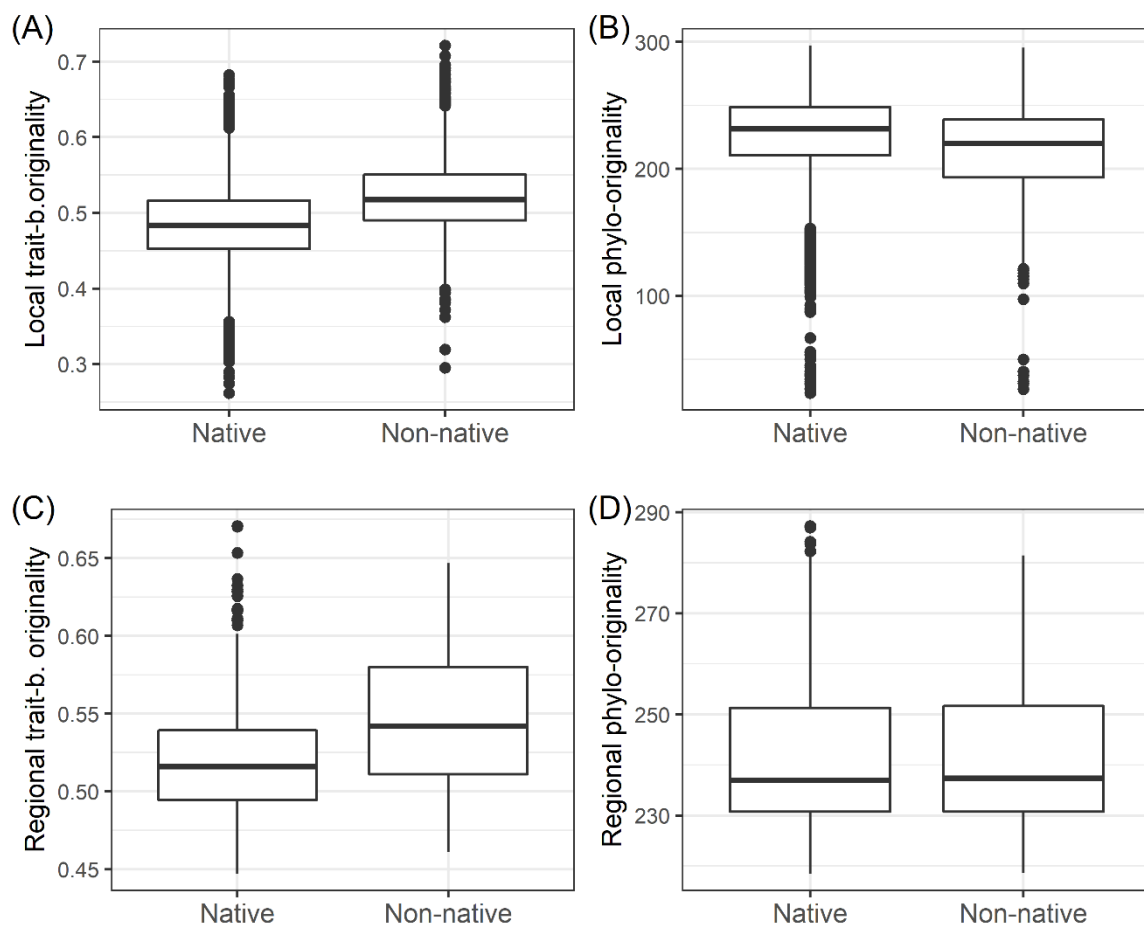
**Supplementary Figure 2.** Variation of species-specific **(A)** trait-based and **(B)** phylogenetic originalities of native and non-native species according to the percentage of urbanization within 1x1km grid cell (grid cell corresponds to the Vigie-flore protocol) (in **(A)**) and species richness (SR) (in **(B)**). One data point represents one species score, including originality values of native and non-native species from all plots across all the years, thus the same species can occur multiple times in different plots. In **(A)**, colors represent a scale of species richness (SR) going from blue as a lower SR (minimum of 3 species) to red as a higher SR (maximum of 50 species). Colored curves represent the estimated trends of an interaction between urbanization and SR for 5 (blue), 25 (salmon) and 40 (red) species retrieved from the mean originality GLMM models. In **(B)**, colors represent a scale of urbanization gradient calculated as the percentage of urbanized area cover; colored lines represent the estimated trends of an interaction between SR and urbanization for 0%, 50% and 100% respectively, within 1x1km grid cell, retrieved from the skewness GLMM model.



**Supplementary Figure 3.** Variation of species (A) trait-based and (B) phylogenetic originality ratio  $LAV_{ip} / rAV_i$  according to the percentage of urbanization within 1x1km grid cell (grid cell corresponds to the Vigie-flore protocol) (in (A)) and species richness (SR) (in (B)). One data point represents one species score, including originality values of native and non-native species from all plots across all the years, thus the same species can occur multiple times in different plots. Originality ratio superior to one indicate species that became original at the local scale at least once although they were not at the regional scale. In (A), colors represent a scale of species richness (SR) going from blue as a lower SR (minimum of 3 species) to red as a higher SR (maximum of 50 species). Colored curves represent the estimated trends of an interaction between urbanization and SR for 5 (blue), 25 (salmon) and 40 (red) species retrieved from the mean originality GLMM models. In (B), colors represent a scale of urbanization gradient calculated as the percentage of urbanized area cover; colored lines represent the estimated trends of an interaction between SR and urbanization for 0%, 50% and 100% respectively, within 1x1km grid cell, retrieved from the skewness GLMM model.



**Supplementary Figure 4.** Variation of species (A) trait-based and (B) phylogenetic originality ratio  $LAV_{ip} / rAV_i$  according to the percentage of urbanization within 1x1km grid cell (grid cell corresponds to the Vigie-flore protocol). One data point represents one species score, including originality values of native and non-native species from all plots across all the years, thus the same species can occur multiple times in different plots. Originality ratio superior to one indicate species that became original at the local scale at least once although they were not at the regional scale. Colors represent a scale of species urbanity going from blue for urbanophobic species to green for urbanophilic species. Species urbanity was calculated as the mean percentage of urbanized area in all grid cells where a species occurred.



**Supplementary Figure 6.** Box plots for (A) local trait-based, (B) local phylogenetic, (C) regional trait-based and (D) regional phylogenetic originality of 586 species. Regional originality values were calculated regarding all species in the dataset; local originality values were calculated for each of 586 species regarding only coexisting species in a particular plot where species occurred. Bold black lines represent medians, boxes are 25-75% interquartile, moustache (vertical lines) are samples with less than 1.5 times of the interquartile range and dots are outliers. Native and Non-native species differed significantly in their median (Wilcoxon rank test,  $W=5819854$ ,  $W=11347166$ ,  $W=18069$ ,  $p < 0.001$ ) in (A), (B), and (C).

**Supplementary Table 5.** Plant species dataset used in this study and corresponding amount of missing trait data

<i>Species name</i>	<i>Missing trait data (%)</i>
<i>Clinopodium menthifolium</i>	35,80%
<i>Bromus diandrus</i>	28,40%
<i>Hypochaeris radicata</i>	24,69%
<i>Verbena bonariensis</i>	22,22%
<i>Dactylis glomerata</i>	20,99%
<i>Hypericum hirsutum</i>	20,99%
<i>Vulpia muralis</i>	20,99%
<i>Centaurea microptilon</i>	16,05%
<i>Eschscholzia californica</i>	14,81%
<i>Helictotrichon sempervirens</i>	14,81%
<i>Catapodium rigidum</i>	13,58%
<i>Crepis bursifolia</i>	13,58%
<i>Duchesnea indica</i>	13,58%
<i>Trifolium pallidum</i>	13,58%
<i>Antinoria agrostidea</i>	12,35%
<i>Carex mairei</i>	12,35%
<i>Centaurea debeauxii</i>	12,35%
<i>Triticum monococcum</i>	12,35%
<i>Exaculum pusillum</i>	11,11%
<i>Minuartia viscosa</i>	11,11%
<i>Pulmonaria longifolia</i>	11,11%

<i>Solanum melongena</i>	11,11%
<i>Bellis annua</i>	9,88%
<i>Bromus rubens</i>	9,88%
<i>Bryonia cretica</i>	9,88%
<i>Centaurea nigra</i>	9,88%
<i>Erigeron sumatrensis</i>	9,88%
<i>Nigella damascena</i>	9,88%
<i>Potentilla montana</i>	9,88%
<i>Vicia parviflora</i>	9,88%
<i>Allium cepa</i>	8,64%
<i>Amaranthus hybridus</i>	8,64%
<i>Arum italicum</i>	8,64%
<i>Carduus tenuiflorus</i>	8,64%
<i>Carum verticillatum</i>	8,64%
<i>Ornithogalum pyrenaicum</i>	8,64%
<i>Orobanche picridis</i>	8,64%
<i>Veronica acinifolia</i>	8,64%
<i>Amaranthus cruentus</i>	7,41%
<i>Impatiens balfourii</i>	7,41%
<i>Neottia ovata</i>	7,41%
<i>Petroselinum crispum</i>	7,41%
<i>Pilosella officinarum</i>	7,41%
<i>Pisum sativum</i>	7,41%

<i>Vicia faba</i>	7,41%
<i>Amaranthus deflexus</i>	6,17%
<i>Artemisia verlotiorum</i>	6,17%
<i>Betonica officinalis</i>	6,17%
<i>Bromus catharticus</i>	6,17%
<i>Calamintha nepeta</i>	6,17%
<i>Calendula officinalis</i>	6,17%
<i>Campanula patula</i>	6,17%
<i>Carex cuprina</i>	6,17%
<i>Chaenorrhinum minus</i>	6,17%
<i>Crassula tillaea</i>	6,17%
<i>Crepis foetida</i>	6,17%
<i>Cucubalus baccifer</i>	6,17%
<i>Festuca filiformis</i>	6,17%
<i>Ficaria verna</i>	6,17%
<i>Gnaphalium uliginosum</i>	6,17%
<i>Helminthotheca echioides</i>	6,17%
<i>Lepidium didymum</i>	6,17%
<i>Linaria arvensis</i>	6,17%
<i>Lolium rigidum</i>	6,17%
<i>Melissa officinalis</i>	6,17%
<i>Muscari comosum</i>	6,17%
<i>Muscari neglectum</i>	6,17%



<i>Scilla autumnalis</i>	6,17%
<i>Sedum forsterianum</i>	6,17%
<i>Sisymbrium irio</i>	6,17%
<i>Taraxacum ruderales</i>	6,17%
<i>Tripleurospermum inodorum</i>	6,17%
<i>Vicia grandiflora</i>	6,17%
<i>Zea mays</i>	6,17%
<i>Amaranthus albus</i>	4,94%
<i>Arabis turrata</i>	4,94%
<i>Arctium nemorosum</i>	4,94%
<i>Arenaria serpyllifolia</i>	4,94%
<i>Brassica napus</i>	4,94%
<i>Carex leporina</i>	4,94%
<i>Crepis setosa</i>	4,94%
<i>Cynosurus echinatus</i>	4,94%
<i>Dipsacus sativus</i>	4,94%
<i>Elymus caninus</i>	4,94%
<i>Festuca heterophylla</i>	4,94%
<i>Glebionis segetum</i>	4,94%
<i>Goodyera repens</i>	4,94%
<i>Lathyrus aphaca</i>	4,94%
<i>Lathyrus hirsutus</i>	4,94%
<i>Lepidium draba</i>	4,94%

<i>Luzula forsteri</i>	4,94%
<i>Matricaria recutita</i>	4,94%
<i>Mycelis muralis</i>	4,94%
<i>Oxalis stricta</i>	4,94%
<i>Persicaria amphibia</i>	4,94%
<i>Persicaria hydropiper</i>	4,94%
<i>Persicaria lapathifolia</i>	4,94%
<i>Persicaria maculosa</i>	4,94%
<i>Sedum telephium</i>	4,94%
<i>Setaria verticillata</i>	4,94%
<i>Sison amomum</i>	4,94%
<i>Sorghum halepense</i>	4,94%
<i>Stellaria alsine</i>	4,94%
<i>Verbascum blattaria</i>	4,94%
<i>Viola alba</i>	4,94%
<i>Amaranthus blitoides</i>	3,70%
<i>Atriplex prostrata</i>	3,70%
<i>Cirsium palustre</i>	3,70%
<i>Epilobium tetragonum</i>	3,70%
<i>Erodium cicutarium</i>	3,70%
<i>Euphorbia cyparissias</i>	3,70%
<i>Festuca ovina</i>	3,70%
<i>Geranium rotundifolium</i>	3,70%

<i>Matricaria discoidea</i>	3,70%
<i>Odontites vernus</i>	3,70%
<i>Ononis spinosa</i>	3,70%
<i>Polygonum mite</i>	3,70%
<i>Ranunculus reptans</i>	3,70%
<i>Rubia peregrina</i>	3,70%
<i>Solanum tuberosum</i>	3,70%
<i>Vicia villosa</i>	3,70%
<i>Anagallis minima</i>	2,47%
<i>Anchusa arvensis</i>	2,47%
<i>Armeria maritima</i>	2,47%
<i>Aster lanceolatus</i>	2,47%
<i>Avenula pubescens</i>	2,47%
<i>Beta vulgaris</i>	2,47%
<i>Brassica oleracea</i>	2,47%
<i>Buglossoides arvensis</i>	2,47%
<i>Centaurea jacea</i>	2,47%
<i>Chaerophyllum hirsutum</i>	2,47%
<i>Elymus pungens</i>	2,47%
<i>Kickxia spuria</i>	2,47%
<i>Lamiaeum galeobdolon</i>	2,47%
<i>Limonium vulgare</i>	2,47%
<i>Lythrum portula</i>	2,47%

<i>Medicago minima</i>	2,47%
<i>Medicago polymorpha</i>	2,47%
<i>Medicago sativa</i>	2,47%
<i>Melica uniflora</i>	2,47%
<i>Ranunculus lanuginosus</i>	2,47%
<i>Reynoutria japonica</i>	2,47%
<i>Securigera varia</i>	2,47%
<i>Sedum acre</i>	2,47%
<i>Setaria italica</i>	2,47%
<i>Tamus communis</i>	2,47%
<i>Trifolium aureum</i>	2,47%
<i>Turritis glabra</i>	2,47%
<i>Utricularia australis</i>	2,47%
<i>Veronica spicata</i>	2,47%
<i>Agrostemma githago</i>	1,23%
<i>Amaranthus blitum</i>	1,23%
<i>Anagallis arvensis</i>	1,23%
<i>Anemone sylvestris</i>	1,23%
<i>Anthoxanthum odoratum</i>	1,23%
<i>Antirrhinum majus</i>	1,23%
<i>Brachypodium pinnatum</i>	1,23%
<i>Brassica rapa</i>	1,23%
<i>Bromus erectus</i>	1,23%

<i>Bromus inermis</i>	1,23%
<i>Bromus japonicus</i>	1,23%
<i>Bromus ramosus</i>	1,23%
<i>Bromus squarrosus</i>	1,23%
<i>Bromus sterilis</i>	1,23%
<i>Campanula rapunculoides</i>	1,23%
<i>Carex arenaria</i>	1,23%
<i>Carex muricata</i>	1,23%
<i>Carex vulpina</i>	1,23%
<i>Centaurea cyanus</i>	1,23%
<i>Centranthus ruber</i>	1,23%
<i>Cerastium arvense</i>	1,23%
<i>Cerastium semidecandrum</i>	1,23%
<i>Chenopodium hybridum</i>	1,23%
<i>Chenopodium polyspermum</i>	1,23%
<i>Chenopodium rubrum</i>	1,23%
<i>Conyza canadensis</i>	1,23%
<i>Coronopus squamatus</i>	1,23%
<i>Crepis vesicaria</i>	1,23%
<i>Daucus carota</i>	1,23%
<i>Deschampsia cespitosa</i>	1,23%
<i>Deschampsia flexuosa</i>	1,23%
<i>Dipsacus pilosus</i>	1,23%

<i>Elymus repens</i>	1,23%
<i>Epipactis atrorubens</i>	1,23%
<i>Erophila verna</i>	1,23%
<i>Euphorbia peplus</i>	1,23%
<i>Festuca arundinacea</i>	1,23%
<i>Festuca gigantea</i>	1,23%
<i>Festuca pratensis</i>	1,23%
<i>Festuca rubra</i>	1,23%
<i>Galinsoga parviflora</i>	1,23%
<i>Galium saxatile</i>	1,23%
<i>Hieracium umbellatum</i>	1,23%
<i>Hordeum vulgare</i>	1,23%
<i>Inula conyza</i>	1,23%
<i>Juncus effusus</i>	1,23%
<i>Kickxia elatine</i>	1,23%
<i>Knautia arvensis</i>	1,23%
<i>Lemna minor</i>	1,23%
<i>Leontodon autumnalis</i>	1,23%
<i>Leontodon hispidus</i>	1,23%
<i>Lepidium latifolium</i>	1,23%
<i>Leucanthemum vulgare</i>	1,23%
<i>Linaria repens</i>	1,23%
<i>Linum usitatissimum</i>	1,23%

<i>Lolium multiflorum</i>	1,23%
<i>Lotus pedunculatus</i>	1,23%
<i>Luzula pilosa</i>	1,23%
<i>Luzula sylvatica</i>	1,23%
<i>Malva sylvestris</i>	1,23%
<i>Medicago arabica</i>	1,23%
<i>Melica nutans</i>	1,23%
<i>Melilotus albus</i>	1,23%
<i>Melittis melissophyllum</i>	1,23%
<i>Myosotis discolor</i>	1,23%
<i>Myosoton aquaticum</i>	1,23%
<i>Papaver somniferum</i>	1,23%
<i>Phacelia tanacetifolia</i>	1,23%
<i>Pimpinella major</i>	1,23%
<i>Poa bulbosa</i>	1,23%
<i>Poa compressa</i>	1,23%
<i>Poa trivialis</i>	1,23%
<i>Polygonatum multiflorum</i>	1,23%
<i>Potentilla anserina</i>	1,23%
<i>Potentilla sterilis</i>	1,23%
<i>Pulsatilla vulgaris</i>	1,23%
<i>Sagina apetala</i>	1,23%
<i>Salvia pratensis</i>	1,23%

<i>Sanguisorba minor</i>	1,23%
<i>Scrophularia auriculata</i>	1,23%
<i>Senecio erucifolius</i>	1,23%
<i>Senecio inaequidens</i>	1,23%
<i>Senecio jacobaea</i>	1,23%
<i>Setaria viridis</i>	1,23%
<i>Silene latifolia</i>	1,23%
<i>Solanum nigrum</i>	1,23%
<i>Solidago canadensis</i>	1,23%
<i>Sonchus palustris</i>	1,23%
<i>Spergularia rubra</i>	1,23%
<i>Stachys palustris</i>	1,23%
<i>Stachys recta</i>	1,23%
<i>Stellaria nemorum</i>	1,23%
<i>Teucrium scordium</i>	1,23%
<i>Tordylium maximum</i>	1,23%
<i>Torilis nodosa</i>	1,23%
<i>Trifolium repens</i>	1,23%
<i>Triticum aestivum</i>	1,23%
<i>Valerianella ramosa</i>	1,23%
<i>Viola tricolor</i>	1,23%
<i>Vulpia bromoides</i>	1,23%
<i>Achillea millefolium</i>	0,00%



<i>Adoxa moschatellina</i>	0,00%
<i>Aethusa cynapium</i>	0,00%
<i>Agrimonia eupatoria</i>	0,00%
<i>Agrostis canina</i>	0,00%
<i>Agrostis capillaris</i>	0,00%
<i>Agrostis gigantea</i>	0,00%
<i>Agrostis stolonifera</i>	0,00%
<i>Ajuga genevensis</i>	0,00%
<i>Ajuga reptans</i>	0,00%
<i>Alisma lanceolatum</i>	0,00%
<i>Alisma plantago-aquatica</i>	0,00%
<i>Alliaria petiolata</i>	0,00%
<i>Allium sphaerocephalon</i>	0,00%
<i>Allium vineale</i>	0,00%
<i>Alopecurus geniculatus</i>	0,00%
<i>Alopecurus myosuroides</i>	0,00%
<i>Alopecurus pratensis</i>	0,00%
<i>Amaranthus retroflexus</i>	0,00%
<i>Ammi majus</i>	0,00%
<i>Anemone nemorosa</i>	0,00%
<i>Angelica sylvestris</i>	0,00%
<i>Anthemis cotula</i>	0,00%
<i>Anthericum ramosum</i>	0,00%

<i>Anthriscus caucalis</i>	0,00%
<i>Anthriscus sylvestris</i>	0,00%
<i>Apera spica-venti</i>	0,00%
<i>Aphanes arvensis</i>	0,00%
<i>Apium nodiflorum</i>	0,00%
<i>Aquilegia vulgaris</i>	0,00%
<i>Arabidopsis thaliana</i>	0,00%
<i>Arctium lappa</i>	0,00%
<i>Arctium minus</i>	0,00%
<i>Arrhenatherum elatius</i>	0,00%
<i>Artemisia vulgaris</i>	0,00%
<i>Arum maculatum</i>	0,00%
<i>Astragalus glycyphyllos</i>	0,00%
<i>Atriplex patula</i>	0,00%
<i>Avena fatua</i>	0,00%
<i>Avena sativa</i>	0,00%
<i>Ballota nigra</i>	0,00%
<i>Barbarea vulgaris</i>	0,00%
<i>Bellis perennis</i>	0,00%
<i>Blackstonia perfoliata</i>	0,00%
<i>Brachypodium sylvaticum</i>	0,00%
<i>Brassica nigra</i>	0,00%
<i>Bromus arvensis</i>	0,00%

<i>Bromus commutatus</i>	0,00%
<i>Bromus hordeaceus</i>	0,00%
<i>Bromus racemosus</i>	0,00%
<i>Bromus secalinus</i>	0,00%
<i>Calamagrostis epigejos</i>	0,00%
<i>Calendula arvensis</i>	0,00%
<i>Calystegia sepium</i>	0,00%
<i>Campanula rapunculus</i>	0,00%
<i>Campanula rotundifolia</i>	0,00%
<i>Capsella bursa-pastoris</i>	0,00%
<i>Cardamine hirsuta</i>	0,00%
<i>Cardamine pratensis</i>	0,00%
<i>Carduus acanthoides</i>	0,00%
<i>Carduus crispus</i>	0,00%
<i>Carduus nutans</i>	0,00%
<i>Carex acutiformis</i>	0,00%
<i>Carex divulsa</i>	0,00%
<i>Carex flacca</i>	0,00%
<i>Carex hirta</i>	0,00%
<i>Carex pendula</i>	0,00%
<i>Carex pilulifera</i>	0,00%
<i>Carex pseudocyperus</i>	0,00%
<i>Carex remota</i>	0,00%

<i>Carex riparia</i>	0,00%
<i>Carex spicata</i>	0,00%
<i>Carex sylvatica</i>	0,00%
<i>Carlina vulgaris</i>	0,00%
<i>Centaurea scabiosa</i>	0,00%
<i>Centaurium erythraea</i>	0,00%
<i>Centaurium pulchellum</i>	0,00%
<i>Cephalanthera damasonium</i>	0,00%
<i>Cerastium fontanum</i>	0,00%
<i>Cerastium glomeratum</i>	0,00%
<i>Cerastium pumilum</i>	0,00%
<i>Chaerophyllum temulum</i>	0,00%
<i>Chelidonium majus</i>	0,00%
<i>Chenopodium album</i>	0,00%
<i>Chenopodium ficifolium</i>	0,00%
<i>Circaea lutetiana</i>	0,00%
<i>Cirsium arvense</i>	0,00%
<i>Cirsium dissectum</i>	0,00%
<i>Cirsium oleraceum</i>	0,00%
<i>Cirsium vulgare</i>	0,00%
<i>Clinopodium vulgare</i>	0,00%
<i>Colchicum autumnale</i>	0,00%
<i>Conium maculatum</i>	0,00%

<i>Convallaria majalis</i>	0,00%
<i>Convolvulus arvensis</i>	0,00%
<i>Crepis capillaris</i>	0,00%
<i>Cruciata laevipes</i>	0,00%
<i>Cymbalaria muralis</i>	0,00%
<i>Cynodon dactylon</i>	0,00%
<i>Cynosurus cristatus</i>	0,00%
<i>Dactylorhiza maculata</i>	0,00%
<i>Danthonia decumbens</i>	0,00%
<i>Datura stramonium</i>	0,00%
<i>Digitalis purpurea</i>	0,00%
<i>Digitaria ischaemum</i>	0,00%
<i>Digitaria sanguinalis</i>	0,00%
<i>Diploaxis tenuifolia</i>	0,00%
<i>Dipsacus fullonum</i>	0,00%
<i>Echinochloa crus-galli</i>	0,00%
<i>Echium vulgare</i>	0,00%
<i>Epilobium angustifolium</i>	0,00%
<i>Epilobium ciliatum</i>	0,00%
<i>Epilobium hirsutum</i>	0,00%
<i>Epilobium montanum</i>	0,00%
<i>Epilobium parviflorum</i>	0,00%
<i>Epipactis helleborine</i>	0,00%

<i>Eragrostis minor</i>	0,00%
<i>Erigeron annuus</i>	0,00%
<i>Eryngium campestre</i>	0,00%
<i>Erysimum cheiri</i>	0,00%
<i>Eupatorium cannabinum</i>	0,00%
<i>Euphorbia amygdaloides</i>	0,00%
<i>Euphorbia exigua</i>	0,00%
<i>Euphorbia helioscopia</i>	0,00%
<i>Fallopia convolvulus</i>	0,00%
<i>Filipendula ulmaria</i>	0,00%
<i>Filipendula vulgaris</i>	0,00%
<i>Fragaria vesca</i>	0,00%
<i>Fragaria viridis</i>	0,00%
<i>Fumaria officinalis</i>	0,00%
<i>Galega officinalis</i>	0,00%
<i>Galeopsis tetrahit</i>	0,00%
<i>Galium aparine</i>	0,00%
<i>Galium mollugo</i>	0,00%
<i>Galium odoratum</i>	0,00%
<i>Galium palustre</i>	0,00%
<i>Galium parisiense</i>	0,00%
<i>Galium uliginosum</i>	0,00%
<i>Galium verum</i>	0,00%

<i>Geranium columbinum</i>	0,00%
<i>Geranium dissectum</i>	0,00%
<i>Geranium molle</i>	0,00%
<i>Geranium pusillum</i>	0,00%
<i>Geranium pyrenaicum</i>	0,00%
<i>Geranium robertianum</i>	0,00%
<i>Geranium sanguineum</i>	0,00%
<i>Geum urbanum</i>	0,00%
<i>Glechoma hederacea</i>	0,00%
<i>Helleborus foetidus</i>	0,00%
<i>Heracleum mantegazzianum</i>	0,00%
<i>Heracleum sphondylium</i>	0,00%
<i>Herniaria glabra</i>	0,00%
<i>Herniaria hirsuta</i>	0,00%
<i>Himantoglossum hircinum</i>	0,00%
<i>Hippocrepis comosa</i>	0,00%
<i>Hirschfeldia incana</i>	0,00%
<i>Holcus lanatus</i>	0,00%
<i>Holcus mollis</i>	0,00%
<i>Hordeum murinum</i>	0,00%
<i>Humulus lupulus</i>	0,00%
<i>Hyacinthoides non-scripta</i>	0,00%
<i>Hydrocotyle vulgaris</i>	0,00%

<i>Hypericum humifusum</i>	0,00%
<i>Hypericum maculatum</i>	0,00%
<i>Hypericum perforatum</i>	0,00%
<i>Hypericum pulchrum</i>	0,00%
<i>Hypericum tetrapterum</i>	0,00%
<i>Illecebrum verticillatum</i>	0,00%
<i>Iris pseudacorus</i>	0,00%
<i>Juncus acutiflorus</i>	0,00%
<i>Juncus bufonius</i>	0,00%
<i>Juncus conglomeratus</i>	0,00%
<i>Juncus inflexus</i>	0,00%
<i>Juncus tenuis</i>	0,00%
<i>Lactuca serriola</i>	0,00%
<i>Lactuca virosa</i>	0,00%
<i>Lamium album</i>	0,00%
<i>Lamium amplexicaule</i>	0,00%
<i>Lamium purpureum</i>	0,00%
<i>Lapsana communis</i>	0,00%
<i>Lathyrus latifolius</i>	0,00%
<i>Lathyrus nissolia</i>	0,00%
<i>Lathyrus pratensis</i>	0,00%
<i>Lathyrus tuberosus</i>	0,00%
<i>Legousia speculum-veneris</i>	0,00%



<i>Linaria vulgaris</i>	0,00%
<i>Linum catharticum</i>	0,00%
<i>Lolium perenne</i>	0,00%
<i>Lotus corniculatus</i>	0,00%
<i>Luzula campestris</i>	0,00%
<i>Luzula multiflora</i>	0,00%
<i>Lycopus europaeus</i>	0,00%
<i>Lysimachia nummularia</i>	0,00%
<i>Lysimachia vulgaris</i>	0,00%
<i>Lythrum salicaria</i>	0,00%
<i>Malva moschata</i>	0,00%
<i>Malva neglecta</i>	0,00%
<i>Medicago lupulina</i>	0,00%
<i>Melampyrum arvense</i>	0,00%
<i>Melampyrum pratense</i>	0,00%
<i>Melilotus officinalis</i>	0,00%
<i>Mentha aquatica</i>	0,00%
<i>Mentha arvensis</i>	0,00%
<i>Mentha suaveolens</i>	0,00%
<i>Mercurialis annua</i>	0,00%
<i>Mercurialis perennis</i>	0,00%
<i>Milium effusum</i>	0,00%
<i>Moehringia trinervia</i>	0,00%

<i>Molinia caerulea</i>	0,00%
<i>Myosotis arvensis</i>	0,00%
<i>Myosotis ramosissima</i>	0,00%
<i>Myosotis scorpioides</i>	0,00%
<i>Myosotis sylvatica</i>	0,00%
<i>Oenanthe lachenalii</i>	0,00%
<i>Onobrychis viciifolia</i>	0,00%
<i>Onopordum acanthium</i>	0,00%
<i>Ophrys apifera</i>	0,00%
<i>Origanum vulgare</i>	0,00%
<i>Ornithopus perpusillus</i>	0,00%
<i>Oxalis corniculata</i>	0,00%
<i>Papaver dubium</i>	0,00%
<i>Papaver rhoeas</i>	0,00%
<i>Parietaria judaica</i>	0,00%
<i>Parietaria officinalis</i>	0,00%
<i>Pastinaca sativa</i>	0,00%
<i>Phalaris arundinacea</i>	0,00%
<i>Phleum pratense</i>	0,00%
<i>Phragmites australis</i>	0,00%
<i>Picris hieracioides</i>	0,00%
<i>Pimpinella saxifraga</i>	0,00%
<i>Plantago coronopus</i>	0,00%

<i>Plantago lanceolata</i>	0,00%
<i>Plantago major</i>	0,00%
<i>Plantago media</i>	0,00%
<i>Platanthera bifolia</i>	0,00%
<i>Platanthera chlorantha</i>	0,00%
<i>Poa annua</i>	0,00%
<i>Poa nemoralis</i>	0,00%
<i>Poa pratensis</i>	0,00%
<i>Polygonatum odoratum</i>	0,00%
<i>Polygonum aviculare</i>	0,00%
<i>Portulaca oleracea</i>	0,00%
<i>Potentilla erecta</i>	0,00%
<i>Potentilla recta</i>	0,00%
<i>Potentilla reptans</i>	0,00%
<i>Primula elatior</i>	0,00%
<i>Primula veris</i>	0,00%
<i>Primula vulgaris</i>	0,00%
<i>Prunella vulgaris</i>	0,00%
<i>Pulicaria dysenterica</i>	0,00%
<i>Ranunculus acris</i>	0,00%
<i>Ranunculus bulbosus</i>	0,00%
<i>Ranunculus flammula</i>	0,00%
<i>Ranunculus repens</i>	0,00%

<i>Ranunculus sardous</i>	0,00%
<i>Raphanus raphanistrum</i>	0,00%
<i>Reseda lutea</i>	0,00%
<i>Reseda luteola</i>	0,00%
<i>Rorippa sylvestris</i>	0,00%
<i>Rumex acetosa</i>	0,00%
<i>Rumex acetosella</i>	0,00%
<i>Rumex conglomeratus</i>	0,00%
<i>Rumex crispus</i>	0,00%
<i>Rumex obtusifolius</i>	0,00%
<i>Rumex sanguineus</i>	0,00%
<i>Sagina procumbens</i>	0,00%
<i>Sambucus ebulus</i>	0,00%
<i>Sanicula europaea</i>	0,00%
<i>Saponaria officinalis</i>	0,00%
<i>Saxifraga tridactylites</i>	0,00%
<i>Scandix pecten-veneris</i>	0,00%
<i>Scleranthus annuus</i>	0,00%
<i>Scrophularia nodosa</i>	0,00%
<i>Sedum album</i>	0,00%
<i>Senecio viscosus</i>	0,00%
<i>Senecio vulgaris</i>	0,00%
<i>Sherardia arvensis</i>	0,00%

<i>Silene vulgaris</i>	0,00%
<i>Sinapis alba</i>	0,00%
<i>Sinapis arvensis</i>	0,00%
<i>Sisymbrium officinale</i>	0,00%
<i>Solidago virgaurea</i>	0,00%
<i>Sonchus arvensis</i>	0,00%
<i>Sonchus asper</i>	0,00%
<i>Sonchus oleraceus</i>	0,00%
<i>Spergula arvensis</i>	0,00%
<i>Stachys arvensis</i>	0,00%
<i>Stachys sylvatica</i>	0,00%
<i>Stellaria graminea</i>	0,00%
<i>Stellaria holostea</i>	0,00%
<i>Stellaria media</i>	0,00%
<i>Symphytum officinale</i>	0,00%
<i>Tanacetum vulgare</i>	0,00%
<i>Teucrium chamaedrys</i>	0,00%
<i>Teucrium scorodonia</i>	0,00%
<i>Thlaspi arvense</i>	0,00%
<i>Torilis arvensis</i>	0,00%
<i>Torilis japonica</i>	0,00%
<i>Tragopogon pratensis</i>	0,00%
<i>Trifolium arvense</i>	0,00%

<i>Trifolium campestre</i>	0,00%
<i>Trifolium dubium</i>	0,00%
<i>Trifolium fragiferum</i>	0,00%
<i>Trifolium hybridum</i>	0,00%
<i>Trifolium pratense</i>	0,00%
<i>Trisetum flavescens</i>	0,00%
<i>Tussilago farfara</i>	0,00%
<i>Urtica dioica</i>	0,00%
<i>Urtica urens</i>	0,00%
<i>Verbascum nigrum</i>	0,00%
<i>Verbascum thapsus</i>	0,00%
<i>Verbena officinalis</i>	0,00%
<i>Veronica agrestis</i>	0,00%
<i>Veronica arvensis</i>	0,00%
<i>Veronica beccabunga</i>	0,00%
<i>Veronica chamaedrys</i>	0,00%
<i>Veronica filiformis</i>	0,00%
<i>Veronica hederifolia</i>	0,00%
<i>Veronica montana</i>	0,00%
<i>Veronica officinalis</i>	0,00%
<i>Veronica persica</i>	0,00%
<i>Veronica polita</i>	0,00%
<i>Veronica serpyllifolia</i>	0,00%

<i>Vicia cracca</i>	0,00%
<i>Vicia hirsuta</i>	0,00%
<i>Vicia lathyroides</i>	0,00%
<i>Vicia sativa</i>	0,00%
<i>Vicia sepium</i>	0,00%
<i>Vicia tetrasperma</i>	0,00%
<i>Vincetoxicum hirundinaria</i>	0,00%
<i>Viola arvensis</i>	0,00%
<i>Viola hirta</i>	0,00%
<i>Viola odorata</i>	0,00%
<i>Viola reichenbachiana</i>	0,00%
<i>Viola riviniana</i>	0,00%
<i>Vulpia myuros</i>	0,00%

## SUPPLEMENTARY RESULTS AND COMMENTARIES TO CHAPTER 2

### Variation of local species-specific originality

In this chapter we investigated species' local and regional originality variation along an urbanization gradient. However, as the originality of a given species is relative to the other species in a set, its value should be different from one set to another. We wondered whether the most original species were always highly original despite different composition of species sets or whether its originality varied. I present below the results obtained with trait-based originality scores (unpublished results).

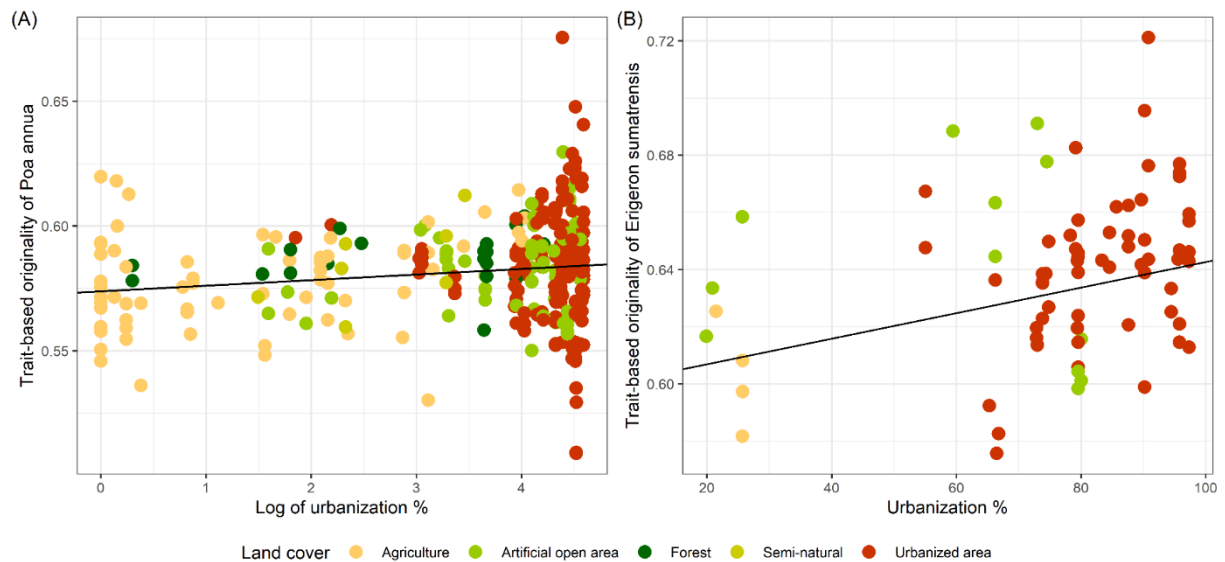
We noticed that many of the top 5% of the most original species occurred only few times in the whole 10-years dataset, making impossible to follow their originality changes. We therefore selected species only occurring more than 50 times from the top 5% trait-based original species and followed their originality variation across land cover types and urbanization gradient. One of such species was a fleabane *Erigeron sumatrensis*. Its high trait-based originality value (max = 0.72, min = 0.57) in highly urbanized areas could reflect the fact that as a non-native species, with original trait values, *E. sumatrensis* is able to colonize novel ecosystems created by urbanization where species competition is low (but see Dawson et al., 2015). Another highly original species in its traits was the annual bluegrass *Poa annua*, native and highly common species on the Ile-de-France region (appeared within every land cover type). *P. annua* trait-based originality increased with urbanization and varied between 0.50 and 0.68 with the greatest variation observed in more urbanized areas (Figure S7).

Trait values vary across species and underpin differences in their ecological strategies (Westboy et al., 2002). Such information on originality variation per species could shed light on such questions as what biotic and abiotic conditions make species the most original and what could be the consequences of the loss of these species. In this chapter, we found that some plant species became more original with increased urbanization at local level compared to their regional originality (ratio of local over regional originality). This could indicate that some urbanization processes filter species regarding their traits and evolutionary history from regional pool making



them more or less original at local level (see Williams et al., 2009, 2015). Urbanization is often characterized by habitat fragmentation creating a mosaic of ecological niches and thus promoting coexistence of many species with different trait values in the same place, thus increasing species trait-based originality and diversity. However, higher trait-based originality in urbanized areas could be also explained by higher competition and limiting similarity within species communities. These hypotheses could be tested with null models, comparing the observed originality patterns and those created with random species distribution (Emerson and Gillespie, 2008; Cavender-Bares et al., 2009). Our exploratory results described species-specific originality patterns, yet more data would be needed to test the hypothesis explaining possible factors and causes of these patterns. For example, including species abundance into originality calculation would certainly decrease the originality value of the most common and the most original species like *Poa annua* (see Chapter1, Figure 5 and Table 2) depending on the abundances of the other species. Thus, using species abundance could reveal influence of interspecific originality and change our conclusions about causes of species originality variation.

Nevertheless, species trait-based originality in this study was inferred from the mean trait values for each species and its variation only depended on the composition of species plots and land cover type. A step further to investigate species originality would be indeed to measure trait-based originality at the individual level, i.e. giving an originality value to each individual of a species. Indeed, intraspecific originality variation could better quantify the variation of trait-based (functional) diversity patterns in communities and increase our understanding of major questions of community ecology (Violle et al., 2012; Siefert et al., 2015; Zhao et al., 2019).



**Figure S7.** Variation of trait-based originality of (A) *Poa annua* (n=319) and (B) *Erigeron sumatrensis* (n=76) according to the percentage of urbanization within 1x1km grid cell (grid cell corresponds to the Vigie-flore protocol). One data point represents one species score. Colors represent the land cover type of a plot where species occurred. Black curves represent the estimated trends retrieved from the mean originality GLMM models.

### Comparing data issues and methodological choices with study of Veron et al.

Here I would like to briefly compare some common problems found in our study presented in this chapter and Veron's et al. (In prep., see Appendix A) study, where I contributed as a co-author by conceiving ideas at early stages and participating into manuscript editing. Veron et al. explored the originality of insular and continental monocotyledon species by using six quantitative traits (from TRY database). Amongst questions they investigated was the relation of a species' trait-based and phylogenetic originality with its geographical range. Are original species also rare? Veron et al. used the same originality index that we used for urban originality study (the AV index) and same taxonomic group of plants (although in Veron et al. there were only Monocotyledon species). Therefore, I will compare problems concerning the sensitivity of trait-based originality measures to the amount missing trait data, to the method of missing data correction and its relation to species abundance.

#### Sensitivity of originality indices to missing trait data

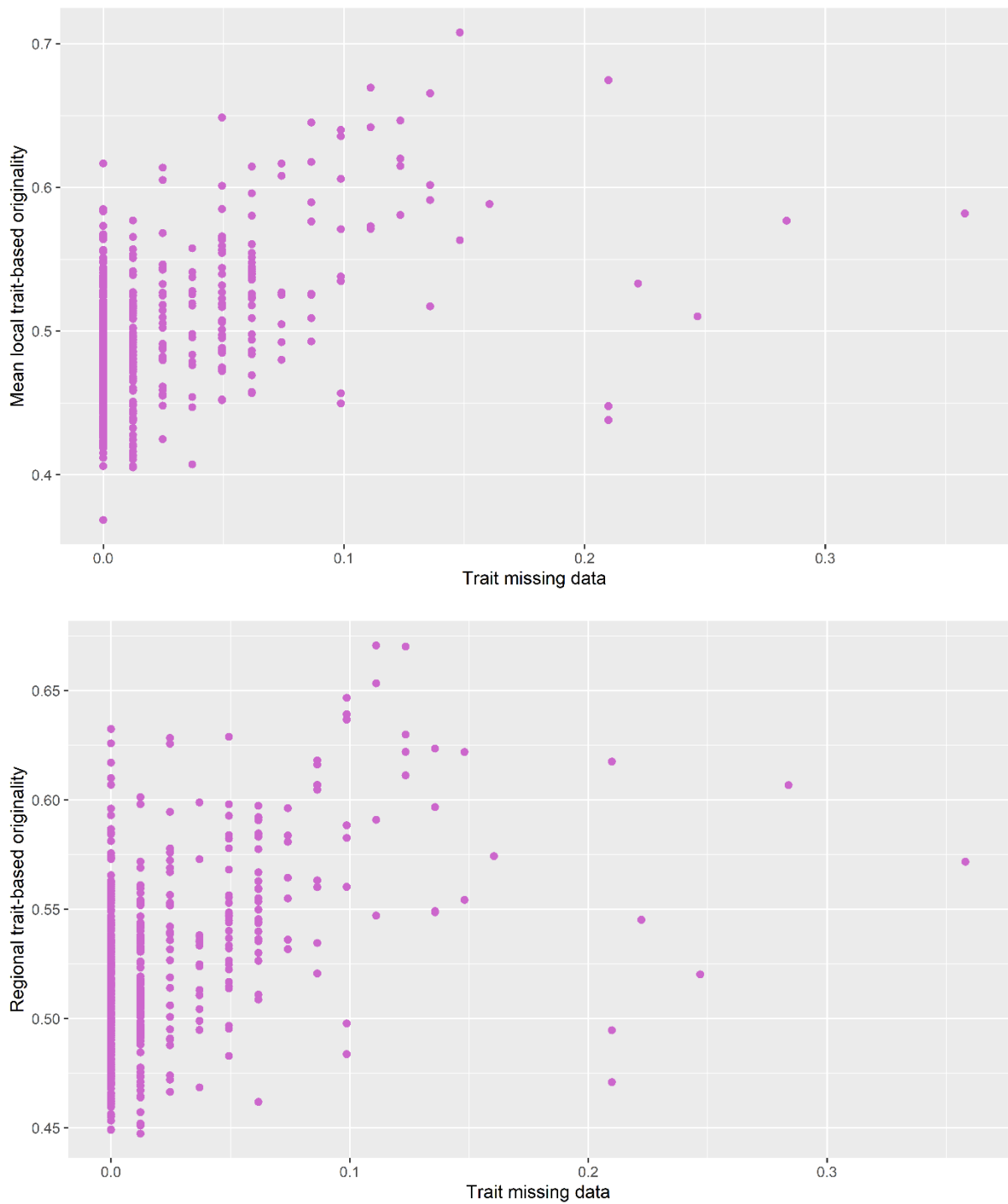
In this chapter, we did not explicitly investigate the sensitivity of originality metric AV to the

amount of missing trait data. However, we found that the top 5% of the most original plant species had a relatively high amount of "Not-Attributed" trait values (see Table S5). Moreover, the missing trait data could influence not only species originality but also a possible biological interpretation of its distribution. We tested a correlation between trait-based originality and the amount of total missing trait data per species which was positive but moderate (Spearman's correlation coefficient  $r = 0.47$ ,  $p < 0.001$  with mean local originality,  $r = 0.43$ ,  $p < 0.001$  with regional originality, Figure S8). To the contrary, Veron et al. (In prep.) found a negative low correlation between trait-based originality and the amount of total missing trait data per species ( $r = -0.14$ ,  $p < 0.001$ , see Appendix 1 in Veron et al.). The difference between the negative correlation between trait-based originality and the percentage of total missing trait data found by Veron et al. and the positive correlation found by Kondratyeva et al. is more likely due to difference in the type of traits considered (Table S6).

**Table S6.** Comparison of data and methods between 2 case studies, Kondratyeva et al. (in revision) and Veron et al. (in preparation) which could have implications in species originality variation. For more details see respectively Table 1 (above) and Appendix A.

	Kondratyeva et al. (in revision for <i>Frontiers Ecol. and Evol.</i> )	Veron et al. (in preparation)
Number of traits	11	6
Type of traits	multiple (see table S1)	quantitative (see Diaz et al., 2016)
Amount of overall missing trait data per species	min = 0%, max = 35%	min=0%, max = 83%
Amount of overall missing trait data per trait	min = 0% , max = 19.8%	min = 23%, max = 81% (see Appendix 1)
Number of species with missing trait data	250/586	2022/2281
Method of distance matrix creation	Extended Gower's distance adapted to trait types (Pavoine et al., 2009)	Gower's distance (Gower, 1971)

There is also a quite possible an effect of the geographical origin of species and trait data. In urban originality study, we used species common to northern-western region of Europe with well-established trait data coming from multiple standardized field and experimental sources. Veron et al. studied species dispersed across the world; therefore, differences in data collecting and availability are probably high. A frequently encountered sampling condition, where rarer species are often poorly studied and likely characterized by missing trait information, is particularly true for some insular endemic species. For example, by creating a growing missing trait information by deleting trait values from a complete data set, Méjeková et al. (2016) found that some trait-based diversity indices were more sensitive than others were to missing trait data. They supposed that the missing trait information in rare species would remove extreme or outlier trait values, thus it could underestimate species originality. Indeed, because of interdependence of diversity and originality metrics, findings on trait-based diversity could be tested on the originality measures as well.



**Figure S8.** Relation between trait-based mean local originality per species (top), regional originality (bottom) and the percentage of total missing trait data. Each data point corresponds to a plant species.

To deal with the problem of missing data, several strategies can be adopted (Pakeman, 2014; Penone et al., 2014). The first one would be to reinforce trait-sampling effort for studies contributing to global databases and to maximize research effort of potential data sources of trait

data. In this chapter we adopted the approach of maximizing data sources, as our species pool was relatively small (586 species). However, Veron et al. had much larger number of species (2281 species for trait-based originality) with many poorly studied species. Therefore, Veron et al. chose to impute missing trait values. Among the multiple methods of imputation described in the literature, they selected a random forest method, which imputes missing trait values by comparing it to the non-missing values of other species (missForest package, Stekhoven and Buehlmann, 2012). After running 1000 random imputations of this type, they found a significant positive but moderate Spearman's correlation between the trait-based originality computed before and after trait data imputation ( $r=0.65$ ,  $p<0.001$ ). To test the influence of imputation method one can also artificially delete at random some amount of trait data and compare it to the complete dataset (Mějčcova et al., 2016), however, complete trait datasets are usually not available. In Veron et al. only 259 species had all values of all traits for which they deleted at random 2/3 of trait data and then proceeded an imputation of deleted values. There was a positive moderate correlation between trait-based originality computed on a subset of species with complete trait data and on species with created and then imputed missing data (Pearson's correlation coefficient  $r=0.41$ ,  $p<0.001$ ). Hence, the chance of imputing an original trait value for a poorly known species is more likely smaller than the chance of imputing a common trait value. Depending on how missing data are distributed across species (at random, not at random or partially at random), the result of the imputation would therefore differ (Penone et al., 2014). Thus, more research is needed to increase our knowledge on the sensitivity of species originality measures to missing trait data.

#### Relation of species abundance-based and trait-based originality

We noticed that the majority of the top 5% of trait-based original species appeared only few times across local communities (Table S3). It could confirm previous findings that more original species are rarer than expected by chance (Mouillot et al., 2013a) and are more likely to become extinct (Gaston 1994). However, we did not have information on species abundances, as abundance data are not required by the Vigie-flore protocol (see Appendix B). The only information we disposed

was species frequency in each of the 1m<sup>2</sup> sampling quadrats of one plot. This led to species characterized by 10 possible cases of "rarity". Therefore, we did not find any strong pattern in the correlation between this discretized estimation of rarity and species local trait-based originality (the Spearman's correlation coefficient  $r = -0.06$ ,  $p < 0.001$ ; unpublished results). That is to say, one of the main objectives of Veron's et al. (in prep) study was to compare insular plant species geographical range (or rarity) with its originality. They found that the most phylogenetically original species (especially those in the top 5%) were more range-restricted and occurred on fewer islands than expected by chance, although this originality-abundance relation was not strong (Pearson's correlation coefficient  $r = -0.15$ ,  $p < 0.001$ , see Table 5 in Veron et al, in prep.). However, only endemic (rare) species with moderate trait-based originality had a smaller geographic range than expected, and very original species occurred on more islands than expected by chance (Pearson's correlation coefficient  $r = 0.43$ ,  $p < 0.001$ ; see also Table 3 in Veron et al., in prep.). Many causes could explain the differences of restrictedness of trait-based and phylogenetic originality discussed by Veron et al. (in prep.).

More generally, species originality and abundance-based rarity were investigated by several studies (e.g., Jetz et al., 2014; Thullier et al., 2015; Grenié et al., 2018; Stein et al., 2018) and their contradictory results may have been influenced by the same data problems as discussed above. If the most original species are actually under extinction risk or are at least more vulnerable to environmental changes because of their rarity, this information could be crucial to guide conservation priorities.

GENERAL DISCUSSION

**ALL IS A MATTER OF CHOICE AND  
AVAILABILITY**

---



*“Just when you think you know something, you have to look at in another way. Even though it may seem silly or wrong, you must try.”*  
John Keating in Dead poets’ society.

In this thesis, we have sought to confine the attention on the originality metrics in ecology and conservation biology. We have demonstrated main benefits of originality measures completing our understanding of biodiversity. In 1991 Vane-Wright et al. stated that measures of species diversity at that time were inadequate for biodiversity conservation priorities because they did not use species distinctness treating different species as equally valuable. Vane-Wright et al. developed a measure of taxonomic distinctness, called taxic weight, improved by May (1990), which opened a way to a new group of metrics in ecology – species originality. The “agony of choice” in conservation would push policymakers to unreasonable wish to protect one member of each taxonomic group. According to the originality measures, the priority areas for conservation could be designed by accounting for species differences within species communities, on which they are dependent, and to which they contribute by being different in their traits and evolutionary history. That is, species should not be equally considered. We explored in this thesis, phylogenetic and trait-based differences between species and thus phylogenetic and trait-based originality.

When they appeared in the 90’s, originality measures were erroneously classified as diversity measures, which they are not. Hence, their use was abandoned in community ecology and conservation biology for the benefit of alternative diversity measures, such as Faith’s (1992) Phylogenetic Diversity index, considered more convenient (Pavoine et al., 2005). However, species originality metrics do not measure an amount of diversity at the species level; instead, they evaluate the contribution of each species to the diversity of a delineated set of species, defined at specific spatial and taxonomic levels. That is, the originality of a given species is dependent on a reference set of species and varies if the composition of the set changes. Even if species originality is not a measure of diversity per se, the diversity of a species set, however, depends implicitly on species originality. The more species are different from each other, the more they are original, thus

increasing the diversity of the species set. We discussed in the Chapter 1 some theoretical examples of such a dependence. Indeed, many diversity measures can be seen as functions of species originality (mean or sum of species originalities) even though they were not developed as such. The link between diversity and originality is thus both conceptual and mathematical, as indices of species diversity and originality are interdependent and use common concepts and data. Performing two-steps analysis by first measuring species originality and then evaluating species diversity thanks to these originalities has the advantage of giving simultaneous information on the level of diversity of a species set and on the contribution of each species to that level. In other words, we can easily switch between originality and diversity measures. In Chapter 2, we used such a switch by calculating the mean and the skewness (asymmetry) of species originality values to evaluate how urbanization affects the trait-based and phylogenetic compositions of plant communities. In this case study, the mean of species originalities corresponded to a measure of trait-based or phylogenetic diversity, depending on whether species were characterized by their traits or by their phylogeny. The skewness (or asymmetry) of species' originalities informed on the shape of the distribution of originality values and can be interpreted as the proportion of uniqueness versus redundancy in a set of species.

This thesis started with the objective of evaluating an impact that the concept of originality could have in ecological studies. Two chapters that are at the core of this thesis highlighted links between the concepts and measurements of diversity, originality and rarity and provided a first test of the joint analysis of originality and diversity in urban ecology. During the last three years, we were, however, repeatedly confronted to the reality of data analysis, with missing data and plethora of potential mathematical formulas to evaluate diversity and originality, each formula having its own interests and drawbacks. In addition, our knowledge on biodiversity, although always growing, is still incomplete. Our conservation efficiency thus depends upon series of choices made in scientific studies, notably: the choices made about which species we focus on, which biological characters we consider, at which spatial scales, and with which measures of species diversity and

originality. We will discuss below how some of these choices affect originality measures and their interest for ecological and conservation studies.

## **1. CHOOSING AND OBTAINING THE DATA FOR ORIGINALITY MEASURES**

Biodiversity is declining across the planet, but its taxonomic, trait-based, and evolutionary aspects are declining at different paces. Thus, understanding how these aspects respond to the global change and what effect they have on ecosystem functioning is crucial to better prevent the ecological consequences of biodiversity loss. Most current studies of species originality consider species' phylogenetic relations at large spatial scales for studying evolutionary history of entire groups of species in a context of conservation biology (Safi et al., 2013; Jetz et al., 2014; Pollock et al., 2017). In contrast, community ecologists report processes of species co-existence acting at local scale (Mouillot et al., 2013b; Garnier et al., 2016; see also Chapter 2, section II.5). However, evolutionary and ecological processes both structure species local communities (Lessard et al., 2012). There are more and more studies coupling ecological and evolutionary sides of species biology with the common aim to protect biodiversity at global scale (e.g., Mazel et al., 2014; Brum et al., 2017; Le Bagousse-Pinguet et al., 2019; Kosman et al., 2019). Besides, multifaceted and multiscale approaches for biodiversity evaluation have been adapted by Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) in Conceptual Framework for nature benefits to people (Diaz et al., 2015; Brondizio et al., 2019, in prep.). Indeed, as described in the Introduction, every species (if we talk about biological concept of species in macro-ecology) can be described by its abundance, phylogeny and trait values or at least one of those. Below, we will discuss the benefits and shortcomings of each facet of species originality measuring.

### **1.1 On the benefits and limits of using different aspects of species' biology**

#### 1.1.1 On the benefits of species abundance analysis

More informative than species occurrence data (presence/absence), abundance can be measured by total number of individuals, but also by its density (number of individuals within a standard

sampled quadrat), percentage cover or total biomass. Species abundance is the first and simplest step to estimate how much species contribute to biological diversity. We should, however, be aware of sampling biases when estimating species abundance, depending on where, when and how species were observed (Brown, 1984; Gaston, 1997, Buckland et al., 2005, see Chapter 1).

Depending on how it is measured, abundance informs on species' commonness (and rarity), distribution range and habitat preferences. Furthermore, the notions of species' dominance and specificity to habitat type are also based on species abundance. According to the “mass ratio” hypothesis (Grime, 1998), species influence ecosystem functioning in proportion to their biomass and to the extent of their functional trait values. Thus, dominant species would have the most of the influence on ecosystem functioning independently of the richness of subordinate species. All information about community change is therefore contained within the patterns of species abundance distributions (Buckland et al., 2005).

Species abundance is determined by local environmental conditions and regional species pool (Shipley et al., 2006). In return, species' relative abundances would determine the number and strength of interactions between species (Vázquez et al., 2007). Hence, abundance-based rarity of species is tightly involved into relationship between species diversity and ecosystem functioning. The reasons why species are dominant or rare in their abundance is related to their biological characters and their ecological performance, constrained by community assembly mechanisms (e.g., Adler et al., 2014). Thus, the distribution of species in a certain environment depends not only on their abundance but also on the degree of dissimilarity with other species in their biological characters (Shipley et al., 2006), i.e. their originality. Hence, species affects ecosystem functioning via phylogenetic and trait-based originalities. Analyzing both abundance and species' biological characteristics may thus inform on the processes underlying variations in biodiversity.

As shown in Chapter 1, originality measures can integrate abundance data, leading to species abundance-based originality. The originality value of a given species can change if weighted by species abundance (see Figure 5 in Chapter 1) and thus represents a tool to emphasize the rarest

species whose individuals have the rarest biological characteristics (Cadotte and Davies, 2010; Mouillot et al., 2013b; Pavoine et al., 2017). An increase through time of the abundance-based rarity of a species might indicate that profound changes in the ecosystem functioning are ongoing. Some even included extinction risks in phylogenetic originality measures, where extinction risk is partly determined by a decrease in population size, i.e. in a number of individuals (Isaac et al., 2007; Steel et al., 2007). These measures allow to detect species that are both original and threatened (Isaac et al., 2007) as well as species that are original and belong to phylogenetic clades where many species have a high risk of extinction (Steel et al., 2007). If conservation actions are based on such measures, they could prevent the loss of species with unique biological characteristics.

#### 1.1.2 On the benefits of trait-based analysis of diversity and originality

Biological characteristics of species can be linked to environmental variables, illustrating species adaptations to environmental conditions (Diaz and Cabido, 2001). Trait-based (functional) diversity and originality inform us on species' ecological dissimilarities, related ecological niche breadth and ecosystem functions and services (Violle and Jiang, 2009; Cadotte et al., 2011, Mouillot et al., 2013a, Hidasi-Neto et al., 2015). Therefore, trait-based approaches have increasingly gained many areas of ecological research going from community assembly (McGill et al., 2006, Cavender-Bares et al., 2009, Mason et al., 2013), ecosystem functioning (Herben and Goldberg, 2014), ecosystem services (Lavorel et al., 2011; Diaz et al., 2007b, 2018) and species' invasions (Rejmanek and Richardson, 1996; Kühn et al., 2017) to conservation biology (Grenié et al., 2018; Kosman et al., 2019), urban ecology (Vallet et al., 2010; Knapp et al., 2008, 2012; Palma et al., 2017) and many other areas. One of the common goals of those studies is to understand how species' traits mediate community assembly and species coexistence and thus to predict the effects of global change on biodiversity.

Today it is still unclear which traits should be defined as functional (Mlambo, 2014). Violle et al. (2007) provided one of the most comprehensive and most cited definitions to date: species functional trait should be measurable on individual level and have the influence on species fitness,

i.e. its ability to survive and reproduce. Even if there is no universal definition of term "function" itself in ecology (see Bellwood et al., 2019), the use of any trait defined as functional should go with an explanation of a function related to that trait. Classifying species' traits into response traits (involved into species' responses to environmental processes, measured by its overall fitness) and effect traits (involved into mechanisms by which species influence ecosystem processes) is a good beginning (Hooper et al., 2005). That is, in plants, such traits as specific leaf area, plant height, and seed mass are used as the principal traits to describe three fundamental axes of phenotypic variation amongst plants, which are respectively related to the acquisition and use of resources, the competitive ability, and the capacity for sexual regeneration (Diaz et al., 2016; Garnier et al., 2016). Species' traits without clear functional significance should not be referred as functional (Cadotte et al., 2011), as for example Ellenberg values (Ellenberg et al., 1991; Garnier et al., 2017) and Grime strategies (Grime et al., 1988) or else life-history traits (González-Suárez and Revilla, 2013). However, originality can be measured with any trait data making sense in a specific context. This is why in this thesis we decided to use more inclusive terms of "traits" or "biological characters".

The trait-based approach in biodiversity studies turns out to be particularly useful in establishing links between species performances and ecosystem functioning processes at local scale (Loreau et al., 2001). Indeed, the range of trait values present in a community was found to be a good predictor of ecosystem functioning in several studies (Tilman et al., 1997; Petchey & Gaston, 2006). Trait-based or functional ecology offers a promising way to model deterministic processes of local community assembly and improves our way of understanding local diversity variations (Diaz and Cabido, 2001) and increases the amount of benefits that humanity could gain from biodiversity (Diaz et al., 2007b). In line with trait-based diversity, the main assumption in trait-based originality approach is that the most original species with more unique combinations of traits would contribute more to the ecosystem functioning or may maintain specific, irreplaceable functions and processes (Mouillot et al., 2013b; Brandl et al., 2016; Leitão et al., 2016). Recently, models were developed in functional ecology in order to depict the mechanisms determining

species trait-based differences and to identify traits involved into community assembly processes (Jabot, 2010; Münkemüller and Gallien, 2015; Munoz et al., 2017). That is, mechanistic insights into patterns of community assembly may rely more on the trait-based originality of species than on their abundance alone (Cadotte & Davies, 2016).

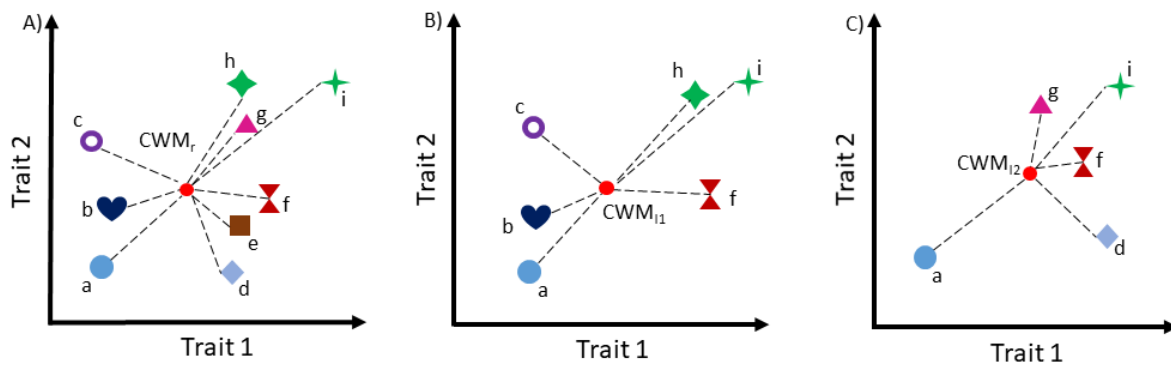
Many community ecology studies have investigated a community-weighted mean (CWM) of trait values (a mean of community-level trait values weighted by species abundances) along abiotic gradients to describe species' trait distribution compared to the mean value of community (Garnier et al., 2004; Cornwell et al., 2006; Diaz et al., 2007b). Other researchers related specific trait variations between and within species to particular environmental conditions (Dwyer et al. 2014; Siefert et al., 2014). A recent work on trait-environment relationship proposed a concept of species functional optimality, where species trait values are distributed according to a Gaussian function around a community optimal value of functional traits (Mucarella and Uriarte, 2016). The community-weighted mean (CWM) and variance (CWV) of trait values are thus expected to depict the optimum and the intensity of environmental filtering, respectively (Denelle et al, 2019). Trait-based originality measures could be seen here as a connection between species trait-based diversity and trait distribution of a community. Indeed, the deviation between species trait value and CWM, as a measure based on trait-based distances between species, is similar to originality (Violle et al., 2017a). In addition, the CWM is the weighted mean of the squared deviations between species trait values and CWM and thus a mean of originality values, equal to the variance of trait values and related to a well-known index of diversity – Rao's quadratic entropy (e.g., Pavoine et al. 2004).

When several traits are jointly considered, the use of multidimensional functional trait space may allow visualizing a species position relatively to the other species (Villéger et al., 2008). In multidimensional trait space, trait-based originality can be measured as the distance from a species' position to the community multivariate centroid (the hypothetical average species, Buisson et al., 2013), which actually represents the multivariate CWM of the community (Muscarella and Uruarte, 2016) or else the optimal value of species traits enabling species to persist in the community

(Denelle et al., 2019). Species located closer to the CWM would be less original, and contribute less to the community trait-based diversity (Figure 1). For example, Muscarella and Uriarte (2016) found that 25% of the tree species in Puerto Rico significantly opposed CWM-optimality correlation for at least one trait (by being distant from CWM), thereby being more original and contributing greatly to local trait-based diversity. To the contrary, 75% of species were closer to the community optimal trait value, i.e. they were more redundant in their traits. Therefore, variation in trait distributions of a community influences species originality, as it is relative to other species' trait values. Thus, species trait-based originality may reflect the trait-based structure of communities and help to assess the relative part of each species into the trait-based diversity variation under future global changes. In this manner, species trait-based approach is mostly useful at the local community level for determining how species trait differences shape species originality and their capacity to colonize new communities just as much as their influence on ecosystem functioning. Thus, each study can adapt its choice of traits to the underlying assembly processes of interest.

However, traits are related to phylogeny by the underlying evolutionary processes. Indeed, each species is a product of evolutionary processes, during which the species acquires its biological characteristics (Darwin, 1856; Harvey and Pagel, 1991). Thus, while trait-based (functional) diversity and originality may inform us on the contribution of each species to ecosystem functions and services, phylodiversity and originality could provide information on the evolutionary and biogeographic histories of taxa. Thus, species trait evolution must also be taken into account (see article in Chapter 1, Section IV.1) to bridge community trait-based diversity and phylodiversity to community assembly processes (Losos, 2008).





**Figure 1.** Theoretical illustration of species trait-based originality. Nine species presented by geometrical colored figures (a – i) are projected in a two-dimensional trait space (trait 1 = shape and trait 2 = color). The red points represent a community multivariate centroid (the hypothetical average species) or else a Community Weighted Mean of A) a regional species pool ( $CWM_r$ ), and of two local species communities B)  $CWM_{l1}$ , C)  $CWM_{l2}$ . Species originality is measured as a distance from a focal species to the centroid (dashed lines). B) Species i is the most original in the region, however its originality varies between communities and is higher in community B) than in community C). *Inspired from Buisson et al., 2013.*

### 1.1.3 On the benefits of phylo-based analysis of diversity and originality

Today, the large availability of DNA sequences in datasets such as in Genebank together with efficient computer algorithms facilitates the reconstruction of species evolutionary histories represented by phylogenetic trees with previously unseen completeness (Diniz-Filho et al., 2013). Therefore, phylogenetic information in conservation biology and community ecology has become increasingly reliable and meaningful (see Mouquet et al., 2012).

Although Phylogenetic Diversity measure (PD, Faith, 1992) has been recently included to the IPBES Biodiversity Assessment (IPBES, 2019, Diaz et al., in press), underpinning its importance for prioritization of species and areas for conservation, the confusion remains about which biodiversity measure to use in conservation. Due to their phylogenetic relatedness species are not independent units of biodiversity (Felsenstein, 1985) and should not be treated otherwise. Evolutionary-based approaches reveal that species embody different amounts of unique versus shared evolutionary information, which can be measured by species originality. Indeed, species phylogenetic originality has been used to identify sites or taxa with particular conservation value (Barker, 2002; Isaac et al., 2007; Cadotte & Davies, 2010). Yet, current conservation actions often

ignore species phylogenetic originality, targeting areas with high species richness and degree of endemism (Veron et al., 2017). An exception would be the EDGE program (Evolutionarily Distinct and Generally Endangered) at the Zoological Society of London. It is probably the only conservation initiative that really uses species phylogenetic originality measure (Evolutionary Distinctiveness or ED) weighted by species IUCN threat status as an estimate of its extinction risk thus focusing conservation efforts specifically on threatened species (Isaac et al., 2007). However, the EDGE index has been criticized for considering species extinction risks and species' phylogenetic originality as independent variables, whereas the extinction of a given species would increase the originality of its sister species (Faith, 2008). The EDGE index was thus improved into the HEDGE index that integrates the probability of extinction for species and its close relatives (Steel et al., 2007). Therefore, global broad-scale conservation schemes can be established (Isaac et al., 2007).

However, most conservation actions focus on the biodiversity within areas and not on the originality of each species. We must recall that originality is not diversity. Protecting only the most phylogenetically original species will not necessarily maximize phylodiversity. Indeed, the set of species maximizing phylodiversity has to be dispersed across the tree, while phylogenetically original species may in some cases be phylogenetically clustered (Pavoine et al., 2005). Empirical studies have, however, shown that subsets of species with high phylogenetic originality capture more phylodiversity than at random (Redding et al., 2008) and that ranking mammal species for conservation priority with a metric used in the EDGE program, could protect more phylodiversity than if species selection was based on phylogenetic originality value or on threat status alone (Redding and Mooers, 2015b).

In the human-centric benefits of evolutionary history for focusing conservation actions on preserving evolutionary history, Tucker et al. (2019) identified six general arguments (such as "Does evolutionary history have surrogacy value for phenotypic diversity?" or "Does evolutionary history or phenotypic diversity provide greater evolutionary potential and future options for humanity?"),

gave a logical support for or against, and future directions for each argument (Table 2 in Tucker et al., 2019). However, some arguments are still poorly empirically supported, although optimizing the total amount of evolutionary history under protection should enhance the protection of important species features useful to humanity, with features defined as the variety of states across homologous traits (Lean and Maclaurin, 2016).

A common assumption is indeed that maximizing the amount of protected evolutionary history should capture the variation in species characteristics that are otherwise unknown, supporting the intrinsic value of biodiversity (Faith, 1992). Phylogenetic trees have thus been widely used with the assumption that shared ancestry provides shared biological traits (Webb et al., 2002; Redding and Mooers, 2015b). In that way, the evolutionary potential of species represents species' intrinsic and option values for the future potential uses by humanity (Faith, 1992). Thus, taking into account phylogenies in conservation schemes means to plan the protection of potential genetic and trait-based diversity (Sarrazin and Lecomte, 2016). Moreover, when there is a lack of understanding of the relationship between species traits and ecosystem functioning or which traits matter the most in a particular environmental conditions, phylodiversity appears to be a useful tool for estimating species feature diversity (Faith, 1992; Webb et al., 2002; Tucker et al., 2018). It therefore appears to be a reasonable assumption that preserving phylogenetically distant and original species supports trait-based diversity (Redding and Mooers, 2015b; Tucker et al., 2019). However, the use of phylodiversity to represent trait-based diversity has been strongly debated and studies showed contradictory results (Flynn et al., 2011; Mazel et al., 2017, 2018, 2019; Owen et al., 2019, see the next section). Such results brought sometimes an uncertainty about the worth of valuing species phylodiversity as a proxy of feature diversity, creating a gap between produced scientific theory and its application in conservation actions (Winter et al., 2013).

Moreover, phylogenetic originality informs on the evolutionary history events such as immigration, speciation, extinctions and divergence of species characters, which are at the basis of species trait-based differences. However, in search for general patterns and rules governing

community assembly based on species' ecological differences, direct use of trait-based approaches of species' diversity and originality could be more informative (Best et al., 2013, Cadotte et al., 2013).

#### 1.1.4 Are phylogenies good proxies for species traits?

Since two decades, many researchers in ecology and conservation biology have used phylogenies as a proxy for species ecological similarities (Webb et al., 2002; Cavender-Bares et al., 2009; Mouquet et al., 2012; Strivastava et al., 2012, Cadotte et al., 2013). Some studies even argued that phylodiversity and trait-based diversity might be redundant measures (see Tucker et al., 2018; 2019 for the examples). Some even proposed to combine both distances when assessing species assembly rules (“traitgrams” in Cadotte et al., 2013). The assumption made by these studies is that since phylogenies represent the accumulation of trait evolution among species, closely related species should be more redundant in their traits, and thus, phylodiversity could be a proxy for trait-based diversity (Gerhold et al., 2015).

However, recent studies have challenged the assumed relationship between evolutionary (phylogenetic) distance and ecological trait-based similarity. Cadotte et al. (2017) presented seven reasons why the seemingly clear link between species ecological differences and their phylogenetic distances may actually not always be straightforward. These reasons are related to evolutionary and ecological processes and events (high evolution rates, species assemblage rules different from expected) or are due to shortcomings of experimental design (use of only one evolutionary model of traits, poor species phylogenetic resolution or wrong evolutionary processes tested). More generally, the phylogeny-traits relationship is inherently complex and can take many forms. It is thus important to analyze trait evolution rather than assuming that phylogenetic distances among species are proportional to trait-based distances (Letten and Cornwell, 2015). Models of trait evolution can help us to understand the so-called phylogenetic signal – an expected covariation between trait differences and phylogenetic distances – and distinguish it from random trait distributions (Blomberg et al., 2003). Thus, the strength of phylogenetic signal is a measure of the

relationship between phylogeny and trait values, and many metrics of phylogenetic signal exist varying in their sensitivity to the phylogenetic topologies (number of tips, presence of polytomies and availability of branch length information, Münkemüller et al., 2012). Yet, many different models of trait evolution exist (Felsenstein, 1985; Hansen, 1997; Beaulieu et al., 2012) and there is thus no universal model, especially when multiple traits are in focus. In addition, several models may predict the same trait evolution equally well (Revell et al., 2008).

Pavoine et al. (2013) demonstrated that the existence of phylogenetic signal in traits is not sufficient to ensure a strong correlation between phylogenetic diversity and trait-based diversity. Later, Tucker et al. (2018) found that a positive correlation between phylodiversity and trait-based (functional) diversity of simulated data becomes stronger with increasing trait dimensionality, but varies depending on the form of evolutionary model, and on the evolutionary complexity of studied traits. Regarding originality, Pavoine et al. (2017) used simulated and real data to show that the correlation between phylogenetic and trait-based originality increases with the number of traits considered but is dependent upon the evolutionary model (speciation events at random, close to tips or to root) and originality index used. In their real case study on European Carnivores, phylogenetic and trait-based originalities were not significantly correlated and traits had a weak phylogenetic signals.

In a recent theoretical study, Mazel et al. (2017) argued that phylodiversity (measured by Faith's PD) is not a good surrogate for preserving the diversity of traits. By using biologically plausible scenarios of trait evolution, they found that phylodiversity-based conservation priorities selects less of trait-based diversity than at random. In a related empirical study, i.e. using a real dataset, Mazel et al. (2018) found that prioritizing the most phylodiverse sets of species results in an average gain of only 18% of trait-based diversity relative to a random choice of species sets. However, Mazel et al. (2018) stayed aware that they used a limited set of traits and probably more traits information is needed to test the relevance of Faith's PD for predicting features diversity.

For sure, phylodiversity and trait-based diversity are intimately related, but certainly not on

a linear manner and complex evolutionary mechanisms are behind species characteristics underpinned by environmental influence (Tucker et al., 2019). Thus, rather than searching for the phylogeny-traits duality, one should measure both aspects, phylogenetic and trait-based, for diversity and originality, searching how their joint analysis could benefit to ecological studies. In addition, proving that phylodiversity is correlated with a measurable trait-based diversity (diversity of functional traits or trait richness, Pavoine and Bonsall, 2011) is different from assuming that phylodiversity could represent an unmeasured part of trait-based diversity and maintain future benefits to humanity, both known and unknown (“option values”, Faith, 1992, see Introduction, 2.4).

Therefore, while planning conservation actions, policy makers would like to be reassured that, by using a given measure of species diversity or originality, we encompass the right aspects of biodiversity that we want to protect. However, the concepts of diversity and originality are so multidimensional that a single mathematical formula, relying on a finite set of data, would never encompass all these aspects. In a comment to Mazel et al. (2018), Owen et al. (2019) underlined that further exploration of phylogeny-trait link is needed by using different diversity (and originality) measures on multiple spatial scales. Nevertheless, the most important in community ecology studies is to recognize that we use only a small part of measurable traits, and that the remaining features (which are not only functional traits) could completely change the assumption of phylogenetic surrogacy for trait diversity and originality. Thus, it is more precocious to use trait-based and phylogenetic approaches as complementary in ecological and conservational studies, each at the adapted spatial scale: global for evolutionary history and local for ecological processes operating on community level (Losos, 2008; Saito et al., 2016, Tucker et al., 2019). Therefore, each additional aspect of species originality would provide new insights into species contribution to the diversity of a set of species.

## **1.2 Issues related to the availability and to the quality of the data**

Ecology, evolution and conservation science requires an exhaustive knowledge on the distribution

of species and their biological characters. Therefore, standardizing protocols of species inventory is of primary importance to many ecological studies. Standardized data makes possible to combine several datasets in one study and to make generalizable assumptions about biodiversity patterns. Nevertheless, it is not only about collecting data, but also about its representativeness of the available information including raw and processed data, its storage, managing, processing and sharing (Merino et al., 2016). A growing availability of ecological data has broadened the scope of questions in biodiversity assessment. There are several ways to assess data on biodiversity: 1. big online depositories and networks such as Knowledge Network for Biocomplexity<sup>3</sup> (KNB), Global Biodiversity Information Facility<sup>4</sup> (GBIF) mostly for species occurrence data, TRY<sup>5</sup> for plant trait data (Kattge et al., 2011), PanTHERIA for mammal traits (Jones et al., 2009) or Dyad<sup>6</sup> where multiple contributors and stakeholders share data; 2. multiple softwares allowing finding, extracting and cleaning data easily (i.e. R (R Core Team, 2019), Morpho (Higgins et al., 2002), Data Retriever<sup>7</sup>) and techniques of machine learning allowing to generate more data (Peters et al., 2014; Thessen, 2016); 3. network science approaches assembling researchers to collect data together (LTER<sup>8</sup>, DataONE<sup>9</sup>). A global objective of scientific community is to share standardized data from different sources stored at interoperable portals to improve data availability and quality (Hampton et al., 2013). Thus, the use of ecological data serves to analyze specific hypotheses (by estimating model parameters and predicting complex processes) and to explore and generate new ideas and models. However, many ecological datasets are still hold by individual scientists and laboratories, often available on request, but lacking of specific data policy and peer-reviewing (Roche et al., 2015). Some other datasets, which could be useful to ecological research, comes from old collections, museums and herbariums, and were not necessarily designed for modern research use (Pearman et al., 2006). Finally, in Chapter 2, we used the citizen science program of vascular flora survey, “Vigie-

---

<sup>3</sup> knb.ecoinformatics.org

<sup>4</sup> gbif.org

<sup>5</sup> try-db.org

<sup>6</sup> datadryad.org

<sup>7</sup> <https://www.data-retriever.org/>

<sup>8</sup> lternet.edu

<sup>9</sup> dataone.org

flore”. Citizen science was created to assist scientists in a hard task of large data collection, where quantity is as important as quality. Citizen science data is increasingly used to answer important ecological questions and possess a strong educational outcome (Bonney et al., 2009; Dickinson et al., 2010; McKinley et al., 2017; Martin et al., 2019).

Any biological dataset, whether it includes species occurrences, trait values, gene sequences and other aspects of biological information can include missing or erroneous information. It is related to several issues in ecology sometimes called the “dark data” (ODE Report, 2011). In fact, the path of data transformation from its raw to published form still includes hidden processes of data transformation, making it sometimes inaccessible and creating time-consuming and unnecessary replications between studies (Hampton et al., 2013). A new challenge for ecologists is how to better aggregate, search, cross-reference, and mine large volumes of data (LaDeau et al., 2017). Thus, a crucial point before analyzing any scientific data is to determine the maximum of possible pitfalls and limits relative to the study. One of the most common bias of biodiversity data is the information on species geographic range (“Wallacean shortfall”, Hortal et al., 2015). Usually species are recorded in the most accessible locations depending on spatial and temporal sampling effort and even politico-economic situation of a country (Rodrigues et al., 2010). Along with general problems of data availability and reliability, we will briefly discuss the most relevant problems of the data related to the measurement of originality.

### 1.2.1 Phylogenetic bias

If a fully resolved phylogenetic tree existed, it would still possess some topological uncertainties (Swenson, 2009). Each phylogeny is a hypothesis about the real relationships between species represented by specific tree topology, thus different tree topologies might represent differently species evolutionary history (Knapp et al., 2012). Usually phylogenetic trees have to be rooted for calculation of species phylogenetic originality (Thuillier et al., 2015). Sometimes phylogenies also have to be ultrametric (i.e. the shortest path from root to tip is constant). For example, the originality index  $Q_b$  would estimate some species’ originalities as equal to zero if non-ultrametric



trees were used (Pavoine et al., 2005; Pavoine et al., 2017; see Chapter 1). Moreover, the topology of phylogenetic trees changes due to updates in scientific knowledge in some taxonomic groups more than in others because of recent diversifications and extinctions (Morlon et al., 2010; Billaud et al., 2019), resolution of soft polytomies (Kuhn et al., 2011) and lateral transfers of characters between species (Gribaldo and Philippe, 2002). For example, recent extinctions would make the remaining closely related species more isolated on the tree and make them more phylogenetically original (Chapter 1, Figures 7 and 8).

Phylogenetic soft polytomies reflect unresolved nodes that support more than two descendent branches in a tree. Such low phylogenetic resolution can affect species phylogenetic originality (Isaac et al., 2007) by wrongly giving different species the same originality value. This bias would be greater with polytomies occurring deeper in the tree if the originality index is sensitive to the information nearer the root (such as the AVerage index), and greater with polytomies in the terminal nodes for measures that overweight information nearer the tips (such as Equal Splits). Even though polytomies can be resolved by several methods (Lewis et al., 2005; Davies et al., 2008; Kuhn et al., 2011), their real impact in conservation priorities is still poorly understood.

Furthermore, if polytomies are present or if molecular information is missing for some species, terminal branch lengths can be poorly estimated, what would cause wrong representation of species relatedness (Roquet et al., 2013). Phylogenetic branch unit may have length set to 1, or may have units of observable changes in traits (or features; Faith, 1992), or genetic distances based on the molecular markers, or may be inferred from node age in millions of years (Cavender-Bares et al., 2006) or estimated with molecular clock models calibrated by fossils presence (Tucker et al., 2019). The differences in branch-length estimations would thus under/overestimate species phylogenetic originality based on phylogenetic distances.

The whole phylogenetic tree may change depending on the molecular data available (Sanderson and Shaffer, 2002). Plant phylogenies are especially well known to suffer from frequent topological changes with species replacements. The definitions of the species, placed as the tips of

phylogenetic trees can also sometimes be very instable. Therefore, a species could be placed differently on the phylogeny if its definition (limits) change. For example, the split of a species into two species due to updates in taxonomic knowledge modifies estimations of species phylogenetic originality (Robuchon et al., 2019). Therefore, one should use the most robust and recent version of phylogeny available. Some studies have hopefully shown that wrongly placed species did not affect phylogenetic originality substantially (Collen et al., 2011; Isaac et al., 2012), which is true if taxonomic errors are spread across phylogeny and not concentrated within close relatives. However, more studies are necessary to evaluate the effect of phylogenetic quality on species' phylogenetic originalities.

### 1.2.2 Trait data bias

Prioritizing species conservation based on their trait-based originality is a promising strategy; however, it remains challenging in practice because we have imperfect knowledge about species traits relevant for conservation and how they should be integrated into conservation actions. By working with species biological traits, researchers explore trait-based diversity of biological assemblages and seek for generalizable predictions of species ecological functions highly constrained by the type and number of traits considered (Petchey and Gaston, 2006, Adler et al., 2014). Information on functional traits and the extent to which they relate to an ecological function is available mostly for plant species, and still remains limited for other taxonomical groups, although other non-functional traits are available (Lavorel and Garnier, 2002; Diaz et al., 2016). Many studies in functional ecology and community ecology are thus dominated by plant diversity studies and it is not well understood whether patterns observed with plants can be generalized to the other taxa.

Plants are the most studied taxonomic group, because it is the most documented one. Most plant trait databases include mean trait values per species with no information on trait intraspecific variability (Diaz et al., 2016). Thus, it seems sometimes difficult to establish relationships between the performance of a species and a particular habitat if species' traits have been averaged across a

range of different environmental conditions. For example, in Chapter 2 we used trait data from several countries of northern-western Europe, but excluded those from the other European countries, where same species occurred (available in TRY database). Moreover, the majority of species trait values are available for the most common species and for the most easily measurable traits (Violle et al., 2015), which bias species originality to the most abundant species. Therefore, values of trait-based originality are relying on the choice of traits, their number and type, how they were measured and on the way these traits are used to represent dissimilarities between species.

In particular, species traits can be of different statistical types. In Chapter 2, we used quantitative, categorical, ordinal, circular, binary and nominative traits. The type of data will have an influence on how species' originality is perceived. For example, categorical traits lead to binary distances between species: two species are similar (distance=0) if they belong to the same category or they are maximally distant (distance=1) if they belong to distant categories. In that case, species belonging to the less frequent category will be considered highly original compared to others, as they are maximally distant from other species. Quantitative traits lead to more continuous differences between species and thus to more continuous and symmetric distributions of species originalities. Discrete traits could thus lead to the highest distances between species and thus to an overestimation of their trait-based originality (Maire et al., 2015). When working with several species traits, one should also test for collinearity between trait values and exclude highly correlated ones. Then, Gower's measure of species distances (Gower, 1971) allows taking into account multiple trait types and handles missing trait data when calculating trait-based dissimilarities between species. A generalization of Gower's metric into a mixed-variable coefficient of dissimilarity (Pavoine et al., 2009; see Chapter 2 for use) improved the consideration of traits of different types. Although, if the amount of missing data is relatively small to the size of the dataset, then leaving out the few samples with missing features may be the best strategy in order not to bias the analysis, computational bias is still inevitable.

Species trait-based originality measures are sensitive to the missing data, some more than

the others, as found in the case study of Veron et al., (In prep, see Appendix A). Similarly, in our case study in Chapter 2 species with the most of missing trait data had the highest trait-based originality scores, even if in reality they may not be that different from the other species. For sure the best solution to missing trait data would be not having any. Experimentally controlled studies could probably get nearly complete datasets. However, most ecological datasets combine many data sources and thus many potential sources of missing data. The common habit is to delete units with missing observation, which likely increases estimation bias particularly if data is not missing completely at random (Nakagawa and Freckleton, 2008; Penone et al., 2014). For example, Penone et al. (2014) found that removing species with missing trait values (especially with >30%) created more bias than trying to estimate (impute) missing traits.

To avoid inconsistent results, one could try to improve the effort of data collection, for example by combining multiple databases as we did by combining four plant trait databases in urban originality study, and completing it with data from literature, museums and fieldwork. If an important amount of missing data persists, imputation strategies exist (see Chapter 2), replacing missing data points with an estimated value relative to the given data input (Soley-Bori, 2013). Penone et al. (2014) compared the power of four imputation methods of missing trait values by artificially removing an increasing amount of trait data from a complete dataset of Carnivora species. Penone et al. (2014) also included phylogenetic information into the imputation algorithm thus improving trait values estimation; however, this could introduce circularity if we calculated trait-based originality from phylogenetically constrained data and then compared it to the phylogenetic originality. Penone et al. (2014) found that up to 60% of missing data would be an acceptable threshold for a meaningful imputation with four methods they tested. However to date, there were no studies evaluating species originality sensitivity to the missing trait data and to the mentioned imputation methods (but see Veron et al., in prep in Appendix A).

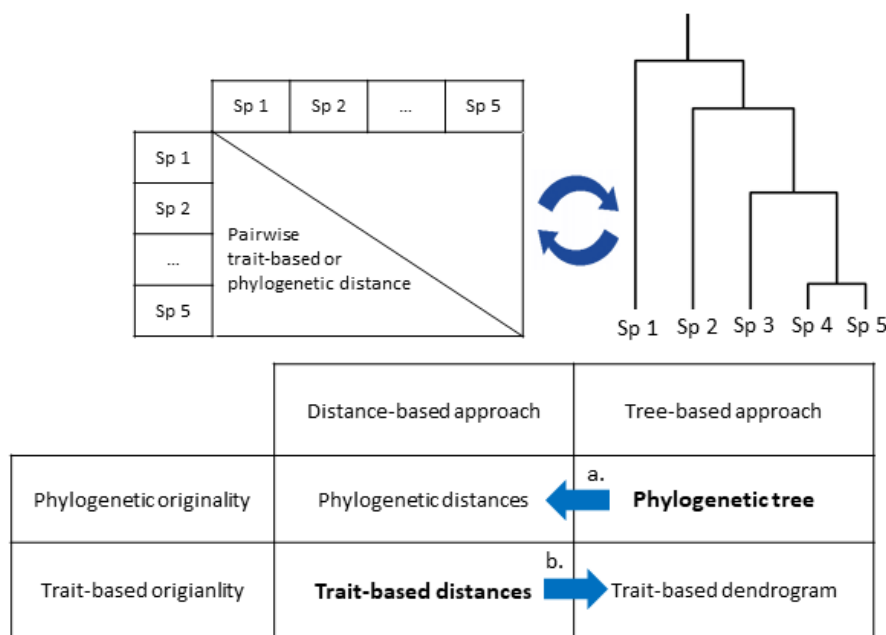
Another source of trait-based bias is the way of representation of species differences. There are three main approaches for representing trait-based differences between species: tree-based,

distance-based and trait-value ordinations in multidimensional space (Chao et al., 2014). Because originality indices were first developed for the phylogenetic tree structure, the most used method to measure trait-based originality is the tree-based approach, by constructing functional dendrograms with hierarchical clustering methods (Figure 2.b; Mouchet et al., 2008). Trait-based dendrograms consider all possible interrelations between species of a set (Chao et al., 2014). Yet, there is no rationale for representing trait-based relatedness between species with a dendrogram (Villéger et al., 2017) and it could misrepresent species' raw distances and bias the estimates of species trait-based originality (Pavoine et al., 2017). Nevertheless, there is a lack of research investigating how trait-based clustering approaches influence trait-based originality values (but see Pavoine et al., 2017; Veron et al., in prep in Appendix A.). We would thus recommend to test the quality of several approaches for trait-based distances representation compared to raw trait differences and to choose the most appropriate one. For example, Veron et al. (in prep, Appendix A) found a low Pearson's correlation between species originality computed on trait and phylogeny data.

In a study case, Maire et al. (2015) found that the use of a tree-based approach could artificially increase species trait-based distances and that the use of multidimensional space with axes derived from principal component analysis (PCoA), seems to better represent species trait-based distances. Species originality, calculated in the multidimensional space as a distance from focal species to the centroid of all species, would avoid tree-based approach problems; but, if there are too many traits, the visualization of functional space can somehow become tricky and difficult to interpret (Saito et al., 2016). It is still not clear, however, how big is the difference between tree-based and ordination approaches for traits-based distances representation (Weinstein et al., 2014; Cianciaruso et al., 2017).

As species trait-based originality is drastically dependent on the accuracy of clustering algorithm (Podani and Schimera, 2007) and the quality of multidimensional space (see Maire et al., 2015), yet it could be difficult to choose between the two approaches. An alternative is to apply

originality measures directly to a distance matrix (Ricotta and Szeidl, 2009) as we did in urban plant originality in Chapter 2. Using the distance matrices for trait-based species originality measuring avoids us the arbitrariness of choice of clustering and ordination methods and simplifies the interpretation of multiple trait-based differences between species. Such an approach also makes it easier to compare trait-based originality and phylogenetic originality as phylogenetic distances can easily be extracted from a phylogenetic tree (Figure 2.a).



**Figure 2.** Illustration of interchangeable use of tree-based and distance-based approaches of species biological data representation for originality measuring: **a.** methods of pairwise phylogenetic distances calculation from phylogenetic trees, **b.** methods of hierarchical clustering from trait distance matrix to trait-based dendrogram. Methods presented in a. and b. can also be reversed. Sp represents five different species in a set.

## 2. THE MATTER OF SCALES FOR ORIGINALITY MEASURES

Levin (1992) noted that “the problem of relating phenomena across scales is the central problem in biology and in all of science.” Hence, the problem of scale at which ecological processes should be considered is critical to generate general predictions in ecology (Chave, 2013). As species originality is relative to a reference set of species, its value depends on the scale at which this species set is defined. Most studies on species originality have concerned large clades, like birds or

mammals, and have been done at the global, world scale (Safi et al., 2013; Jetz et al., 2014; Thullier et al., 2015). Such studies allow responding the questions like: which are the most original species on Earth?; Are original species threatened by extinction? In Chapter 2, however, we analyzed originality from local to regional scales and not on global scale. When originality is measured in a local site, it depends on the species composition of this local site, which itself strongly influenced by the regional source pool of species.

## **2.1 Importance of the spatial scale**

As written above, the global world scale has been the most studied scale for calculating species originality (see e.g., Veron et al. 2017 and references therein). This is because most studies focused on conservation biology asking which species are globally original and how many globally original and endangered species are supported by a given region (Safi et al., 2013; Jetz et al., 2014). These studies focus on patterns of the distribution of global originality and use the concept of originality as a criterion for the conservation of biodiversity. However, originality concept could help ecological studies to go beyond describing biodiversity patterns to determination of key processes underlying these patterns.

Ecological and evolutionary processes that act at different scales of biodiversity organization are interdependent. Thus, species ecological properties shape processes of species' evolution that in return influence the distribution and abundance of species in communities (Ackerly et al., 2003; Cavender-Bares et al., 2004). Local patterns of species originalities are determined by species ability to disperse, install and survive in local communities constrained by local community assembly processes (Chase and Myers, 2011). At the same time, local originality patterns are also constrained by macro-evolutionary processes (Ricklefs, 1987). Moreover, species' local originality is likely to vary across different specific compositions of communities, constrained by environmental gradients, and different evolutionary histories of coexisting species.

As a result, in Chapter 2, we chose to explore the concept of originality at two restricted scales: local and regional. It is indeed possible to calculate the originality of each species in reference

to the set of species present in a local area. If a species is present in several local areas, then its local originality will change depending on the local area considered (see discussion of Chapter 2). In addition, it is also possible to calculate the originality of each species in reference to all species present in a region and to make the mean of these regional originality values ( $MrO_p$ ) in each local community as we did for urban plant species. This measure ( $MrO_p$ ) represents the amount of the regional diversity expressed by species in the local community. If the regional diversity is well represented at the local level, then a local conservation effort could be enough to protect a local hotspot of regional biodiversity (Margules and Pressey, 2000; Veron et al., 2016). Multi-level originality approaches from local to global scales could allow determining the scale at which environmental factors have the strongest effect and where conservation should be prior. Such a multi-scale approach however questions on the definitions of regional and local species pools.

The regional pool of species was subject to controversial definitions since nearly 80 years (Graves and Rahbek, 2005). Usually, the composition of observed regional pool is limited to a list of species known to occur in a given biogeographic area, yet often without the reference to habitat or distance from the local assemblage (Graves and Rahbek, 2005). With such a definition, some species would appear highly original within the regional pool. In urban originality study (Chapter 2) for example, some of non-native species introduced by humans from other biogeographical zones were top ranked by trait-based originality measures. In another study, introduced Carnivore species were characterized by high phylogenetic and trait-based originality scores in Europe (Pavoine et al., 2017). Therefore, a regional pool should represent a reference set of species susceptible to colonize and compete for resource use in local communities (Lortie et al. 2004; Lessard et al, 2012; Cornell and Harrison 2014). The regional pool composition thus reflects distinct assembling processes acting hierarchically at local and regional scales and greatly influences local communities' formation (Zobel, 1997; Lessard et al., 2012). Hence, depending on how regional pool and local community are defined in terms of the taxa included, originality measures would give different lists of the most original species and hence, will reveal signatures of different



ecological processes and indicate different priority areas for conservation.

## **2.2 Taxonomical scale**

The decision about which taxonomic group to include into a study is based on a number of theoretical and practical considerations. Usually community ecology studies are interested in a finite set of traits and consider species from a delimited taxonomical group that occur in a given place. More fundamental evolutionary studies are searching for global patterns of species diversity and usually work with entire clades of species. Use of an inappropriate taxonomic scale can alter patterns detected during analysis (Cavender-Bares et al., 2006). Decisions on which species are included into the regional pool can fundamentally change the interpretation of species originality patterns. For example in Chapter 2 we decided to exclude species from Bryophytes (mosses) and Pteridophytes (ferns) from plant communities, as they might have very different trait values and evolutionary history compared to Angiosperm species retained. However, this choice was arbitrary and not easily justified, which is often the case (Cadotte et al., 2017). If there is no indications on which taxa to include, one could start with the largest possible species pool and test whether patterns of species originality distribution change within smaller species pools.

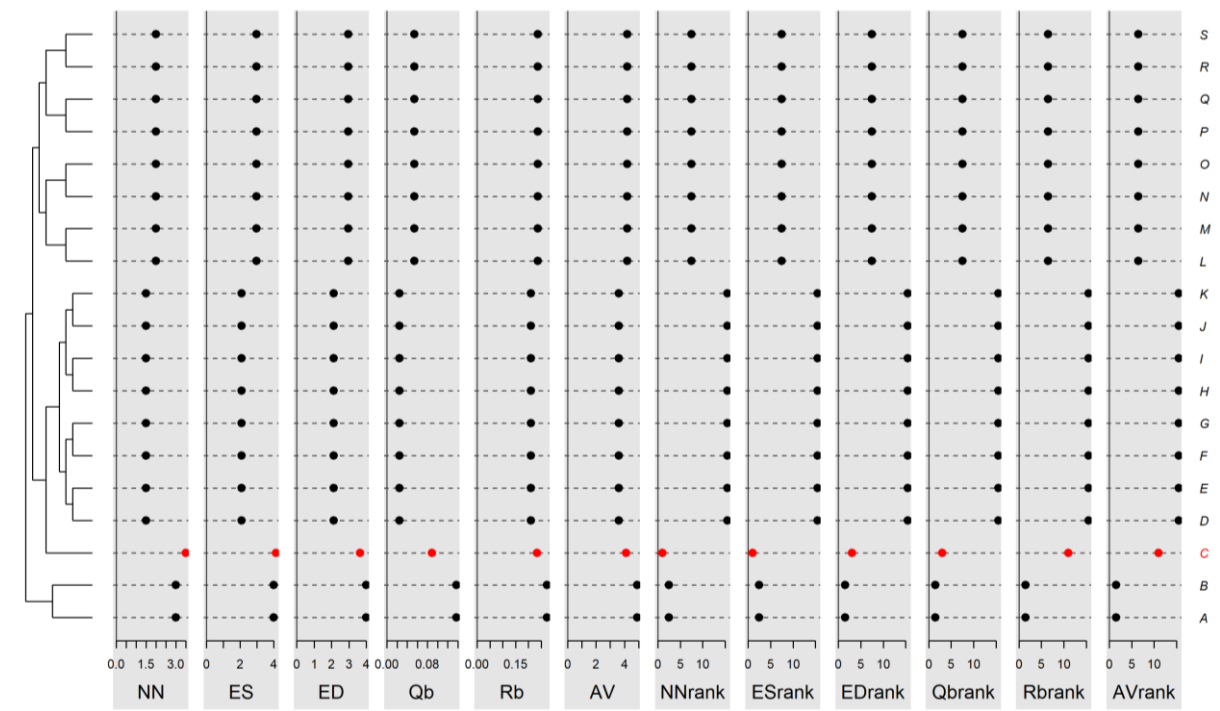
## **3. CHOOSING THE MATHEMATICAL FORMULA TO MEASURE ORIGINALITY**

A significant part of the efforts made by evolutionary and ecology scientists is guided by a common aim to preserve biodiversity and the benefits it provides to humanity. This ambitious intention demands not only a deep knowledge and understanding of what we want to protect but also an adequate tool helping humanity to make the decision. As seen above, the difficult task of selecting among different facets of biodiversity (species, sets of species as communities) as targets for prioritization depends upon the choice of data type and that of the study scale. It also depends on the measure selected. In this thesis we wanted to emphasize species originality measures that come to complete well-known species diversity measures. Species originality can be seen as a rarity of species characteristics (Pavoine et al., 2005, Chapter 1), completing species abundance-based rarity

with information on species' biological differences. Originality indices are numerous, some of them are redundant, some are using different types of data and differ in the way they weight species. These differences, sometimes subtle, can still substantially influence the patterns of originality scores observed. In Chapter 1, we already presented the relationship between species originality and diversity measures. Here we overview some methodological aspects that one should be aware of before choosing the right originality index.

### **3.1 Relations between species originality measures**

The idea that certain species contribute disproportionately to phylodiversity and trait-based diversity of species set is based on the assumption that those species possess an unusual set of traits unshared by the rest of species set (see Chapter 1). Studying species originality may reveal whether or not species with high originality values also maintain important ecosystem functions and large amounts of evolutionary history and are therefore worth to be protected in priority. In contrast to widely used diversity metrics, measured at the level of species set (local, regional, or global set), originality gives a value to each species in a set (Chapter 1). It is important to remember that originality is a relative measure, thus it can be described as a gradient (Chapter 1, Figure 2), with two extreme yet probably impossible cases: unique species on one side, not sharing any biological characteristic with others (100% of unique contribution to the diversity of a set) and redundant species on another side, sharing all its characters with other species (0% of unique contribution to the diversity of a set). Such a gradient encompasses all the possibilities of species differences in a set and its contribution to the diversity. An ideal originality measure would thus be able to position species along the whole gradient. Some originality measures however are biased along this gradient, being more sensitive to species uniqueness than to species redundancy (Figure 3).



**Figure 3.** Illustration of differences between originality indices in their sensitivity to uniqueness versus redundancy. Here six originality indices are considered. We considered a theoretical rooted phylogenetic tree with 19 species on tips (capital letters). The root is on the left. The distance from tip to root is equal to 5 units. We calculated for each species the NN index (nearest neighbor = pendant edge (PE)), the Equal-Split (ES) and Evolutionary-Distinctiveness (ED) indices (see Chapter 1), the indices derived from the quadratic entropy (Rb and Qb, see Chapter 1), and AV index (average distance to all other species). Then species were ranked from the highest (rank = 1) to the lowest (rank = 19) value (variables NNrank, ESrank, EDrank, Qbrank, Rbrank and AVrank). Species C is considered as the most original according to NN and ES. These indices overweight uniqueness (the branch that leads to a single species) in comparison to redundancy (shared branches) (NN even discards any redundancy). Species C is considered the third most original species by indices ED and Qb, which acknowledges the phylogenetic isolation of species A and B. In contrast, species C is considered the 11th most original species (so among the least original species) according to Rb and AV.

We have already presented a large variety of originality measures in Chapter 1. Redding et al. (2014) showed that some indices were always very closely related (see Table 1 in Redding et al., 2014). As described in the Appendix 1 of Chapter 1 (Table S1.1), there are basically two types of originality measures, nested inside of the species characters type used – phylogenetic or trait-based. We showed in Chapter 1 that many indices independently developed to measure either phylogenetic originality or trait-based originality are actually similar. This is the case for the pendant edge measure of phylogenetic originality (PE, Redding et al., 2014) and the nearest neighbor (NN) used to measure trait-based originality (Pavoine et al., 2017). The PE index does not include the root of phylogenetic tree in the calculation of phylogenetic originality. Therefore, PE reflects a recent dimension of species evolution and is sensible only to changes in the lengths of terminal branches on a phylogenetic tree (Redding et al., 2015a). PE overweighs species uniqueness because it fails to compare the amount of unique evolution with the amount of shared evolution between species, the latter being discarded. Same arguments are valuable for the NN measure. Hence, according to PE and NN indices, a species may appear more original than it actually is (Figure 3).

Among indices that were inferred from phylogenetic trees, the most used ones (Fair Proportion FP, Isaac et al., 2007, and Equal Splits ES, Redding et al., 2006) include the root of the phylogenetic tree in originality calculation and partition the total branch length between all species from the set (Figure 1 from Pavoine et al., 2017). In contrast to PE, they include all branches of the phylogenetic tree from tips to root while calculating species originality. Despite that, it has been shown that these measures are strongly correlated with PE and thus also overweigh species uniqueness over species redundancy (e.g., Jono and Pavoine, 2012).

In Chapter 2, we used an alternative index that averages dissimilarity between the target species and all others in a set. When applied to phylogenetic distances, this index (named AV for Average, Pavoine et al. 2017; or Average Pairwise Distance (APD), Redding et al., 2015a) represents more ancient dimensions of species evolution. Hence, it is sensible to deeper phylogenetic structure changes and errors in branch lengths (particularly the terminal branch) will provide only a limited

amount of incorrect information to a final originality value (Redding et al., 2015a). Redding et al. (2015a) also found that APD, working with phylogenetic information, was strongly correlated with trait-based measures of originality and thus could also capture well the ecosystem trait-based diversity. Thus, APD could represent an analytical link between phylogenetic and ecological conservation studies (Redding et al., 2015a). Pavoine et al. (2017) showed, however, that indices like AV and APD tend to give close values to all species in a set. Thus, if we were interested in distinguishing better between species redundancy and uniqueness of species, another related index named Rb would be more appropriate (Figure 3; Pavoine et al., 2017). This index, derived from Rao's quadratic entropy index (Rao, 1982), is strongly correlated to AV, but leads to higher variances in the values of species originality (Pavoine et al., 2017). Because of their various sensitivity to species uniqueness versus redundancy, the choice of an originality index is crucial to better represent and understand phylogenetic and trait-based structure of species community. For example, applying 11 different indices of phylogenetic originality to world's phylogenies of Mammals and Amphibians created different lists of the top 100 most original species, which also differed from the top species list created by EDGE program (Table 2 in Redding et al., 2014). However, the metrics that were closely related to the ED measure composing the EDGE formula (Isaac et al., 2007) produced very similar lists of top original species (Redding et al., 2014). Therefore, the more originality metrics are different in their calculation manner, the more varies species originality score.

### **3.2 On the ubiquity of originality measures in general**

We have seen in section 2.1 that continuous indices of species originality could be measured at any spatial scale from local to global. However, should the mathematical formula for originality change depending on the scale of a study? For example, in Violle et al. (2017a) functional rarity framework, the term of species "distinctiveness" was defined only at the local level and designated the mean distance to the other species in a set; in contrast, the term of species "uniqueness" was defined only at the regional level and designated the distance to the  $k$  nearest neighbors within a regional species

pool. Further research is however needed to determine whether the choice of originality indices should be dependent on the spatial scale of a study and more generally to determine which index is most appropriate to respond to which ecological questions (see Carmona et al., 2017; Violle et al., 2017b).

More generally, originality measures are flexible tools and can be applied in many different circumstances of various areas of ecological and conservational research. Originality measures require a set of entities that can be described by a common set of characteristics. In this thesis, we used species as an entity and brought attention to the most available and used species characters – traits and phylogenies. Using species as a unit constrains researchers to stick up to one of species' concept definitions. In this thesis, we used a mix of several definitions of species concept (see Introduction, section 3.1). Moreover, species could be replaced with individuals characterized by traits or DNA sequences (already indirectly used with molecular phylogenies) thus representing intra-specific originality. Species could also be replaced by other taxonomic levels (genera, families, clades; Cornwell et al., 2014) or trait functional groups assembling multiple species (Brandl et al., 2016). More broadly, the originality (and rarity) concepts could be extended to any scale of biological organization from genes to ecosystems (Carmona et al., 2017). As proposed by Pavoine et al. (2017):

*“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.”*

Within a biogeographic zone, some habitats or communities can have a unique composition, as well as a restricted area. Consideration of originality at any organizational scale of biodiversity would broaden the application of originality framework to study differences between habitats within landscapes or regions, regions within countries and biogeographic domains within the world (Carmona et al., 2017). More research is thus needed to determine which originality indices could allow us to integrate nested scales of originality by measuring the originality of an individual within a species, of a species within a site, of a site within a region, etc.

One of global goals of conservation biology, ecology and other disciplines interested in biological diversity is to *"address the underlying causes of biodiversity loss [...] to improve the status of biodiversity by safeguarding ecosystems, species and genetic diversity [...] and mainstreaming biodiversity across government and society"* (Aichi Biodiversity Targets, Strategic Plan for Biodiversity 2011–2020, CBD, 2011). Feasibility of projects considering biodiversity is constrained by available amount of knowledge and by budgetary capacities. At the same time, biodiversity prioritization is guided by nature's benefits to people, reflected in phylodiversity and trait-based diversity of species. Unfortunately, there is still a gap between the scales where biological diversity and ecosystem functioning are studied, scales where conservation actions are taken, and scales where direct benefits to humanity exist. To overpass this gap, an important role of biological research is to bridge theoretical and practical approaches, sometimes by making simplifications and by according theory a practical use. The main objective of this thesis was to evaluate the benefits that can bring the concept and measures of species' originality to different biological disciplines, particularly to conservation biology and community ecology. Originality indices allow treating species within multiple facets of biological diversity. Combining theoretical and empirical evaluations this thesis has contributed to the following key points:

- ∅ Common meaning of words used to designate a concept in Science can differ between researchers, which can be misleading for general understanding. Therefore, we first provided a semantic and historical overview of three commonly used concepts – diversity, rarity and originality – at the level of species assemblage. We propose that studies investigating species contribution to the diversity of a set use a unifying term of originality, so that the comparison between studies and key-words research could be more efficient.
- ∅ We made an explicit differentiation between species diversity, which is a property of a set of species and species originality, which is a property of individual species composing the



set. From there, we were able to describe that originality is a species contribution to the diversity of a reference set of species based on two main attributes used to describe species: traits and phylogeny. Therefore, the concept of species originality can be seen as species rarity based on other attributes than species abundance.

- Ø We then showed that mathematical links exist between associated indices of rarity, originality and diversity. Rarity, which is historically the oldest way to describe differences between species, was explicitly integrated into diversity measures by means of species abundance. Then, based on species' biological differences, diversity indices integrated species' phylogenetic and trait characters. We showed that phylodiversity and trait-based diversity can thus be written as a function of species rarity and originality measures in several ways (notably, as a mean, sum or a weighted mean).
- Ø Following theoretical outlines of originality-diversity concepts we proposed a practical application of a two-step analytical framework to real plant species data, by first measuring species-specific originality and then inferring species diversity with real plant species data. We demonstrated how regional and local species originality contributed to variations in community diversity along an urbanization gradient. The influence of urbanization gradient on community-level diversity differed between trait-based and phylo-based measures and between local and regional scales, which were the area over which originality was calculated (originality relative to the rest of the community or to the rest of the region).
- Ø Our originality framework indicated the most relevant spatial scale for studying trait-based and phylo-based originality. Moreover, we determined which species, original or redundant, are responsible for variations in trait-based diversity and phylodiversity. We notably showed that urbanization causes some species to have higher trait-based originality at the local community scale than at the regional scale. Therefore, variation of species originality across local communities is not random but dependent on regional species pool and community assemblage rules.

- Ø We brought attention to possible pitfalls that can be encountered during species originality analysis. Measured by various approaches based on trees (phylogenetic tree or trait-based dendrogram), dissimilarity matrices or multidimensional space, species originality depends on data completeness and accuracy. This issue has been encountered by many studies, however, its influence on species originality measures has not been addressed in literature. We showed with the very first results that originality dependency on trait missing data requires a deeper consideration and analysis in future studies.
- Ø We underlined that species originality can vary across spatial and taxonomic scales according to different ecological and evolutionary processes acting at these scales. In addition, we argued that including species abundance-based rarity into originality measuring can drastically change species scores and conclusions about ecosystem functioning, where a high number of individuals supporting the same trait can ensure greater stability. However, we showed that it is possible to develop originality measures that are regulating the relative importance given to rare vs. abundant species.
- Ø As it is not possible (and not meaningful) to include all species' attributes into one metric, we stressed the need for a joint use not only of different species' biological characters but also of different originality and diversity measures in order to understand mechanisms shaping species communities at different spatial and temporal scales. Different aspects of species characters would reflect different patterns of species evolution and implications into ecosystem functioning.
- Ø Overall, although these methodological considerations will have to be treated in future research, this thesis has demonstrated that the originality concept and its measures could be useful to community ecology, notably to identify the evolutionary events and ecological processes leading to the coexistence of species with different levels of originality. As these processes can include human activity, the originality concept and its measure could be also useful to conservation studies, especially if original species are threatened.

Ø Finally, we underlined the ubiquity of originality measures as they could give a value to any entity in focus compared to other entities of the same type. The focal entity can be a species compared to other species in a set as treated in this thesis. In future research a focal entity could also be a community of species compared to other communities in different regions or habitats, characterized by species composition. Few conservation actions have already shifted from conserving rare threatened species to conserving threatened species with original biological characteristics. As a step further, future actions could shift from conserving more diverse and threatened areas (or hotspots of biodiversity) to conserving areas composed of more original species.

## REFERENCES OF THE THESIS

---

- ACKERLY, D.D. (2003) Community Assembly , Niche Conservatism , and Adaptive Evolution in Changing Environments. *International Journal of Plant Sciences* 164, 164–184.
- ADLER, P.B., SALGUERO-GOMEZ, R., COMPAGNONI, A., HSU, J.S., RAY-MUKHERJEE, J., MBEAU-ACHE, C. & FRANCO, M. (2014) Functional traits explain variation in plant life history strategies. *PNAS* 111, 740–745.
- AGAPOW, P.M., BININDA-EMONDS, O.R.P., CRANDALL, K.A., GITTLEMAN, J.L., MACE, G.M., MARSHALL, J.C. & PURVIS, A. (2004) The impact of species concept on biodiversity studies. *Quarterly Review of Biology* 79, 161–179.
- ALBERT, C., LUQUE, G.M. & COURCHAMP, F. (2018) The twenty most charismatic species. *PLoS ONE* 13. Public Library of Science.
- ALBERT, C.H., DE BELLO, F., BOULANGEAT, I., PELLET, G., LAVOREL, S. & THUILLER, W. (2012) On the importance of intraspecific variability for the quantification of functional diversity. *Oikos* 121, 116–126.
- ANDERSSON, L. (1990) The driving force: species concepts and ecology. *Taxon* 39, 375–382.
- ANDRE, J.-C., MEGIE, G. & SCHMIDT-LAINE, C. (2003), Échelles et changements d'échelles, problématiques et outils, in: *Études sur l'environnement : de l'échelle du territoire à celle du continent*, P. Caseau (ed.) Rapport sur la science et la technologie, Académie des sciences, Tech&Doc, Lavoisier, Paris, pp. 167-199.
- BUCKLAND, S.T., MAGURRAN, A.E., GREEN, R.E. & FEWSTER, R.M. (2005) Monitoring change in biodiversity through composite indices *Monitoring change in biodiversity through composite indices*, 243–254.
- BAGOUSSE-PINGUET LE, Y., SOLIVERES, S., GROSS, N., TORICES, R. & BERDUGO, M. (2019) Phylogenetic, functional, and taxonomic richness have both positive and negative effects on ecosystem multifunctionality. *PNAS* 116, 8419–8424.
- BALMFORD, A. & BOND, W. (2005) Trends in the state of nature and their implications for human well-being. *Ecology Letters* 8, 1218–1234.
- BALMFORD, A., BRUNER, A., COOPER, P., COSTANZA, R., FARBER, S., GREEN, R.E., JENKINS, M., JEFFERISS, P., JESSAMY, V., MADDEN, J., MUNRO, K., MYERS, N., NAEEM, S., PAAVOLA, J., RAYMENT, M., ET AL. (2002) Economic Reasons for Conserving Wild Nature 297, 950–954.
- BARKER, G.M. (2002) Phylogenetic diversity: A quantitative framework for measurement of priority and achievement in biodiversity conservation. *Biological Journal of the Linnean Society* 76, 165–194.
- BEAULIEU, J.M., JHWUENG, D.C., BOETTIGER, C. & O'MEARA, B.C. (2012) Modeling stabilizing selection: Expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* 66, 2369–2383.

- BELLWOOD, D.R., STREIT, R.P., BRANDL, S.J. & TEBBETT, S.B. (2019) The meaning of the term ‘function’ in ecology: A coral reef perspective. *Functional Ecology* 33, 948–961.
- BEST, R.J., CAULK, N.C. & STACHOWICZ, J.J. (2013) Trait vs. phylogenetic diversity as predictors of competition and community composition in herbivorous marine amphipods. *Ecology Letters* 16, 72–80.
- BILLAUD, O., MOEN, D.S., PARSONS, T.L. & MORLON, H. (2019) Estimating Diversity Through Time using Molecular Phylogenies: Old and Species-Poor Frog Families are the Remnants of a Diverse Past. *Systematic Biology*.
- 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.
- BONNEY, R., COOPER, C.B., DICKINSON, J., KELLING, S., PHILLIPS, T., ROSENBERG, K. V. & SHIRK, J. (2009) Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience* 59, 977–984.
- BOON, E. (2019) Considering Intra-individual Genetic Heterogeneity to Understand Biodiversity. In: Casetta E., Marques da Silva J., Vecchi D. (eds.) *From Assessing to Conserving Biodiversity. History, Philosophy and Theory of the Life Sciences*, 24. Springer, Cham
- BRANDL, S.J., EMSLIE, M.J. & CECCARELLI, D.M. (2016) Habitat degradation increases functional originality in highly diverse coral reef fish assemblages. *Ecosphere*.
- BRÄUER, I. (2003) Money as an indicator : to make use of economic evaluation for biodiversity conservation 98, 483–491.
- BRONDIZIO ET AL. (2019, in press) IPBES Global Assessment on Biodiversity and Ecosystem Services, Chapter 1. Assessing a planet in transformation: Rationale and approach of the IPBES Global Assessment on Biodiversity and Ecosystem Services
- BROWN, J.H. (1984) On the relationship between abundance and distribution of species. *American Naturalist* 124, 255–279.
- BRUM, F.T., GRAHAM, C.H., COSTA, G.C., HEDGES, S.B., PENONE, C., RADELOFF, V.C., RONDININI, C., LOYOLA, R. & DAVIDSON, A.D. (2017) Global priorities for conservation across multiple dimensions of mammalian diversity. *Proceedings of the National Academy of Sciences* 114, 7641–7646.
- CADOTTE, M.W., DAVIES, T.J. & PERES-NETO, P.R. (2017) Why phylogenies do not always predict ecological differences. *Ecological Monographs* 87, 535–551.
- CADOTTE, M. W. & DAVIES, T. J. (2016) *Phylogenies in Ecology: A Guide to Concepts and Methods*. Princeton University Press, Princeton.
- CADOTTE, M., ALBERT, C.H. & WALKER, S.C. (2013) The ecology of differences: Assessing community assembly with trait and evolutionary distances. *Ecology Letters* 16, 1234–1244.
- CADOTTE, M.W., CARSCADDEN, K. & MIROTCHEV, N. (2011) Beyond species: Functional diversity and the maintenance of ecological processes and services. *Journal of Applied*

Ecology 48, 1079–1087.

CADOTTE, M.W., JONATHAN DAVIES, 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. *Ecology Letters* 13, 96–105.

CARDINALE, B.J., PALMER, M.A. & COLLINS, S.L. (2002) Species diversity enhances ecosystem functioning through interspecific facilitation. *Nature* 415, 426–429.

CARMONA, C.P., DE BELLO, F., SASAKI, T., UCHIDA, K. & PÄRTEL, M. (2017) Towards a Common Toolbox for Rarity: A Response to Violle et al. *Trends in Ecology and Evolution* xx, 11–12. Elsevier Ltd.

CASSETTA, E., MARQUES DA SILVA, J. & VECCHI, D. (2019) From Assessing to Conserving Biodiversity. Conceptual and Practical Challenges. Springer International Publishing.

CAVENDER-BARES, J., KOZAK, K.H., FINE, P.V.A. & KEMBEL, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters* 12, 693–715.

CAVENDER-BARES, J., KEEN, A. & MILES, B. (2006) Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* 87, 109–122.

CAVENDER-BARES, J., ACKERLY, D.D., BAUM, D.A. & BAZZAZ, F.A. (2004) Phylogenetic overdispersion in Floridian oak communities. *American Naturalist* 163, 823–843.

CHAO, A., CHIU, C. & JOST, L. (2014) Unifying Species Diversity, Phylogenetic Diversity, Functional Diversity, and Related Similarity and Differentiation Measures through Hill Numbers. *Annual Review of Ecology, Evolution, and Systematics* 45, 297–324.

CHASE, J.M. & MYERS, J.A. (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366, 2351–2363.

CHAVE, J. (2013) The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecology Letters* 16, 4–16.

CHIU, C.H., JOST, L. & CHAO, A. (2014) Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs* 84, 21–44.

CIANCIARUSO, M. V., SOBRAL, F.L. & LEES, A.C. (2017) On dendrograms, ordinations and functional spaces: Methodological choices or pitfalls? *Frontiers of Biogeography* 9, 1–4.

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. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366, 2611–2622.

CORNELL, H. V. & HARRISON, S.P. (2014) What Are Species Pools and When Are They Important? *Annual Review of Ecology, Evolution, and Systematics* 45, 45–67.

CORNWELL, W.K., WESTOBY, M., FALSTER, D.S., FITZJOHN, R.G., O'MEARA, B.C.,

PENNELL, M.W., MCGLINN, D.J., EASTMAN, J.M., MOLES, A.T., REICH, P.B., TANK, D.C., WRIGHT, I.J., AARSSSEN, L., BEAULIEU, J.M., KOOYMAN, R.M., ET AL. (2014) Functional distinctiveness of major plant lineages. *Journal of Ecology* 102, 345–356.

CORNWELL, W.K., SCHWILK, D.W. & ACKERLY, D.D. (2006) a Trait-Based Test for Habitat Filtering: Convex Hull Volume. *Ecology* 87, 1465–1471.

COSTANZA, R., J, R.A., GROOT, R. DE, FARBERLL, S., GRASSOT, M., HANNON, B., LIMBURG, K., NAEEM, S., NEILLTT, R.V.O., J, J.P.J., RASKIN, R.G., SUTTONLLLL, P. & BELT, M. VAN DEN (1997) The value of the world ' s ecosystem services and natural capital. *Nature* 387, 253–260.

CRUTZEN, P.J. (2006) Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma? *Climatic Change* 77, 211–219.

DARWIN, C. (1859) On the origin of species by means of natural selection: or, the preservation of favoured races in the struggle for life.

DAVIES, T.J., FRITZ, S.A., GRENYER, R., ORME, C.D.L., BIELBY, J., BININDA-EMONDS, O.R.P., CARDILLO, M., JONES, K.E., GITTLEMAN, J.L., MACE, G.M. & PURVIS, A. (2008) Phylogenetic trees and the future of mammalian biodiversity. In *the Light of Evolution* 2, 263–280.

DELONG, D.C.J. (1996) Defining biodiversity. *Wildlife Society Bulletin (1973-2006)* 24, 738–749.

DENELLE, P., VIOLLE, C. & MUNOZ, F. (2019) Distinguishing the signatures of local environmental filtering and regional trait range limits in the study of trait–environment relationships. *Oikos* 128, 960–971.

DIAZ, B.S., PASCUAL, U., STENSEKE, M., MARTIN-LOPEZ, B., WATSON, R.T., MOLNAR, Z., HILL, R., CHAN, K.M.A., BASTE, I.A., BRAUMAN, K.A., POLASKY, S., CHURCH, A., LONSDALE, M., LARIGAUDERIE, A., LEADLEY, P.W., ET AL. (2018) Assessing nature's contributions to people. *Science* 359, 270–272.

DIAZ, S., KATTGE, J., CORNELISSEN, J.H.C., WRIGHT, I.J., LAVOREL, S., DRAY, S., REU, B., KLEYER, M., WIRTH, C., COLIN PRENTICE, I., GARNIER, E., BÖNISCH, G., WESTOBY, M., POORTER, H., REICH, P.B., ET AL. (2016) The global spectrum of plant form and function. *Nature* 529, 167–171.

DIAZ, S., DEMISSEW, S., CARABIAS, J., JOLY, C., LONSDALE, M., ASH, N., LARIGAUDERIE, A., ADHIKARI, J.R., ARICO, S., BALDI, A., BARTUSKA, A., BASTE, I.A., BILGIN, A., BRONDIZIO, E., CHAN, K.M.A., ET AL. (2015) The IPBES Conceptual Framework - connecting nature and people. *Current Opinion in Environmental Sustainability* 14, 1–16.

DIAZ, S., GRIGULIS, K., ROBSON, T.M., LAVOREL, S., BELLO, F. DE, QUE, F., DIAZ, S., LAVOREL, S., DE BELLO, F., QUETIER, F., GRIGULIS, K. & ROBSON, T.M. (2007b) Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences of the United States of America* 104, 20684–20689.

DIAZ, S., FARGIONE, J., CHAPIN, F.S. & TILMAN, D. (2006) Biodiversity loss threatens

human well-being. *PLoS Biology* 4, 1300–1305.

DIAZ, S. & CABIDO, M. (2001) Vive la difference: Plant functional diversity matters to ecosystem processes. *Trends in Ecology and Evolution* 16, 646–655.

DIAZ ET AL. (in press) Summary for policymakers of the global assessment report on biodiversity and ecosystem services. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.

DESCOLA, P. (2015) *Par-delà nature et culture*. Paris, Gallimard, coll. Folio essais.

DICKINSON, J.L., ZUCKERBERG, B. & BONTER, D.N. (2010) Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics* 41, 149–172.

DINIZ-FILHO, J.A.F., LOYOLA, R.D., RAIÁ, P., MOOERS, A.O. & BINI, L.M. (2013) Darwinian shortfalls in biodiversity conservation. *Trends in Ecology and Evolution* 28, 689–695.

DURANT, S.M. (2014) Building Sustainable National Monitoring Networks. In: Collen, B., Pettorelli, N., Baillie, J.E.M., Durant, S.M (eds.). *Biodiversity Monitoring and Conservation: Bridging the Gap between Global Commitment and Local Action*. Wiley-Blackwell

DWYER, J.M., HOBBS, R.J. & MAYFIELD, M.M. (2014) Specific leaf area responses to environmental gradients through space and time 95, 399–410.

ELANDS, B.H.M., WIERSUM, K.F., BUIJS, A.E. & VIERIKKO, K. (2015) Policy interpretations and manifestation of biocultural diversity in urbanized Europe: conservation of lived biodiversity. *Biodiversity and Conservation* 24, 3347–3366. Springer Netherlands.

ELLENBERG, H., WEBER, H.E., DÜLL, R., WIRTH, V., WERNER, W., PAULISSEN, D. (1991). *Zeigerwerte von Pflanzen in Mitteleuropa*. *Scr. Geobot.* 188, 1–248.

EMERSON, B.C. & GILLESPIE, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution* 23, 619–630.

ESCUEDRO, A., IRIONDO, J.M. & TORRES, M.E. (2003) Spatial analysis of genetic diversity as a tool for plant conservation. *Biological Conservation* 113, 351–365.

FAITH, D.P., MAGALLON, S., HENDRY, A.P. & DONOGHUE, M.J. (2017) Future Benefits from Contemporary Ecosystem Services: A Response to Rudman et al. *Trends in Ecology and Evolution* 32, 717–719. Elsevier Ltd.

FAITH, D.P. (2008) Threatened species and the potential loss of phylogenetic diversity: Conservation scenarios based on estimated extinction probabilities and phylogenetic risk analysis. *Conservation Biology* 22, 1461–1470.

FAITH, D.P. & WALKER, P.A. (1996b) Environmental diversity: On the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiversity and Conservation* 5, 399–415.

FAITH, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*



61, 1–10.

FELSENSTEIN, J. (1985) Phylogenies and the comparative method. *American Naturalist*, 125, 1–15.

FLYNN, D.F.B., MIROTCHEV, N., JAIN, M., PALMER, M.I. & NAEEM, S. (2011) Functional and phylogenetic diversity as predictors of biodiversity–ecosystem–function relationships. *Ecology* 92, 1573–1581.

GARNIER, E., STAHL, U., LAPORTE, M. A., KATTGE, J., MOUGENOT, I., KÜHN, I., et al. (2017). Towards a thesaurus of plant characteristics: an ecological contribution. *J. Ecol.* 105, 298–309. doi:10.1111/1365-2745.12698

GARNIER, E., NAVAS, M.-L. & GRIGULIS, K. (2016) *Plant Functional Diversity*. Oxford University Press.

GARNIER, E., CORTEZ, J., BILLES, G., NAVAS, M.-L., ROUMET, C., DEBUSSCHE, M., LAURENT, G., BLANCHARD, A., AUBRY, D., BELLMAN, A., NEILL, C. & TOUSSAINT, J.-P. (2004) Plant Functional Markers Capture Ecosystem Properties. *Ecology* 85, 2630–2637.

GASTON, K.J. (1994) Progress reports Biodiversity - measurement. *Progress in Physical Geography* 18, 565–574.

GASTON, K. J. (1997). What is rarity? In Kunin, W. E. and Gaston, K. J. (eds), *The Biology of Rarity* pp. 30–47. Chapman & Hall, London.

GASTON, K.J., SPICER, J.I. (2004) *Biodiversity: An Introduction*. Oxford: Blackwell Science Ltd  
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). *Functional Ecology*, 29, 600–614.

GINI, C. (1912). Variabilità e Mutabilità. *Studi economica giuridici delle facoltà di giurisprudenza dell'Università di Cagliari III, Parte II*.

GONZÁLEZ-SUÁREZ, M. & E., R. (2013) Variability in life-history and ecological traits is a buffer against extinction in mammals. *Ecology Letters* 16, 242–251.

GOWER, J.C. (1971) A GENERAL COEFFICIENT OF SIMILARITY AND SOME OF ITS PROPERTIES. *Biometrics* 27, 837–842.

GRAVES, G.R. & RAHBEK, C. (2005) Source pool geometry and the assembly of continental avifaunas. *Proceedings of the National Academy of Sciences of the United States of America* 102, 7871–7876.

GREEN, J. & BOHANNAN, B.J.M. (2006) Spatial scaling of microbial biodiversity. *Trends in Ecology and Evolution* 21, 501–507.

GRENÍÉ, M., MOUILLOT, D., VILLÉGER, S., DENELLE, P., TUCKER, C.M., MUNOZ, F. & VIOLLE, C. (2018) Functional rarity of coral reef fishes at the global scale: Hotspots and challenges for conservation. *Biological Conservation* 226, 288–299.

- GRIBALDO, S. & PHILIPPE, H. (2002) Ancient Phylogenetic Relationships. *Theoretical Population Biology* 61, 391–408.
- GRIME, J.P. (1998) Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *Journal of Ecology* 86, 902–910.
- GROSS, N., LIANCOURT, P., BUTTERS, R., DUNCAN, R.P. & HULME, P.E. (2015) Functional equivalence, competitive hierarchy and facilitation determine species coexistence in highly invaded grasslands. *New Phytologist* 206, 175–186.
- HAILA, Y. & KOUKI, J. (1994) The phenomenon of biodiversity in conservation biology. *Finnish Zoological Publishing Board* 31, 5–18.
- HAMPTON, S.E., STRASSER, C.A., TEWKSBURY, J.J., GRAM, W.K., BUDDEN, A.E., BATCHELLER, A.L., DUKE, C.S. & PORTER, J.H. (2013) Big data and the future of ecology. *Frontiers in Ecology and the Environment* 11, 156–162.
- HANSEN, T.F. (1997) Stabilizing Selection and the Comparative Analysis of Adaptation. *Evolution* 51, 1341–1351.
- HARVEY, P. H. & PAGEL, M. D. (1991) *The Comparative Method in Evolutionary Biology* Oxford: Oxford University Press, pp. 239
- HAY, J.M., SARRE, S.D., LAMBERT, D.M., ALLENDORF, F.W. & DAUGHERTY, C.H. (2010) Genetic diversity and taxonomy: A reassessment of species designation in tuatara (*Sphenodon*: Reptilia). *Conservation Genetics* 11, 1063–1081.
- HERBEN, T. & GOLDBERG, D.E. (2014) Community assembly by limiting similarity vs. competitive hierarchies: Testing the consequences of dispersion of individual traits. *Journal of Ecology* 102, 156–166.
- HIDASI-NETO, J., LOYOLA, R. & CIANCIARUSO, M. V. (2015) Global and local evolutionary and ecological distinctiveness of terrestrial mammals: Identifying priorities across scales. *Diversity and Distributions* 21, 548–559.
- HIGGINS, D., BERKLEY, C. & JONES, M.B. (2002) Managing heterogeneous ecological data using Morpho. In *Proceedings of the 14th International Conference on Scientific and Statistical Database Management, SSDBM* pp. 69–76.
- HOLLING, C.S. (1973) Resilience and stability of ecological systems. *Annu.Rev.Ecol.Syst.* 4, 1–23.
- HOOPER, D.U., CHAPIN, 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. *Ecological Monographs* 75, 3–35. Ecological Society of America.
- HUMPHRIES, C., WILLIAMS, P.H. & VANE-WRIGHT, R.I. (1995) MEASURING BIODIVERSITY VALUE FOR CONSERVATION. *Annu. Rev. Ecol. Syst.*, 93–111.

IPBES (2019) Global Assessment on Biodiversity and Ecosystem Services. Chapter 1. Assessing a planet in transformation: Rationale and approach of the IPBES Global Assessment on Biodiversity and Ecosystem Services.

ISAAC, N.J.B., REDDING, D.W., MEREDITH, H.M. & SAFI, K. (2012) Phylogenetically-Informed Priorities for Amphibian Conservation. *PLoS ONE* 7, 1–8.

ISAAC, N.J.B., TURVEY, S.T., COLLEN, B., WATERMAN, C. & BAILLIE, J.E.M. (2007) Mammals on the EDGE: Conservation priorities based on threat and phylogeny. *PLoS ONE* 2.  
IVES, A.R. & CARPENTER, S.R. (2007) Stability and Diversity of Ecosystems Concepts of Stability. *Science* 317, 58–63.

IVES, A.R. & HELMUS, M.R. (2010) Phylogenetic metrics of community similarity. *American Naturalist* 176.

JABOT, F. (2010) A stochastic dispersal-limited trait-based model of community dynamics. *Journal of Theoretical Biology* 262, 650–661. Elsevier.

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. *Current Biology*.

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., ET AL. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90, 2648–2648.

JONO, C.M.A.A. & PAVOINE, S. (2012) Threat Diversity Will Erode Mammalian Phylogenetic Diversity in the Near Future. *PLoS ONE* 7.

KATTGE, J., DIAZ, S., LAVOREL, S., PRENTICE, I.C., LEADLEY, P., BÖNISCH, G., GARNIER, E., WESTOBY, M., REICH, P.B., WRIGHT, I.J., CORNELISSEN, J.H.C., VIOLLE, C., HARRISON, S.P., VAN BODEGOM, P.M., REICHSTEIN, M., ET AL. (2011) TRY - a global database of plant traits. *Global Change Biology* 17, 2905–2935.

KNAPP, S., DINSMORE, L., FISSORE, C., HOBBIE, S.E., JAKOBSDOTTIR, I., KATTGE, J., KING, J.Y., KLOTZ, S., MCFADDEN, J.P. & CAVENDER-BARES, J. (2012) Phylogenetic and functional characteristics of household yard floras and their changes along an urbanization gradient. *Ecology* 93, 83–98.

KNAPP, S., KÜHN, I., SCHWEIGER, O. & KLOTZ, S. (2008) Challenging urban species diversity: Contrasting phylogenetic patterns across plant functional groups in Germany. *Ecology Letters* 11, 1054–1064.

KOSMAN, E., BURGIO, K.R., PRESLEY, S.J., WILLIG, M.R. & SCHEINER, S.M. (2019) Conservation prioritization based on trait-based metrics illustrated with global parrot distributions. *Diversity and Distributions* 25, 1156–1165.

KRAUSS, J., BOMMARCO, R., GUARDIOLA, M., HEIKKIEN, K., HELM, A., KUUSSAARI, M., LINDBORG, R., OCKINGER, E., PARTEL, M., PINO, J., POYRY, J., RAATIKAINEN,

- K.M., SANG, A., STEFANESCU, C., TEDER, T., ET AL. (2010) Habitat fragmentation causes immediate and time- delayed biodiversity loss at different trophic levels. *Ecology Letters* 13, 597–605.
- KÜHN, I., WOLF, J. & SCHNEIDER, A. (2017) Is there an urban effect in alien plant invasions? *Biological Invasions* 19, 3505–3513.
- KUHN, T.S., MOOERS, A. & THOMAS, G.H. (2011) A simple polytomy resolver for dated phylogenies. *Methods in Ecology and Evolution* 2, 427–436.
- KYLAFIS, G. & LOREAU, M. (2008) Ecological and evolutionary consequences of niche construction for its agent. *Ecology Letters* 11, 1072–1081.
- LADÉAU, S.L., HAN, B.A., ROSI-MARSHALL, E.J. & WEATHERS, K.C. (2017) The Next Decade of Big Data in Ecosystem Science. *Ecosystems* 20, 274–283. Springer US.
- LAURILA-PANT, M., LEHIKOINEN, A., UUSITALO, L. & VENESJÄRVI, R. (2015) How to value biodiversity in environmental management? *Ecological Indicators* 55, 1–11. Elsevier Ltd.
- LAVOREL, S., GRIGULIS, K., LAMARQUE, P., COLACE, M.P., GARDEN, D., GIREL, J., PELLET, G. & DOUZET, R. (2011) Using plant functional traits to understand the landscape distribution of multiple ecosystem services. *Journal of Ecology* 99, 135–147.
- LAVOREL, S., GRIGULIS, K., MCINTYRE, S., WILLIAMS, N.S.G., GARDEN, D., DORROUGH, J., BERMAN, S., QUETIER, F., THEBAULT, A. & BONIS, A. (2008) Assessing functional diversity in the field - Methodology matters! *Functional Ecology* 22, 134–147.
- LAVOREL, S. & GARNIER, E. (2002) Predicting Changes in Community Composition and Ecosystem Functioning from Plant Traits : Revisiting the Holy Grail, *Functional Ecology* 16, 545–556.
- LEAN, C. & MACLAURIN, J. (2016) The value of phylogenetic diversity. In Pellens, R. and Grandcolas, P. (eds), *Biodiversity Conservation and Phylogenetic Systematics* pp. 19–37. Springer International Publishing, Basel.
- LEHMAN, C.L. & TILMAN, D. (2000) Biodiversity , Stability , and Productivity in Competitive Communities. *The American Naturalist* 156, 534–552.
- LEITÃO, R.P., ZUANON, J., VILLEGGER, S., WILLIAMS, S.E., BARALOTO, C., FORTUNEL, C., MENDONÇA, F.P. & MOUILLOT, D. (2016) Rare species contribute disproportionately to the functional structure of species assemblages. *Proc. R. Soc. B* 283: 238.
- LESSARD, J.P., BORREGAARD, M.K., FORDYCE, J.A., RAHBEK, C., WEISER, M.D., DUNN, R.R. & SANDERS, N.J. (2012) Strong influence of regional species pools on continent-wide structuring of local communities. *Proceedings of the Royal Society B: Biological Sciences* 279, 266–274.
- LETTEN, A.D. & CORNWELL, W.K. (2015) Trees, branches and (square) roots: Why evolutionary relatedness is not linearly related to functional distance. *Methods in Ecology and Evolution* 6, 439–444.

- LEVIN, S.A. (2000) Multiple scales and the maintenance of biodiversity. *Ecosystems* 3, 498–506.
- LEVIN, S.A. (1992) The problem of pattern and scale in ecology 73, 1943–1967.
- LEWIS, P.O., HOLDER, M.T. & HOLSINGER, K.E. (2005) Polytomies and Bayesian phylogenetic inference. *Systematic Biology* 54, 241–253.
- LOREAU, M. (2010) Linking biodiversity and ecosystems: Towards a unifying ecological theory. Royal Society. *Phil. Trans. R. Soc. Lond. B*.
- LOREAU, M., NAEEM, S., INCHAUSTI, P., BENGTSSON, J., GRIME, J.P., HECTOR, A., HOOPER, D.U., HUSTON, M.A., RAFFAELLI, D., B.SCHMID, B., TILMAN, D. & WARDLE, D.A. (2001) Biodiversity and Ecosystem Functioning : Current Knowledge and Future Challenges. *Science* 294, 804–809.
- LORTIE, C.J., BROOKER, R.W., CHOLER, P., KIKVIDZE, Z., MICHALET, R., PUGNAIRE, F.I. & CALLAWAY, R.M. (2004) Rethinking plant community theory. *Oikos* 107, 433–438.
- LOSOS, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters* 11, 995–1003.
- LOVEJOY, T.E. (1980) The technical report. In: Barney, G. O., (ed.) *The Global 2000 Report to The President* (Vol. 2, pp. 327-332). Penguin.
- MACE, G.M. (2014) Whose conservation? *Science* 345, 1558–1560.
- MAGURRAN, A. E. (2004). *Measuring Biological Diversity*. Blackwell Publishing, Oxford, UK
- MAIRE, E., GRENOUILLET, G., BROSSE, S. & VILLEGGER, S. (2015) How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. *Global Ecology and Biogeography*.
- MÉJEKOVÁ, M., PAAL, T., PLOWMAN, N.S., BRYNDOVA, M., KASARI, L., NORBERG, A., WEISS, M., BISHOP, T.R., LUKE, S.H., SAM, K., LE BAGOUSSE-PINGUET, Y., LEPS, J., GÖTZENBERGER, L. & DE BELLO, F. (2016) Evaluating Functional diversity: Missing trait data and the importance of species abundance structure and data transformation. *PLoS ONE* 11, 1–17.
- MARGULES, C.R. & PRESSEY, R.L. (2000) Systematic conservation planning. *Nature* 405, 243–253.
- MARTIN, G., DEVICTOR, V., MOTARD, E., MACHON, N. & PORCHER, E. (2019) Short-term climate-induced change in French plant communities, 3–7.
- MASON, N.W.H., DE BELLO, F., MOUILLOT, D., PAVOINE, S. & DRAY, S. (2013) A guide for using functional diversity indices to reveal changes in assembly processes along ecological gradients. *Journal of Vegetation Science* 24, 794–806.
- MAY, R.M. (1990) Taxonomy as destiny. *Nature*, 1–4.

- MAYDEN, R.L. (1999) Consilience and a hierarchy of species concepts: Advances toward closure on the species puzzle. *Journal of Nematology* 31, 95–116.
- MAYER, P. (2006) Biodiversity — The Appreciation of Different Thought Styles and Values Helps to Clarify the Term. *Restoration Ecology* 14, 105–111.
- MAYR, E. (1942) *Systematics and the Origin of Species*. Columbia Univ. Press, New York.
- MAYR, E. (1963) *Animal Species and Evolution*. Harvard Univ. Press, Cambridge, MA
- MAZEL, F., PENNELL, M.W., CADOTTE, M.W., DIAZ, S., RIVA, G.V.D., GRENYER, R., LEPRIEUR, F., MOOERS, A.O., MOUILLOT, D., TUCKER, C.M. & PEARSE, W.D. (2019) Reply to: “Global conservation of phylogenetic diversity captures more than just functional diversity”. *Nature Communications* 10, 9–10. Springer US.
- MAZEL, F., PENNELL, M.W., CADOTTE, M.W., DIAZ, S., DALLA RIVA, G.V., GRENYER, R., LEPRIEUR, F., MOOERS, A.O., MOUILLOT, D., TUCKER, C.M. & PEARSE, W.D. (2018) Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nature Communications* 9, 1–9. Springer US.
- MAZEL, F., MOOERS, A., RIVA, G.V.D. & PENNELL, M.W. (2017) Conserving Phylogenetic Diversity Can Be a Poor Strategy for Conserving Functional Diversity 66, 1019–1027.
- MAZEL, F., GUILHAUMON, F., MOUQUET, N., DEVICTOR, V., GRAVEL, D., RENAUD, J., CIANCIARUSO, M.V., LOYOLA, R., DINIZ-FILHO, J.A.F., MOUILLOT, D. & THUILLER, W. (2014) Multifaceted diversity-area relationships reveal global hotspots of mammalian species, trait and lineage diversity. *Global Ecology and Biogeography* 23, 836–847.
- MCGILL, B.J., ENQUIST, B.J., WEIHER, E. & WESTOBY, M. (2006) Rebuilding community ecology from functional traits 21.
- MCKINLEY, D.C., MILLER-RUSHING, A.J., BALLARD, H.L., BONNEY, R., BROWN, H., COOK-PATTON, S.C., EVANS, D.M., FRENCH, R.A., PARRISH, J.K., PHILLIPS, T.B., RYAN, S.F., SHANLEY, L.A., SHIRK, J.L., STEPENUCK, K.F., WELTZIN, J.F., ET AL. (2017) Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation* 208, 15–28. Elsevier Ltd.
- MERINO, J., CABALLERO, I., RIVAS, B., SERRANO, M. & PIATTINI, M. (2016) A Data Quality in Use model for Big Data. *Future Generation Computer Systems* 63, 123–130. Elsevier B.V.
- MEA (2005). *Ecosystems and Human Wellbeing: Current State and Trends: Findings of the Condition and Trends Working Group*. Washington: Island Press. <http://www.millenniumassessment.org>.
- MILNER-GULLAND, E. J., AND MACE, R. (1998). *Conservation of Biological Resources*. Blackwell Science, Oxford, United Kingdom.
- MITTELBACH, G.G., STEINER, C.F., SCHEINER, S., GROSS, K.L., REYNOLDS, H.L., WAIDE, R.B., WILLIG, M.R., DODSON, S.I. & GOUGH, L. (2001) CONCEPTS & SYNTHESIS. *Ecology* 82, 2381–2396.

- MITTERMEIER, R.A., MITTERMEIER, C.G., BROOKS, T.M., PILGRIM, J.D., KONSTANT, W.R., FONSECA, G.A.B. & KORMOS, C. (2003) Wilderness and biodiversity conservation. *PNAS* 100, 10309–10313.
- MITTERMEIER R.A., TURNER W.R., LARSEN F.W., BROOKS T.M., GASCON C. (2011) Global Biodiversity Conservation: The Critical Role of Hotspots. In: Zachos F., Habel J. (eds.) *Biodiversity Hotspots*. Springer, Berlin, Heidelberg
- MLAMBO, M.C. (2014) Not all traits are ‘functional’: Insights from taxonomy and biodiversity-ecosystem functioning research. *Biodiversity and Conservation* 23, 781–790.
- MORLON, H., POTTS, M.D. & PLOTKIN, J.B. (2010) Inferring the dynamics of diversification: A coalescent approach. *PLoS Biology* 8.
- MOUCHET, M., GUILHAUMON, F., VILLEGGER, S., MASON, N.W.H., TOMASINI, J.-A. & MOUILLOT, D. (2008) Towards a consensus for calculating dendrogram-based functional diversity indices. *Oikos* 117, 794–800.
- 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. (2013a) Rare Species Support Vulnerable Functions in High-Diversity Ecosystems. *PLoS Biology* 11.
- MOUILLOT, D., GRAHAM, N.A.J., VILLEGGER, S., MASON, N.W.H. & BELLWOOD, D.R. (2013b) A functional approach reveals community responses to disturbances. *Trends in Ecology and Evolution* 28, 167–177.
- MOUILLOT, D., CULIOLI, J.M., PELLETIER, D. & TOMASINI, J.A. (2008) Do we protect biological originality in protected areas? A new index and an application to the Bonifacio Strait Natural Reserve. *Biological Conservation* 141, 1569–1580.
- MOUQUET, N., DEVICOR, V., MEYNARD, C.N., MUNOZ, F., BERSIER, L.F., CHAVE, J., COUTERON, P., DALECKY, A., FONTAINE, C., GRAVEL, D., HARDY, O.J., JABOT, F., LAVERGNE, S., LEIBOLD, M., MOUILLOT, D., ET AL. (2012) Ecophylogenetics: Advances and perspectives. *Biological Reviews*.
- MÜNKEMÜLLER, T. & GALLIEN, L. (2015) VirtualCom: a simulation model for eco-evolutionary community assembly and invasion. *Methods in Ecology and Evolution* 6, 735–743.
- MÜNKEMÜLLER, T., LAVERGNE, S., BZEZNIK, B., DRAY, S., JOMBART, T., SCHIFFERS, K. & THUILLER, W. (2012) How to measure and test phylogenetic signal. *Methods in Ecology and Evolution* 3, 743–756.
- MUNOZ, F., GRENIÉ, M., DENELLE, P., TAUDIERE, A., LAROCHE, F., TUCKER, C. & VIOLLE, C. (2017) ecolottery: Simulating and assessing community assembly with environmental filtering and neutral dynamics in R. *Meth* 9, 693–703.
- MUSCARELLA, R. & URIARTE, M. (2016) Do community-weighted mean functional traits reflect optimal strategies? *Proceedings of the Royal Society B: Biological Sciences* 283, 20152434.
- MYERS, N., MITTERMEIER, R.A., MITTERMEIER, C.G., FONSECA, G.A.B. & KENT, J.

- (2000) Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858.
- NAEEM, S., KNOPS, J.M.H., TILMAN, D., HOWE, K.M., KENNEDY, T. & GALE, S. (2000) Plant diversity increases resistance to invasion in the absence of covarying extrinsic factors. *Oikos* 91, 97–108.
- NAKAGAWA, S. & FRECKLETON, R.P. (2008) Missing inaction: the dangers of ignoring missing data. *Trends in Ecology and Evolution* 23, 592–596.
- NAOMI, S.I. (2011) On the integrated frameworks of species concepts: Mayden’s hierarchy of species concepts and de Queiroz’s unified concept of species. *Journal of Zoological Systematics and Evolutionary Research* 49, 177–184.
- NEWMAN, E.A. (2019) Disturbance Ecology in the Anthropocene. *Frontiers in Ecology and Evolution* 7.
- NIXON, K. C. & WHEELER, Q.D. (1992). Measures of phylogenetic diversity. In Novacek, M. J. and Wheeler, Q. D. (eds), *Extinction and Phylogeny*, pp. 216–234. Columbia University Press, New York.
- NOSS, R.F. (1990) Indicators for Monitoring Biodiversity: A Hierarchical Approach. *Conservation Biology* 4, 355–364.
- NUNES, P.A.L.D. & BERGH, J.C.J.M. VAN DEN (2001) Economic valuation of biodiversity : sense or nonsense ? *Ecological Economics* 39, 203–222.
- Opportunities for Data Exchange (ODE) Report on Integration of data and publications, 2011
- OWEN, N., GUMBS, R., GRAY, C.L. & FAITH, D.P. (2019) Global conservation of phylogenetic diversity captures more than just functional diversity. *Nature Communications* 10, 1–3. Springer US.
- PACHEPSKY, E., CRAWFORD, J.W., BOWN, J.L. & SQUIRE, G. (2001) Towards a general theory of biodiversity for the Anthropocene. *Nature* 410, 923–926.
- PAGIOLA, S., VON RITTER, K., BISHOP, J. (2004) Assessing the Economic Value of Ecosystem Conservation. Environment Department Papers. Environmental Economics; no. 101. World Bank, Washington, DC. © World Bank. <https://openknowledge.worldbank.org/handle/10986/18391> License: CC BY 3.0 IGO
- PAKEMAN, R.J. (2014) Functional trait metrics are sensitive to the completeness of the species’ trait data ? *Methods in Ecology and Evolution* 5, 9–15.
- PAKEMAN, R.J., EASTWOOD, A. (2011) Leaf dry matter content as a predictor of grassland litter decomposition: a test of the ‘mass ratio hypothesis’. *Plant Soil* 342, 49–57.
- PALMA, E., CATFORD, J.A., CORLETT, R.T., DUNCAN, R.P., HAHS, A.K., MCCARTHY, M.A., MCDONNELL, M.J., THOMPSON, K., WILLIAMS, N.S.G. & VESK, P.A. (2017) Functional trait changes in the floras of 11 cities across the globe in response to urbanization. *Ecography* 40, 875–886.
- PAPAUX, A., BOURG, D. (2015) Dictionnaire de la pensée écologique, PUF



PAVOINE, S., BONSALE, M.B., DUPAIX, A., JACOB, U. & RICOTTA, C. (2017) From phylogenetic to functional originality: Guide through indices and new developments. *Ecological Indicators* 82, 196–205.

PAVOINE, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, 1–14.

PAVOINE, S. & RICOTTA, C. (2013) Testing for phylogenetic signal in biological traits: The ubiquity of cross-product statistics. *Evolution* 67, 828–840.

PAVOINE, S. & BONSALE, M.B. (2011) Measuring biodiversity to explain community assembly: A unified approach. *Biological Reviews* 86, 792–812.

PAVOINE, S., LOVE, M.S. & BONSALE, 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. *Ecology Letters* 12, 898–908.

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.

PEARMAN, P.B., PENSKAR, M.R., SCHOOLS, E.H. & ENANDER, H.D. (2006) Identifying potential indicators of conservation value using natural heritage occurrence data. *Ecological Applications* 16, 186–201.

PEARSON, R.G. (2016) *Reasons to Conserve Nature*. *Trends in Ecology & Evolution* 31, 366–371. Elsevier Ltd.

PENONE, C., DAVIDSON, A.D., SHOEMAKER, K.T., DI MARCO, M., RONDININI, C., BROOKS, T.M., YOUNG, B.E., GRAHAM, C.H. & COSTA, G.C. (2014) Imputation of missing data in life-history trait datasets: Which approach performs the best? *Methods in Ecology and Evolution* 5, 1–10.

PETCHEY, O.L. & GASTON, K.J. (2002) Extinction and the loss of functional diversity. *Proceedings of the Royal Society B: Biological Sciences* 269, 1721–1727.

PETCHEY, O.L. & GASTON, K.J. (2006) Functional diversity: Back to basics and looking forward. *Ecology Letters* 9, 741–758

PETERS, D.P.C., HAVSTAD, K.M., CUSHING, J., TWEEDIE, C., FUENTES, O. & VILLANUEVA-ROSALES, N. (2014) Harnessing the power of big data: Infusing the scientific method with machine learning to transform ecology. *Ecosphere* 5, 1–15.

PIMM, S.L., AYRES, M., BALMFORD, A., BRANCH, G., BRANDON, K., BROOKS, T., BUSTAMANTE, R., COSTANZA, R., COWLING, R., CURRAN, L.M., DOBSON, A., FARBER, S., DA FONSECA, G.A.B., GASCON, C., KITCHING, R., ET AL. (2001) Can we defy nature's end? *Science* 293, 2207–2208.

PIMM, S.L., JENKINS, C.N., ABELL, R., BROOKS, T.M., GITTLEMAN, J.L., JOPPA, L.N.,

- RAVEN, P.H., ROBERTS, C.M. & SEXTON, J.O. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science* 344.
- PODANI, J. & SCHMERA, D. (2007) How should a dendrogram-based measure of functional diversity function? A rejoinder to Petchey and Gaston. *Oikos*.
- POIANI, K.A., RICHTER, B.D., ANDERSON, M.G. & RICHTER, H.E. (2000) Biodiversity Conservation at Multiple Scales: Functional Sites, Landscapes, and Networks. *BioScience* 50, 133.
- POLLOCK, L.J., THUILLER, W. & JETZ, W. (2017) Large conservation gains possible for global biodiversity facets. *Nature* 546, 141–144. Nature Publishing Group.
- PRANCE, G.T. (2002) Species survival and carbon retention in commercially exploited tropical rainforest. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 360, 1777–1785.
- PROULX, R., WIRTH, C., VOIGT, W., WEIGELT, A., ROSCHER, C., ATTINGER, S., BAADE, J., BARNARD, R.L., BUCHMANN, N., BUSCOT, F., EISENHAUER, N., FISCHER, M., GLEIXNER, G., HALLE, S., HILDEBRANDT, A., ET AL. (2010) Diversity promotes temporal stability across levels of ecosystem organization in experimental grasslands. *PLoS ONE* 5, 1–8.
- PURVIS, A. & HECTOR, A. (2000) Getting the measure of biodiversity. *Nature* 405, 212–219.
- R Development Core Team. 2008 R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. See <http://www.r-project.org/>.
- RAO, C. R. (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology* 21, 24–43.
- RABINOWITZ, D. (1982) Seven forms of rarity. In: *The Biological Aspects of Rare Plant Conservation* (ed. H. Synge), Wiley, Chichester, 205–217.
- REDDING, D.W., MOOERS, A.O., ŞEKERCIOĞLU, Ç.H. & COLLEN, B. (2015a) Global Evolutionary Isolation Measures Can Capture Key Local Conservation Species in Nearctic and Neotropical Bird Communities. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, 1–7.
- REDDING, D.W. & MOOERS, A.O. (2015b) Ranking mammal species for conservation and the loss of both phylogenetic and trait diversity. *PLoS ONE* 10, 1–11.
- REDDING, D.W., MAZEL, F. & MOOERS, A. (2014) Measuring evolutionary isolation for conservation. *PLoS ONE* 9, 1–15.
- REDDING, D.W., HARTMANN, K., MIMOTO, A., BOKAL, D., DEVOS, M. & MOOERS, A. (2008) Evolutionarily distinctive species often capture more phylogenetic diversity than expected. *Journal of Theoretical Biology* 251, 606–615.
- REDDING, D.W. & MOOERS, A.O. (2006) Incorporating evolutionary measures into conservation prioritization. *Conservation Biology* 20, 1670–1678.

- REID, W., MILLER, K. (1989). Keeping options alive. The scientific basis for conserving biodiversity. World Resources Institute, a center for policy research (p. 130).
- REJMANEK, M. & RICHARDSON, D.M. (1996) What Attributes Make Some Plant Species More Invasive? WHAT ATTRIBUTES MAKE SOME PLANT SPECIES MORE INVASIVE? I. *Ecology* 77, 1655–1661.
- REVELL, L.J., HARMON, L.J. & COLLAR, D.C. (2008) Phylogenetic signal, evolutionary process, and rate. *Systematic biology* 57, 591–601.
- RICKLEFS, R.E. (1987) Community diversity: relative roles of local and regional processes. *Science* 235, 167–171.
- RICOTTA, C. & SZEIDL, L. (2009) Diversity partitioning of Rao's quadratic entropy. *Theoretical Population Biology* 76, 299–302. Elsevier Inc.
- RIPPLE, W.J., WOLF, C., NEWSOME, T.M., GALETTI, M., ALAMGIR, M., CRIST, E., MAHMOUD, M.I. & LAURANCE, W.F. (2017) World scientists' warning to humanity: A second notice. *BioScience* 67, 1026–1028.
- ROBUCHON, M., FAITH, D.P., JULLIARD, R., LEROY, B., PELLENS, R., ROBERT, A., THEVENIN, C., VERON, S. & PAVOINE, S. (2019) Species splitting increases estimates of evolutionary history at risk. Elsevier. *Biological Conservation*.
- ROCHE, D.G., KRUK, L.E.B., LANFEAR, R. & BINNING, S.A. (2015) Public Data Archiving in Ecology and Evolution: How Well Are We Doing? *PLoS Biology* 13, 1–12.
- RODRIGUES, A.S.L., GRAY, C.L., CROWTER, B.J., EWERS, R.M., STUART, S.N., WHITTEN, T. & MANICA, A. (2010) A Global Assessment of Amphibian Taxonomic Effort and Expertise. *BioScience* 60, 798–806.
- ROQUET, C., THUILLER, W. & LAVERGNE, S. (2013) Building megaphylogenies for macroecology: Taking up the challenge. *Ecography* 36, 013–026.
- DE QUEIROZ, K. (2005) Ernst Mayr and the modern concept of species. *Systematics and the Origin of Species: On Ernst Mayr's 100th Anniversary*, 243–263.
- DE QUEIROZ, K. (2007) Species concepts and species delimitation. *Systematic Biology* 56, 879–886.
- SAFI, K., ARMOUR-MARSHALL, K., BAILLIE, J.E.M. & ISAAC, N.J.B. (2013) Global Patterns of Evolutionary Distinct and Globally Endangered Amphibians and Mammals. *PLoS ONE* 8, 4–12.
- SAITO, V.S., CIANCIARUSO, M.V., SIQUEIRA, T., FONSECA-GESSNER, A.A. & PAVOINE, S. (2016) Phylogenies and traits provide distinct insights about the historical and contemporary assembly of aquatic insect communities. *Ecology and Evolution* 6, 2925–2937.
- SANDERSON, M.J. & SHAFFER, H.B. (2002) Troubleshooting molecular phylogenetic analyses. *Annual Review of Ecology and Systematics* 33, 49–72.

- SARRAZIN, F. & LECOMTE, J. (2016) Evolution in the Anthropocene. *Science* 351, 922–923.
- SARRAZIN, F. & BARBAULT, R. (1996) Reintroduction: challenges and lessons for basic ecology. *Trends in Ecology and Evolution* 11, 474–478.
- SEDDON, P.J. (2010) From Reintroduction to Assisted Colonization: Moving along the Conservation Translocation Spectrum 18, 796–802.
- SHANNON, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27, 379, 623–423, 656.
- SHIPLEY, B., VILE, D. & GARNIER, E. (2006) From Plant Traits to Plant Communities: A Statistical Mechanistic Approach to Biodiversity. *Science* 314, 812–814.
- SIEFERT, A., VIOLLE, C., CHALMANDRIER, L., ALBERT, C.H., TAUDIERE, A., FAJARDO, A., AARSEN, L.W., BARALOTO, C., CARLUCCI, M.B., CIANCIARUSO, M. V., DE L. DANTAS, V., DE BELLO, F., DUARTE, L.D.S., FONSECA, C.R., FRESCHE, G.T., ET AL. (2015) A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters* 18, 1406–1419.
- SIEFERT, A., FRIDLEY, J.D. & RITCHIE, M.E. (2014) Community Functional Responses to Soil and Climate at Multiple Spatial Scales: When Does Intraspecific Variation Matter? *PLOS One* 9.
- SIMPSON, G.G. (1961) *Principles of Animal Taxonomy*. Columbia University Press, New York.
- SIMPSON, E. H. (1949). Measurement of diversity. *Nature* 163, 688.
- SITES, J.W. & MARSHALL, J.C. (2004) Operational criteria for delimiting species. *Annual Review of Ecology, Evolution, and Systematics* 35, 199–227.
- SOLEY-BORI, M. (2013) Dealing with missing data: Key assumptions and methods for applied analysis. Boston University School of Public Health. <https://www.bu.edu/sph/files/2014/05/Marina-tech-report.pdf>.
- SOULÉ, M.E. (1985) What is Conservation Biology?: A new synthetic discipline addresses the dynamics and problems of perturbed species, communities, and ecosystems, *BioScience*, 35, 727–734
- SRIVASTAVA, D.S., CADOTTE, M.W., MACDONALD, A.A.M., MARUSHIA, R.G. & MIROTNICK, N. (2012) Phylogenetic diversity and the functioning of ecosystems. *Ecology Letters*.
- STALEY, J.T. & GOSINK, J.J. (1999) Poles Apart: Biodiversity and Biogeography of Sea Ice Bacteria. *Annual Review of Microbiology* 53, 189–215.
- STEEL, M., MIMOTO, A. & MOOERS, A.Ø. (2007) Hedging our bets: The expected contribution of species to future phylogenetic diversity. *Evolutionary Bioinformatics* 3, 237–244.
- STEIN, R.W., MULL, C.G., KUHN, T.S., ASCHLIMAN, N.C., DAVIDSON, L.N.K., JOY, J.B., SMITH, G.J., DULVY, N.K. & MOOERS, A.O. (2018) Global priorities for conserving the

evolutionary history of sharks, rays and chimaeras. *Nature Ecology & Evolution* 2, 288–298. Springer US.

STEKHOVEN, D.J. & BÜHLMANN, P. (2012) Miss forest. Non-parametric missing value imputation for mixed-type data. *Bioinformatics* 28, 112–118.

SUKHDEV, P., WITTMER, H., AND MILLER, D., (2014) The Economics of Ecosystems and Biodiversity (TEEB): Challenges and Responses, in D. Helm and C. Hepburn (eds) *Nature in the Balance: The Economics of Biodiversity*. Oxford: Oxford University Press.

SULLIVAN, M.S. & SWINGLAND, I.R. (2006) Extinction risk: Predicting and redressing the threat. *Biodiversity and Conservation* 15, 2009–2016.

SWENSON, N.G. (2009) Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PLoS ONE* 4.

SWINGLAND, I.R. (2013) Biodiversity, definition of. In: Levin, S.A. (ed.) *Encyclopedia of Biodiversity*. vol. 1. Academic, San Diego, pp. 377–391.

TAYLOR, G., CANESSA, S., CLARKE, R.H., INGWERSEN, D., ARMSTRONG, D.P., SEDDON, P.J. & EWEN, J.G. (2017) Is Reintroduction Biology an Effective Applied Science? *Trends in Ecology and Evolution* 32, 873–880. Elsevier Ltd.

THESEN, A.E. (2016) Adoption of Machine Learning Techniques in Ecology and Earth Science. *One Ecosystem* 1, 1–38.

THULLER, W., MAIORANO, L., MAZEL, F., GUILHAUMON, F., FICETOLA, F., LAVERGNE, S., RENAUD, J., ROQUET, C. & MOUILLOT, D. (2015) Conserving the functional and phylogenetic trees of life of European tetrapods. *Phil. Trans. R. Soc. B* 370: 370.

TILMAN, D., ISBELL, F. & COWLES, J.M. (2014) Biodiversity and Ecosystem Functioning. *The Annual Review of Ecology, Evolution, and Systematics* 45, 471–493.

TILMAN, D., KNOPS, J., WEDIN, D., REICH, P., RITCHIE, M. & SIEMANN, E. (1997) The influence of functional diversity and composition on ecosystem processes. *Science* 277, 1300–1302.

TITTENSOR, D.P., WALPOLE, M., HILL, S.L.L., BOYCE, D.G., BRITTEN, G.L., BURGESS, N.D., BUTCHART, S.H.M., LEADLEY, P.W., REGAN, E.C., ALKEMADE, R., BAUMUNG, R., BELLARD, C., BOUWMAN, L., BOWLES-NEWARK, N.J., CHENERY, A.M., ET AL. (2014) A mid-term analysis of progress toward international biodiversity targets. *Science* 346, 241–245.

TUCKER, C.M., AZE, T., CADOTTE, M.W., CANTALAPIEDRA, J.L., CHISHOLM, C., DIAZ, S., GRENYER, R., HUANG, D., MAZEL, F., PEARSE, W.D., PENNELL, M.W., WINTER, M. & MOOERS, A.O. (2019) Assessing the utility of conserving evolutionary history. *Biological Reviews*, brv.12526.

TUCKER, C.M., DAVIES, T.J., CADOTTE, M.W. & PEARSE, W.D. (2018) On the relationship between phylogenetic diversity and trait diversity. *Ecology* 99, 1473–1479.

VAN VALEN, L. (2009) Ecological Species, Multispecies, and Oaks. *Taxon* 25, 233–239.

VALLET, J., DANIEL, H., BEAUJOUAN, V., ROZE, F. & PAVOINE, S. (2010) Using biological traits to assess how urbanization filters plant species of small woodlands. *Applied Vegetation Science* 13, 412–424.

VANE-WRIGHT, R.I., HUMPHRIES, C.J. & WILLIAMS, P.H. (1991) What to Protect? Systematics and the Agony of Choice. *Biological Conservation* 55, 235–254.

VÁZQUEZ, D.P., MELIAN, C.J., WILLIAMS, N., BLUTHGEN, N., KRASNOV, B.R. & POULIN, R. (2007) Species abundance and asymmetric interaction strength in ecological networks. *Oikos* 116, 1120–1127.

VECCHI, D., MILLS, R. (2019) Probing the Process-Based Approach to Biodiversity: Can Plasticity Lead to the Emergence of Novel Units of Biodiversity? In: Casetta E., Marques da Silva J., Vecchi D. (eds.) *From Assessing to Conserving Biodiversity. History, Philosophy and Theory of the Life Sciences*, 24 Springer, Cham.

VELLEND, M., CORNWELL, W. K., MAGNUSON-FORD, K. & MOOERS, A. Ø. (2011). Measuring phylogenetic biodiversity. In *Biological Diversity: Frontiers*, A. E. Magurran and B. J. McGill (eds), In *Measurement and Assessment*, pp. 194–207. Oxford University Press, Oxford.

VERON, S., DAVIES, T.J., CADOTTE, M.W., CLERGEAU, P. & PAVOINE, S. (2017) Predicting loss of evolutionary history: Where are we? *Biological Reviews* 92, 271–291.

VERON, S., CLERGEAU, P. & PAVOINE, S. (2016) Loss and conservation of evolutionary history in the Mediterranean Basin. *BMC Ecology* 16. BioMed Central.

VILLÉGER, S., MAIRE, E. & LEPRIEUR, F. (2017) On the risks of using dendrograms to measure functional diversity and multidimensional spaces to measure phylogenetic diversity: a comment on Sobral et al. (2016). *Ecology Letters* 20, 554–557.

VILLEGER, S., MASON, N. W. & MOUILLOT, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* 89, 2290–2301.

VIOLLE, C., THUILLER, W., MOUQUET, N., MUNOZ, F., KRAFT, N.J.B., CADOTTE, M.W., LIVINGSTONE, S.W. & MOUILLOT, D. (2017a) Functional Rarity: The Ecology of Outliers. *Trends in Ecology & Evolution in Ecology*.

VIOLLE, C., THUILLER, W., MOUQUET, N., MUNOZ, F., KRAFT, N.J.B., CADOTTE, M.W., LIVINGSTONE, S.W., GRENIÉ, M. & MOUILLOT, D. (2017b) A Common Toolbox to Understand, Monitor or Manage Rarity? A Response to Carmona et al. *Trends in Ecology & Evolution*, 2–5. Elsevier Ltd.

VIOLLE, C., ENQUIST, B.J., MCGILL, B.J., JIANG, L., ALBERT, C.H., HULSHOF, C., JUNG, V. & MESSIER, J. (2012) The return of the variance: Intraspecific variability in community ecology. *Trends in Ecology and Evolution* 27, 244–252.

VIOLLE, C. & JIANG, L. (2009) Towards a trait-based quantification of species niche. *Journal of Plant Ecology* 2, 87–93.

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.
- WALTHER, G., ROQUES, A., HULME, P.E., SYKES, M.T., PYS, P., ROBINET, C. & SEMENCHENKO, V. (2009) Alien species in a warmer world : risks and opportunities. *Trends in Ecology and Evolution* 24, 686–693.
- WATERS, C.N., ZALASIEWICZ, J., SUMMERHAYES, C., BARNOSKY, A.D., POIRIER, C., GALUSZKA, A., CEARRETA, A., EDGEWORTH, M., ELLIS, E.C., ELLIS, M., JEANDEL, C., LEINFELDER, R., MCNEILL, J.R., RICHTER, D.D.B., STEFFEN, W., ET AL. (2016) The Anthropocene is functionally and stratigraphically distinct from the Holocene. *Science* 351.
- WEBB, C.O. (2000) Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest. Source: *The American Naturalist* 156, 145–155.
- WEBB, C.O., ACKERLY, D.D., MCPEEK, M.A. & DONOGHUE, M.J. (2002) Phylogenies and community ecology. *Annu. Rev. Ecol. Syst.* 33, 475–505.
- WEINSTEIN, B.G., TINOCO, B., PARRA, J.L., BROWN, L.M., MCGUIRE, J.A., GARY STILES, F. & GRAHAM, C.H. (2014) Taxonomic, phylogenetic, and trait beta diversity in south American hummingbirds. *American Naturalist* 184, 211–224.
- WHITTAKER, A.R.H. (1972) Evolution and Measurement of Species Diversity. *Taxon* 21, 213–251.
- WILEY, E.O. (1978) The evolutionary species concept reconsidered. *Systematic Zoology* 27, 17–26.
- WILLIAMS, M., ZALASIEWICZ, J., HAFF, P.K., SCHWÄGERL, C., BARNOSKY, A.D. & ELLIS, E.C. (2015) The anthropocene biosphere. *Anthropocene Review* 2, 196–219.
- WILLIAMS J.M. (2004) Preface. In: *Missing Links: Connecting Science with Environmental Policy*. Parliamentary Commissioner for the Environment, Te Kaitiaki Taiao a Te Whare Paremata PO Box 10-241, Wellington, New Zealand.
- WILSON, E. O. (1984). *Biophilia*. Cambridge: Harvard University press.
- WILSON, E.O. & BROWN, W.L. (1953) The Subspecies Concept and Its Taxonomic Application Author(s): *Systematic zoology* 2, 97–111.
- WINTER, M., DEVICTOR, V. & SCHWEIGER, O. (2013) Phylogenetic diversity and nature conservation: Where are we? *Trends in Ecology and Evolution* 28, 199–204.
- ZHAO, T., VILLEGER, S. & CUCHEROUSSET, J. (2019) Accounting for intraspecific diversity when examining relationships between non-native species and functional diversity. *Oecologia* 189, 171–183. Springer Berlin Heidelberg.
- ZOBEL, M. (1997) The relative role of species pools in determining plant species richness: An alternative explanation of species coexistence? *Trends in Ecology and Evolution* 12, 266–269.

Extra projects and other doings, which are indirectly related to a research activity of PhD students and researches in general are often forgotten. These usually voluntary work represents, however, an important investment of time and energy. I would like to briefly present here some main activities, which I was happy to be a part of. All of them provided me a great personal experience, whether it was by a popularization of evolution and ecology to a primary school class, by giving botany classes to people passionate by nature and biology during public lectures of MNHN or by leading an association of Master and PhD students (BDEM) during two past years. It was a great pleasure to have a possibility to participate to all those projects, which also opened me different worlds of research, teaching, administration, event organization and many others. As I am planning to continue my past activities before I started to work on this thesis project, which excites me particularly, that is to spread the scientific knowledge to kids, I also participated to a meeting of EvoKE (Evolutionary Knowledge for Everyone) association in Croatia.

Therefore, I would like to thank all the people who I was very glad to meet during these three past years, people which contributed to my personal development in any form.

- January - June 2018: Teaching Theory of Evolution at primary school "Parmentier", Paris, during science classes (activity report can be found in Appendix C)
- 26 - 29 September 2019: EvoKE 2019 meeting, Split, Croatia
- May 2017 – May 2019: Head of PhDs and Master Students association, "Bureau des Doctorants et Etudiants du Muséum" (BDEM) of MNHN
- General organization and coordination of Young Natural History scientists Meeting (YNHM) in 2018 and 2019: <https://ynhm2019.sciencesconf.org/>; <https://ynhm2018.sciencesconf.org/>
- Public lectures for National Museum of Natural History in botany (Angiosperms reproduction, Plant community ecology, Urban plants ecology)
- Participation to the Vigie-Nature stand during "Fête de la Nature" week-ends in 2018 and 2019



## LIST OF PUBLICATIONS

---

**Kondratyeva, A.**, Grandcolas, P., and Pavoine, S. (2019). Reconciling the concepts and measures of diversity, rarity and originality in ecology and evolution. *Biol. Rev.* 94, 1317–1337. doi:10.1111/brv.12504

**Kondratyeva, A.**, Knapp, S., Durka, W., Kühn, I., Vallet, J., Machon, N., Martin, G., Motard, E., Grandcolas, P., and Pavoine, S. (2019). Diversity in the city: urbanization affects differently trait-based and phylogenetic plant originality at local and regional scales. In revision for *Frontiers Ecology and Evolution*, section Urban Ecology.

S. Veron, R. Pellens, **A. Kondratyeva**, S. Diaz, W. Illes, P. Grandcolas, S. Pavoine, M. Robuchon and M. Mouchet. (2019). On islands, evolutionary but not functional originality is rare. In prep.

## LIST OF COMMUNICATIONS

---

**Chaire Modélisation mathématique et Biodiversité (MMB)**, Aussois, 28 mai – 1 juin 2017, oral « *Mesures d'originalité phylogénétique et fonctionnelle* », Kondratyeva, A., Grandcolas, P., Pavoine, S.

**SFEcologie 2018** – International Conference on Ecological Sciences, Rennes, 22-25 October 2018, poster presentation “*Species’ diversity, rarity and originality. Disentangling concepts and measures in ecology and evolution*” Anna Kondratyeva, Philippe Grandcolas, Sandrine Pavoine (award of the best young researcher poster); Oral presentation « *Originalité urbaine des plantes herbacées, région Ile-de-France* », Kondratyeva, A., Knapp, S., Durka, W., Kühn, I., Vallet, J., Machon, N., Martin, G., Motard, E., Grandcolas, P., et Pavoine, S.

**Vigie-flore program days** (27-28 April 2019), oral presentation « *Originalité urbaine des plantes herbacées, région Ile-de-France* », Kondratyeva, A., Knapp, S., Durka, W., Kühn, I., Vallet, J., Machon, N., Martin, G., Motard, E., Grandcolas, P., et Pavoine, S.

**Petit pois déridé**, 4-6 June 2019, oral presentation « *Originalité urbaine des plantes herbacées, région Ile-de-France* », Kondratyeva, A., Knapp, S., Durka, W., Kühn, I., Vallet, J., Machon, N., Martin, G., Motard, E., Grandcolas, P., et Pavoine, S.

## APPENDIX A

# ON ISLANDS, EVOLUTIONARY BUT NOT FUNCTIONAL ORIGINALITY IS RARE

---

# On islands, evolutionary but not functional originality is rare

Authors: S. Veron<sup>1,2,\*</sup>, R. Pellens<sup>1</sup>, A. Kondratyeva<sup>2</sup>, P. Grandcolas<sup>1</sup>, Rafaël Govaerts<sup>3</sup>, M. Robuchon<sup>2</sup>, M. Mouchet<sup>2</sup>

<sup>1</sup>Institut de Systématique Evolution et Biodiversité (ISYEB UMR7205), Muséum National d'histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, CP51, 47 rue Buffon, 75005, Paris, France

<sup>2</sup>Centre d'Ecologie et des Sciences de la Conservation (CESCO UMR7204) MNHN, CNRS, Sorbonne Université - CP51, 55-61 rue Buffon, 75005, Paris, France

<sup>3</sup>Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AE, UK.

\*Author for correspondence: e-mail: [sveron@edu.mnhn.fr](mailto:sveron@edu.mnhn.fr); Tel. +33 1 40 79 57 63; Fax: +33 1 40

79 38 35

## **ABSTRACT**

Functionally and evolutionary original species are those whose traits or evolutionary history are shared by few others in a given set. These original species promote ecosystem multifunctionality, the ability to cope with an uncertain future, future benefits to society and therefore have a high conservation value. A potential signal of their extinction risks is their rarity (stating for geographic range-restriction in this study). In islands life in isolation conducted to the rise of a multitude of original forms and functions as well as to high rates of endemism. Not only patterns and processes of insular originality are unexplained but the relationship between originality and rarity is still unknown. The aim of this study is to assess how original insular species are, to explore whether original species are rare or not and to investigate the factors that may explain the rarity of original species. We first compared the functional and evolutionary originality of monocotyledon species and whether continental or insular species were more original. We found that species restricted to islands were more original than continental species and, although functionally and evolutionary original species were dissimilar, many occurred on similar territories so that regional conservation strategies may allow to conserve these distinct forms. Yet, evolutionary original species were significantly more range-restricted than those which were distinct in their traits. Reflecting their rarity, evolutionary original species had low dispersal abilities and were found on islands where settlement may have been facilitated. On the opposite, functionally original species could reach a wider set of islands by being transported on long-distances. While some mechanisms may both explain rarity and originality such as extinctions, others may be specific to each of these biodiversity facets, in particular diversification, niche shift and expansion, and dispersal power. Implications for conservation are huge: original species are range-restricted and mostly found in the most threatened systems of the world, i.e. islands, endangering the reservoir of features against an uncertain future.

## INTRODUCTION

Originality of a given species in a set is the oddity of its biological characteristics, whatever they are, relatively to all other species of the set (Sandrine Pavoine, Bonsall, Dupaix, Jacob, & Ricotta, 2017). Originality can refer to evolutionary history or to functional traits (e.g., Pavoine, Ollier, & Dufour, 2005; Pavoine et al., 2017; Marc W. Cadotte & Jonathan Davies, 2010; Redding, Mazel, & Mooers, 2014 ; D. Mouillot, Culioli, Pelletier, & Tomasini, 2008 ; Violle et al., 2017 ; Cornwell et al., 2014). Originality has been poorly investigated in community ecology or biogeography (but see David Mouillot, Graham, Villéger, Mason, & Bellwood, 2013) but its value for biodiversity conservation is well supported. Indeed, in addition to capture significant amount of phylogenetic or functional diversity (D. Mouillot et al., 2008; Faith, 1992), evolutionary and functionally original species may insure key ecosystem processes and provide services to humanity. Functional and evolutionary originality may sometimes be decoupled (Faith, 1992; Pollock, Thuiller, & Jetz, 2017) and should be differentiated due to the distinct benefits they may provide. Regarding measures of functional originality, they are based on a reduced number of traits generally related to ecosystem functioning. For example, (Petchey, Hector, & Gaston, 2004) used 12 plant traits and showed that functionally original species may have a large contribution to ecosystem functions by increasing plant biomass production. Loss of functionally original species may then directly impact ecosystem stability, resilience and multifunctionality (Fonseca & Ganade, 2001; Pendleton, Hoetinghaus, Gomes, & Agostinho, 2014; Bracken & Low, 2012; Oliver et al., 2015). Evolutionary original species may also have a leading contribution to ecosystem functioning because phylogenetic isolation may be an indicator of distinct species ecological roles (M. W. Cadotte, Cardinale, & Oakley, 2008; (Redding et al., 2008; Marc W. Cadotte & Jonathan Davies, 2010). Most of all, phylogenetic diversity represents a reservoir of yet-to-be-discovered resources to humanity and evolutionary original species may highly contribute to these option-values (Faith, 2017). Using together functional and evolutionary originality may allow to identify species that capture irreplaceable ecosystem functions and services to humanity. However, functional and evolutionary

originality have rarely been studied jointly. Systems in which the study of both biodiversity facets could be highly valuable are islands. Insular systems have a huge importance in biodiversity conservation to preserve the heritage of our planet and they have a leading role of “natural laboratory” for biogeography and evolutionary ecology (Robert J. Whittaker, Fernández-Palacios, Matthews, Borregaard, & Triantis, 2017; Warren et al., 2015). Islands are home of species which are found nowhere else on Earth and which represent a unique evolutionary history as well as a multitude of forms and functions. Original species may have a high contribution to the island biota, but their role in shaping insular diversity just begin to be investigated (Veron, Haeevermans, Govaerts, Mouchet, & Pellens, 2019; Veron et al., 2019). In addition, very little is known about how originality arose and is distributed, so that investigations on this topic may bring insights that go beyond the scope of island biogeography.

A pre-requisite to the joint study of functional and evolutionary originality is to assess how these biodiversity facets are related. As stated above, findings about this issue are contradictory and it has rarely been studied in an insular context. On islands, the relationship between phylogenetic and functional originality may be weak as early diversification and simultaneous cladogenesis resulted in the lack of phylogenetic signal in insular species traits (Losos, 2008). A famous example is the radiation of *Bidens* in Hawaii: closely-related species have greater diversity of traits (e.g., growth form, floral morphology) than in the rest of the world probably due to the diversity of habitat types and to loss of dispersability (Knopé, Morden, Funk, & Fukami, 2012). As for Hawaiian *Drosophila*, sexual selection has resulted in clearly distinguishable traits among closely-related (Gillespie & Clague, 2009). In addition, phenotypic character changes of insular species do not necessarily involve speciation or a phylogenetic branching pattern so that variation in evolutionary history do not reflect the variation of some specific traits (R. J. Whittaker & Fernández-Palacios, 2007). The first issue we investigated was therefore the extent to which functional and evolutionary originality are related and we specifically tested the assumption that evolutionary and functional originality are less correlated than on the mainland.

Another unresolved question is whether originality is higher on islands in comparison to continental areas. On islands, a long evolutionary history in isolation coupled with unfilled niches and oceanic climate may have given rise to many evolutionary and functionally original species. Extinctions on the continents may also have isolated species in the tree of life and in the functional trait space, creating evolutionary and functionally original insular species (Rosemary G. Gillespie & Roderick, 2002; Grandcolas, Nattier, & Trewick, 2014)). However, whether insular species are more original than continental has been poorly studied so far and maybe highly taxa specific (Jetz et al., 2014). We therefore explored whether insular species are more original than continental ones, which may add a new line of argument for the preservation of insular biodiversity.

Some original species may be at risk due to their rarity (here stating for geographic range-restriction). Range-restriction is indeed a key factor of extinction risks. Many research found that low geographic range size was the main predictor of these risks and it is also one of the main factors of the threat status of species in the IUCN RedList (Rodrigues, Pilgrim, Lamoreux, Hoffmann, & Brooks, 2006; Bland, Collen, Orme, & Bielby, 2015; Veron et al., 2016). Like original species, range-restricted species provide unique functions in an ecosystem as well as unique services to humanity (David Mouillot, Graham, et al., 2013; Leitão et al., 2016). Species which are both rare and original may then support highly vulnerable and unique functions and option values (Rosauer, Laffan, Crisp, Donnellan, & Cook, 2009; David Mouillot, Graham, et al., 2013; Violle et al., 2017 ; Faith, 2017). Many of such species may be found on islands. Indeed, a remarkable feature of islands is their high rates of endemism which may outcompete these in the mainland by a factor of 9 (Kier et al., 2009). More generally, many insular species are spatially range-restricted and found only on a few islands. Yet, how range-restricted are original species on islands is still poorly known and, as far as we know, has never been tested in the case of functional originality. To fill this gap, we explored the potential risks of losing the most original species by assessing the relationship between species originality and rarity. Beyond practical conservation implications, assessing the relationship between originality and rarity may allow to shed light on the mechanisms shaping both originality

and range-restriction. For example, past extinctions may have both isolated species on a phylogenetic tree and/or function space and shrunk species area of distribution. Rarity of original species can also be influenced by their dispersal capacities (Flather et al. 2007) which may be lineage specific. In an insular context, factors influencing colonization rates, such as the possibility of an island to have been at reach during its geological history and species dispersal traits have a strong influence on the distribution of biodiversity on islands (Rosemary G. Gillespie et al., 2012; Patrick Weigelt et al., 2015). We therefore investigated how the dispersal capacities of original species and the features of the islands they occur on can give evidence of their rarity.

Originality on islands is still a pristine research field and through our analyses we tackled four main questions that may lay the foundations of the study of originality on islands having implications in both conservation and biogeography 1) Are evolutionary and functionally original species similar on islands? 2) Are insular species more evolutionary and functionally original than mainland species? 3) Are insular original species rare? 4) Can the rarity of original species be attributed to their dispersal capacities and to the biogeographic characteristic of islands? To address these issues we focused on the group of Monocotyledons (Monocots), a morphologically and functionally diverse clade representing a quarter of flowering plant diversity such as all orchids, palms and cereals. The origin of monocot species represents a diversity of evolutionary and ecological processes. Monocots species are also wide-spread across islands and continents, their phylogeny is well-resolved and their traits well-documented. The clade of monocotyledon is therefore a well-suited group for the study of originality on islands.

## **METHODS**

### **Occurrence data**

We used the e-monocot database (emonocot.org) to extract data on both the spatial distribution of monocot species and delimitations of 126 islands (TDWG 3rd level). E-monocot compiles records from several botanical institutions for all 70,000 monocotyledons. We kept only native and terrestrial species from this database.



## **Estimating evolutionary and functional originality**

We followed results from (Sandrine Pavoine et al., 2017) to choose the metrics of evolutionary and functional originality (EO and FO, respectively). We followed the recommendations of the authors to use a distance-based metric over dendrogram metrics because the latter may be biased by clustering methods, the use of a low number of traits and of the use of non-ultra-metric trees (Mouchet et al., 2008; Pavoine et al., 2017). We first estimated functional and phylogenetic distances thanks to the Gower's distance (Legendre & Legendre, 2012) and derived an originality score calculated with the `distinctDis` function of the package `adiv` (S. Pavoine, 2019). EO is estimated in million years of evolution and FO has no units and ranges between 0 and 1.

We estimated monocot FO with six traits compiled by (Díaz et al., 2015), adult plant height, stem specific density, leaf size expressed as leaf area, leaf mass per area, leaf nitrogen content per unit mass, and diaspore mass (see Díaz et al., 2015 for a description). The traits we chose have been widely used and recognized as fundamentally representatives of plant ecological strategies (Díaz et al., 2015). Missing values were imputed by performing random forest algorithm a thousand times, and by then estimating the mean of the 1000 imputation. We performed a sensitivity analysis (Appendix 1) to i. assess the extent to which missing values may have influenced our results ii. estimate the correlation between traits iii. Assess the contribution of each trait to FO scores. We did not include dispersal mode in the measure of functional originality because it was used later as an explanatory factor of the originality-rarity relationship. By doing so we assumed that dispersal mode was more related to the geographic extent of a species than to its function in an ecosystem or its response to environmental conditions.

To estimate EO thanks to the `distinctDis` function, the Monocot phylogeny we used was built by extracting monocotyledon species from a larger supertree (Qian & Jin, 2016) which is an updated version of a mega-phylogeny of plant species (Zanne et al., 2013). EO was estimated in million years of evolution and FO has no units and ranges from 0 (lowest originality possible) to 1 (highest originality possible). Due to differences in data availability, and in order to estimate originality on

the largest sets of species possible, evolutionary originality was estimated on a set of 6.682 species and FO on a set of 2.281 species.

### **Are species more original on continents or on islands?**

We distributed species among three non-overlapping categories a) insular endemics, i.e. present only on islands; b) continental endemics, i.e. present only on continents; c) insular non-endemic species, i.e. present on both continents and islands. We then estimated the distribution and the average of originality scores among the three categories. We tested whether the average originality was significantly different between the three categories: we randomized species among categories 1000 times, estimated the new average originalities per category in the randomized set and compared them to the observed originalities in each category. A correction under phylogenetic constraint of the average originality score and its significance per category is provided in Appendix 2 but did not influence the results.

### **Relationship between evolutionary and functional originality**

For each category of species, we performed correlation tests between functional and evolutionary originality and measured the phylogenetic signal of each of the six traits individually (Kstar test). To perform these particular analyses, we excluded species that did not have both phylogenetic and functional information. As EO and FO scores are relative to the set of species used to calculate originality, estimating their correlation required to re-calculate a second EO score from the similar set used to estimate FO (i.e. 2.281 species).

Considering again the original set of species (i.e. 6.682 species for EO and 2.281 species for FO), we then only focused on species present in islands (insular endemic and non-endemic species) and drew maps of the number of top 5% original species, i.e. the 5% most original species, in islands both for FO and EO. This included 110 non-endemic species and 40 insular endemics for EO and 38 insular non-endemic and 15 insular endemics for FO.

### **Are insular original species geographically rare?**

Focusing on species present on islands, i.e. both insular endemics and non-endemics, we tested the correlation (Pearson's test) between the originality score and geographic rarity, measured as the number of islands each species occur on. The fewer islands a species occur on the rarer it is. We also used linear regressions corrected for phylogenetic signal in model residuals (Revell, 2010), when necessary. We then specifically tested whether the most original species occurred on fewer islands than expected. We classified species with "top originality", "high originality", "moderate originality", "low originality", "very low originality" (ranked in the 5%, 25%, 50%, 75% and >75% of the originality score distribution, respectively) and calculated for each class the mean number of islands each species occurred on. We randomized species among classes 1000 times and estimated whether the observed mean number of islands species occurred on per originality class was lower than in the randomized set (see also Appendix 2). Finally, we identified islands where species were both among the top original species and were single endemics (found on a single island). We did not use indices incorporating both originality and geographic rarity as they may sometimes be less relevant for practitioners (Rosauer et al., 2009) and may be difficult to interpret at the functional level.

### **Dispersal modes, biogeographic characteristic of islands and species originality**

Coupling dispersal features of original insular species and the characteristics of the islands where they occur may help to understand the distribution of original species and possibly their rarity (e.g., Veron, Haeevermans, et al., 2019). First, we compiled information on insular species dispersal mode from (Carvajal-Endara, Hendry, Emery, & Davies, 2017) as well as from the TRY, SID, FRUBASE, BROT and LEDA databases. Dispersal modes corresponded to dispersal by animal, wind, water, unassisted and others (mainly transportation by humans). Species dispersal strategy was then estimated as either long-distance dispersal (animal, wind, water), short distance (unassisted) or unknown (others). We calculated the mean species originality per dispersal mode and strategy and

assessed their significance by using null models based on the random distribution of dispersal modes among species. In the case of zoochory, we also explored the effect of the identity of the species dispersing seeds (bird [flying/non flying], mammal [flying/non flying], reptile [terrestrial-marine], insects) on originality (Appendix 3, in prep.). However, diaspore mass is a trait used in our measure of FO that may be related to dispersal strategy and their association thus needed to be tested. We found that long-distance dispersal was moderately related to heavy diaspore mass (Welch two-sided test  $p\text{-val}=0.07$ ). Moreover, as “diaspore mass” did not prevail on other traits in the estimation of FO scores (Appendix 1), the relationship between FO and dispersal strategy may have been poorly influenced by the moderate association between “diaspore mass” and “dispersal strategy”.

Following our investigation on dispersal strategies of original species, we investigated how the possibility of an island to be/have been at reach may explain the distribution of original species. In particular, we focused on the biogeographic characteristics of islands that may be linked to past, present and future dispersal events (Table 1) although we acknowledge that some of these island features may also drive species diversification and therefore originality.

**Table 1:** Assumptions about how island features can relate to species geographic rarity

<i>Island characteristics</i>	<i>Assumption about the relation to dispersal potential</i>	<i>Reference</i>
<i>Distance to the continent (km)</i>	An original species having reached a remote island may be a good disperser	Global Island Database (GID; (UNEP-WCM, 2013))
<i>Proportion of surrounding landmass</i>	Colonization events may be less frequent on/from isolated islands	(P. Weigelt, Jetz, & Kreft, 2013)
<i>Elevation (m)</i>	Elevation may provide a climatic refuge for ancient lineages which may consequently rarely disperse	(P. Weigelt et al., 2013)
<i>Area (m<sup>2</sup>)</i>	Successful colonization events may be more probable towards/from large islands	GID (Depraetere & Dall, 2007)

<i>Oceanic or Continental</i>	Species present on oceanic islands may occur following colonization and are likely to be relatively good dispersers	(P. Weigelt et al., 2013)
<i>Glaciated or non-Glaciated</i>	The presence of ancient species on a glaciated island may be due to colonization and tendency to dispersal	Literature; pers. comm.
<i>Age (million years)</i>	An ancient species occurring on a recent island may reflect colonization, and the probability of the species to be a good disperser may be higher	Literature; pers. com.
<i>Latitude (decimal degrees)</i>	Variable used to highlight a possible spatial effect	GID (UNEP-WCM, 2013)
<i>Longitude (decimal degrees)</i>	Variable used to highlight a possible spatial effect	GID (UNEP-WCM, 2013)

Our aim here was to assess the characteristics of the islands where original species occur, potentially revealing the tendency for species to disperse or not (Table 1). To do so, we estimated for each species the average value of each of the biogeographical features of the islands in which it occurred. Each species was therefore associated to 9 values representing these average island features (Table 1). We then performed generalized linear models and multi-model selection with species originality as the response variable and island average features as the explanatory variables (see also Appendix 2 for models performed under phylogenetic constraint). We identified the strongest interactions thanks to Boosted Regression Trees, which were added in the multi-model selection process. To complement these analyses, we used the originality classification previously described (top, high, moderate, low and very low originality) and estimated which island features had a significant effect in each category of originality (Appendix 4, in prep.).

## RESULTS

### Functional and evolutionary originality are decoupled

The diaspore mass of insular species excepted, none plant functional trait exhibited a phylogenetic

signal (Table 2). Functional and evolutionary originality were weakly correlated, although the correlation coefficient was higher in species occurring on islands (Pearson's correlation test:  $cor=0.16, 0.13, 0.018$  and  $0.093$  in insular endemics, insular non-endemics, continental endemics and all monocot species, respectively). This result was robust to our sensitivity tests (Appendix 1).

**Table 2:** Phylogenetic signal (Kstar) in 6 ecological plant traits. Black values mean that there is no phylogenetic signal and bold red value indicate a phylogenetic signal.

	Insular endemics	Insular non-endemics	Continental endemics	All monocots
	<b>Phylogenetic signal (Kstar)</b>			
Leaf area	0.05***	0.028***	0.02**	0.019***
Nmass	0.002	0.002	0.009***	0.002
LMA	0.003	0.006**	0.008**	0.006***
Plant height	0.06***	0.04***	0.05***	0.04***
Diaspore mass	<b>0.59***</b>	0.03**	0.003	0.003
SSD.combined	0.002	0.001	0.009***	0.002

### Species endemic to islands are more original

Evolutionary originality was estimated on a set of 6.682 species and FO on a set of 2.281 species. Islands with the highest number of evolutionary original endemic monocotyledons were Borneo (being home of 10 top 5% original species), the Philippines (9 species), Sumatra (7 species) and Japan (6 species; Figure 1). The maximum concentration of species with top 5% functional original endemic species in a given island was 4 (Philippines), while several other islands were home of 3 top 5% functionally original endemic species (New Guinea, New Zealand, Norfolk Island). Thirteen islands were home of both top 5% functionally original endemic species and top 5% evolutionary original endemic species.

In non-endemic species, the highest numbers of top 5% evolutionary original species were found

in Taiwan (22 species), Hainan (20 species), Japan (17 species), Borneo and Java (16 species). Regarding FO, Cuba was home of 11 top 5% original species, Java and Dominican republic of 9 species while Hainan, the Bahamas, and Trinidad and Tobago harbored 8 top 5% functionally original species. There were 63 islands where both top 5% evolutionary and functionally non-endemic original species occurred.

By estimating the average originality in function of the endemic or non-endemic status of species we found that mean EO and FO was equal to 236.7 and 0.061 for species endemic to islands, 234.9 and 0.043 for non-endemic insular species and 224.5 and 0.047 for continental species only (Figure 2). Average EO was higher than expected for species endemic to islands and those found only in continents. Using phylogenetic analysis of variance models, EO was the highest for insular endemic species (Appendix 2). Average FO was significantly high for insular species whether they were endemic or not.

### **Evolutionary but not functionally original species are unusually rare**

The top 5% original endemic species regarding EO and FO were present on, on average, 1.7 and 3.4 islands, respectively. The original non-endemic species occurred on 5.0 to 5.4 islands, on average. Geographic range had a significant negative effect on originality for EO of insular non-endemic species (Pearson's test:  $cor = -0.15$   $p\text{-value} = 4.794e-11$ ) and a significant positive effect on FO of insular endemic species ( $cor = 0.43$   $p\text{-value} = 3.794e-7$ ). In other cases, no significant effect was found (EO endemics:  $cor = -0.043$   $p\text{-value} = 0.26$ ; FO non-endemics:  $cor = -0.028$   $p\text{-value} = 0.28$ ).

Regarding the classification of species in function of their originality score, the top 5% EO endemic species on fewer islands than expected by chance while top FO species occurred on more islands than expected (Table 3, Appendix 2). For insular non-endemic species, high EO species also occurred on fewer islands than expected by chance (Table 3, Appendix 2). Conversely, species with low or very low EO occurred on more islands than expected by chance. As for non-endemic FO, those with moderate FO occurred on more islands than expected.

**Table 5:** Average geographic range of species, i.e. the number of islands it occurs on, and whether it is lower or larger than expected by chance per category of originality (one-tailed tests). Red color means that species occurs on fewer islands than expected by chance (p-value<0.05). Blue color means that species occurs on more islands than expected by chance (p-value<0.05).

	Distribution	Top original	Very original	Moderate	Low	Very Low
Evolutionary Originality	Insular endemics	Range=1.7	Range=1.7	Range=2.1	Range=2.0	Range=1.8
	Insular non-endemics	Range=5.0	Range=3.7	Range=5.9	Range=7.4	Range=6.4
Functional Originality	Insular endemics	Range=3.2	Range=2.4	Range=1.7	Range=1.8	Range=1.8
	Insular non-endemics	Range=5.4	Range=6.4	Range=7.2	Range=6.1	Range=6.0

31 species were both endemic to a single island and among the top 5% most evolutionary original species. Among the areas which concentrated the highest number of these rare and original species, Borneo harbor 7 of them, 5 were found in the Philippines, Japan and Sumatra, and 3 in Madagascar (Figure 3). 6 species were single endemics and top 5% functionally original. They occurred on Madagascar (2 species), Lord Howe island (2 species), Bermuda (1 species) and the Canaries (1 species). Regarding non-endemic species 27 were top evolutionary original species and found on a single island. They were mainly concentrated on Equatorial Guinea islands (7 species), Japan (5 species), Cuba (4 species), Hainan (3 species) and Ceylan (2 species). 11 non-endemic species were top functionally original and had a single insular location, in particular Trinidad (3 species), Bali, Tasmania, Sicily, Japan, Equatorial Guinea, Ceylan, Corea, Taiwan (1 species).

### Dispersal abilities of original species

On average, species that dispersed on small distances were more evolutionary original than species dispersing on long-distances (Figure 4a). Endemic evolutionary original species were mainly transported by wind, animals or their mode of dispersal was unassisted. Non-endemic evolutionary original species had mainly an “unassisted” mode of dispersal. Regarding FO, species dispersing on long distances were more functionally original, especially zoochor and anemochor species



(Figure 4c and d).

### Possibility of islands to be at reach and distribution of originality

One common point to functionally and evolutionary original species is that they were located at low latitudes, indicating a spatial effect on originality (Table 4). The magnitude of the longitude effect was generally higher for FO than for EO. We observed a high number of differences between the effects of island features on FO and EO. Whatever they were endemic to islands or not, evolutionary original species tended to be found on islands that were continental, non-glaciated, flat and close to the mainland. In addition, endemic evolutionary original species occurred mainly on old islands while non-endemic evolutionary original species were found on more recent islands. The effect of the proportion of surrounding landmass was not significant on the evolutionary originality of non-endemic species and was negative in the case of endemic species but with a low significance (Table 4a). Area had little effect on evolutionary original species whether they were endemic or non-endemic to islands. Regarding FO we found important differences between endemic and non-endemic species. Endemic functionally original species tended to occur on islands that were oceanic, glaciated, small, relatively elevated, close to other landmasses but far from the continent (Table 4b). On the opposite, non-endemic functionally original species were present on islands of continental origin, close to the mainland, and which were relatively young. (Table 4b).

**Table 4:** Effects and importance of island features on Evolutionary and Functional Originality estimated from multi-model selection.

a)

Evolutionary Originality	Species found in both islands and continents					
	Species restricted to islands					
	<i>Estimate</i>	<i>Pr(&gt;z)</i>	<i>Importance</i>	<i>Estimate</i>	<i>Pr(&gt;z)</i>	<i>Importance</i>
<b>SLMP</b>	-4.2	.	0.62	-0.34		0.29
<b>Distance</b>	-2.6	*	0.89	-2.21	***	1

<b>Area</b>	0.62		0.28	-0.45		0.32
<b>Elevation</b>	-6.3	***	1	-0.31		0.29
<b>Age</b>	5.1	**	0.98	-1.5	*	0.88
<b>GMMC</b>	7.07	***	1	4.9	***	1
<b>Glaciated</b>	-5.9	***	1	-3.5	***	1
<b>Latitude</b>	-6.79	***	1	-2.6	**	1
<b>Longitude</b>	-4.25	**	0.79	0.58		0.35
<b>Latitude:SLMP</b>	7.21	***	1			
<b>Latitude:Age</b>	1.53		0.29			
<b>Distance:Elevation</b>				1.5	.	0.66
<b>Glaciated:Distance</b>				0.21		0.24

b)

<b>Functional Originality</b>	<b>Species found in both islands and continents</b>					
	<b>Species restricted to islands</b>					
	<i>Estimate</i>	<i>Pr(&gt;z)</i>	<i>Importance</i>	<i>Estimate</i>	<i>Pr(&gt;z)</i>	<i>Importance</i>
<b>SLMP</b>	0.33	***	1	-0.087	***	1
<b>Near_dist</b>	0.049	**	0.98	-0.058	***	1
<b>Area</b>	-0.083	***	1	-0.005		0.39
<b>Elevation</b>	0.029		0.42	-0.027	***	1
<b>Age</b>	NA			-0.033	***	1
<b>GMMC</b>	-0.22	***	1	0.013	*	0.85
<b>Glaciated</b>	0.081	***	1	-0.04	***	1
<b>Latitude</b>	-0.26	***	1	-0.005		0.27
<b>Longitude</b>	0.084	***	1	-0.09	***	1
<b>Latitude:Distance</b>	0.01		0.25			

<b>SLMP:Latitude</b>	-0.14	***	1
<b>SLMP:Longitude</b>			0.001
<b>SLMP:Elev</b>			0.029
		***	1

## DISCUSSION

### Decoupling of functional and evolutionary originality

In this study we drew a first picture of species originality on islands. Original species have a strong contribution to the uniqueness of island biota, to divergence with the mainland and to centers of endemism (Veron, Haevermans, et al., 2019; Veron, Mouchet, et al., 2019). Among the islands where concentrate both functionally and evolutionary original species are Indonesian islands, the Philippines, Taiwan, Hainan, New-Guinea, Japan, Sumatra, Madagascar. However, some islands harbor only EO species and others only FO species. Importantly, the sets of the most functionally and evolutionary original species were not similar and functional traits of plants representing their ecological strategies had weak phylogenetic signals. There are many possible reasons why FO and EO are not congruent (e.g., Cornwell et al., 2014; Véron, Saito, Padilla-García, Forest, & Bertheau, 2019). On islands, species representative of adaptive radiation, and consequently being closely related and having very close evolutionary originality values, may have very distinct traits in order to fill the niche space let vacant by the low species richness or/and to avoid competition. On the opposite, the process of relictualization, i.e. extinctions of close-relatives on the continent, may have artificially isolated species in the tree-of-life, among which some have diversified relatively recently and share several traits with current species (Grandcolas et al., 2014; R. J. Whittaker & Fernández-Palacios, 2007). Another reason for this decoupling is convergent evolution which has been reported in multiple lineages on multiple islands (Robert J. Whittaker et al., 2017). This is the process by which evolutionary unrelated organisms show similar features as a result of natural selection and adaptation under similar environmental constraints (Mazel et al., 2017). Although

eco-evolutionary studies would be needed to better understand the evolution of plant species traits in islands, this provided a first evidence that ecological strategies of plants are not conserved in the phylogeny (see also Cornwell et al., 2014).

### **Higher originality on islands than on continents**

A second important finding was that island endemic species were on average more functionally and evolutionary original than those occurring on continents. A possible explanation is the isolated nature of islands which allows evolution to take its own course giving birth to species evolutionary isolated from all others and/or that harbor unique forms and functions. Higher evolutionary originality on islands compared to the mainland was however unexpected. Indeed, the initial insular pool is generally a set of the continental pool, therefore one could expect that insular species would not be more ancient and evolutionary isolated than continental ones. Moreover, high speciation rates in islands may result in the diversification of many closely-related species which is expected to decrease the average originality of insular pools. Past history in continents may be one explanation of our contradictory finding of higher evolutionary originality on islands than on continents. One main process could be continental extinction: species loss on continents may have pruned the tree of life so that several insular species lost some close-relatives and consequently became evolutionary original and sometimes endemic (Gillespie & Roderick, 2002). High climatic fluctuations and extreme events in the mainland concomitant with a relative stable climate in islands drove extinctions on the continent and persistence on islands, especially in plants (Cronk, 1997; R. J. Whittaker & Fernández-Palacios, 2007 but see Jetz et al., 2014). Climate change in continents may not only have conducted to extinctions but also to diversification resulting from natural selection of the most adapted forms. The result is similar as for extinctions: the evolutionary isolation of insular forms. Besides, not only species loss on the continent but the very high number of past extinctions on islands may have conducted to the evolutionary isolation of insular forms (Courchamp, Hoffmann, Russell, Leclerc, & Bellard, 2014; Robert J. Whittaker et al., 2017). This process of isolation depends on the phylogenetic signature of extinctions in insular lineages, which

remains to be explored, but our result suggest that they have been clustered in species-poor clades which originated in evolutionary original species. A phylogenetic tree incorporating extinct species coupled to the identification of their past distribution may help to disentangle the roles of extinctions and ancient diversification on insular originality.

As for functional originality we also showed that it was higher on islands than on the mainland. Two main mechanisms may have resulted in a higher functional originality on islands: niche expansion and niche shift. Niche expansion relates to low species richness and high disharmony in islands which resulted in a diversity of vacant niches to be colonized by functionally diverse species (R. J. Whittaker & Fernández-Palacios, 2007). For example, (Diamond, 1970) reported that pacific birds could occupy higher altitudinal ranges, a higher diversity of habitats and a higher diversity of vertical foraging ranges compared to their continental counterparts after they colonized species-poor islands in the pacific. Niche shift is the changing biological and ecological characteristics of insular species compared to the mainland (MacArthur & Wilson, 1967; Simberloff, 1974). The persistence of insular species and especially plants may have been conditioned to adaptations to the oceanic climate of islands resulting in the evolution of distinct ecological strategies.

### **Range-restriction of original species**

The few studies exploring the relationship between evolutionary originality and geographic rarity showed that original species were not more range-restricted than others (Jetz et al., 2014; Thuiller et al., 2015; Stein et al., 2018). Here, although correlation between rarity and EO was generally low, we found that the most evolutionary original species were rare. Some mechanisms generating originality in insular systems also generate endemism, for example adaptation to particular environmental conditions and the function of refuge of many islands. As stated above, island biota has also suffered a number of extinctions and it is likely that human-induced extinctions on islands have both artificially isolated insular species in the Tree of Life and turned multi-endemic species into single endemics (R. J. Whittaker & Fernández-Palacios, 2007). Extinctions may therefore be both responsible for higher originality and higher range-restriction on islands than on continents

but, as we will discuss further, it is not the only factor explaining these patterns.

On average, functionally original species were found only on a small number of islands, which seems to be in accordance with studies showing that rare functions are usually supported by rare species (Kunin & Gaston, 1993; Violle et al., 2017; Grenié et al., 2018). Yet, using null models we found that range-restriction of functionally original species was not different from random distribution of species across islands, i.e. functionally original monocot species were not rarer than less original ones. This contradicts previous findings that “species that have a low functional redundancy (...) are rarer than expected by chance” (David Mouillot, Bellwood, et al., 2013; see also Kunin & Gaston, 1993). Violle et al. (2017) recently defined a multifaceted framework of “functional rarity”, linking species’ functional originality (called distinctiveness or uniqueness) and rarity. In their framework 12 forms of functional rarity are possible at two spatial scales (local and regional). The less frequent case of few common species with original traits is possible and goes with our findings. A possible explanation would be that functionally original species have a low level of competition with other species so that they can more easily settle and occupy a vacant niche when they disperse towards it.

### **Dispersal capacities of evolutionary and functionally original species**

There are many examples of islands where species distribution is due to dispersal power (R. J. Whittaker & Fernández-Palacios, 2007; Borda-de-Água et al., 2017). For example, the Krakatau island has been colonized by ferns whose spores can be transported over long-distances whereas other lineages are absent due to their lack of “ocean-going” dispersal mechanisms (R. J. Whittaker, Jones, & Partomihardjo, 1997). The modes of dispersal of original species and their association with short and long-distance dispersal strongly support our results on the spatial distribution of these species. Highly evolutionary original species are transported predominantly on short distances and their mode of dispersal is generally “unassisted”, which prevents ocean crossing. More strikingly, the stronger the association between originality and the “unassisted” mode of dispersal, the higher is the correlation between originality and rarity. A possibility explaining the

“unassisted” mode of dispersal of original species may be related to the loss of dispersal of plants occurring on islands due to a reduced herbivory pressure and lower competition. Yet, this loss of dispersal syndrome is generally tied to species having recently diversified (Robert J. Whittaker et al., 2017). Consequently, two assumptions that can be made but need to be tested are, first, that evolutionary original species have low dispersal abilities because they have diversified in recent times but became isolated through relictualization, and, second, that low dispersal ability is a relictual trait rather than a derived one in original species.

As for functionally original species, they were more widespread than non-original ones and this also goes with our findings about their mode of dispersal. Indeed, only few of them had an “unassisted” mode of dispersal. Rather they were transported through zoochory (especially birds) and anemochory and may have been able to colonize several islands isolated from each other. Additional important factors of dispersal include the direction, seasonality, strength of dispersal agents (Gillespie et al., 2012). Yet, our results show that the mode of dispersal and the distance on which seeds can be transported are key determinants of the distribution of original species on islands.

The dispersal strategies of original species can also be seen from the features of the islands where they are found. The most evolutionary original species are often single-endemics and rarely occur on islands where recent and/or long-distance dispersal would be needed to establish: they are uncommon on remote, glaciated, young and oceanic islands. Evolutionary original species also occur on islands surrounded by a relatively low proportion of landmass, decreasing the possibilities of dispersal. On the contrary, functionally original endemic species, which are more widespread, occurred on islands with opposite features and that may only be colonized by species having high dispersal abilities. Together, the dispersal abilities of original species and the characteristics of the islands where they are present are important factors explaining the rarity of plants on islands.

### **Conservation**

Originality may have an intrinsic value and its conservation is essential to preserve the breadth of

functions necessary against future uncertainty and to maintain the reservoir of future benefits to people (Faith, 1992); (Violle et al., 2017). If original species go extinct first - and previous considerations shows that this is a very likely assumption - some irreplaceable traits and evolutionary branches would be lost, strongly affecting ecosystem processes and services. Preserving EO species is essential due to their high contribution to “option-values” for humanity and the conservation of FO species is needed to maintain ecosystem functions and processes. This is all the more true on islands as, at least for monocots, the world most original species are insular. Preserving species evolutionary history has sometimes been assumed to be a proxy for the conservation of species traits. However, several reasons such as convergent evolution, the consideration of a low number of traits, a fast pace of diversification, sexual selection etc. may blur this correlation (Gerhold, Cahill, Winter, Bartish, & Prinzing, 2015; Faith, 2018; Véron, Saito, Padilla-García, Forest, & Bertheau, 2019). In our study we emphasized a lack of phylogenetic signal showing that the originality of plant ecological strategies may not be reflected by their evolutionary relationships. Therefore, a ‘silver-bullet’ strategy aiming at the protection of both species functions and evolutionary history may be inadequate. However, in spite of being different, EO and FO species tend to co-occur on many similar islands which give the opportunity for combined conservation strategies that may protect both functional and evolutionary originality. Further research could then look at the overlap between original species range on a given island which may help to prioritize areas for conservation.

A highly important consideration for conservation is that, among endemic species, the most evolutionary original ones are also the rarest. This was also true for non-endemic species but to a lesser degree. Rare and original species may cumulate a diversity of threats that make them at high risks of extinctions (Jono & Pavoine, 2012). First, rarity may be a signal of higher extinction probabilities (e.g., Harnik, Simpson, & Payne, 2012), in particular for insular endemic species for which none population may survive on the mainland. Second, species occurring on islands are among the most vulnerable on Earth due to the combined effects of human-induced threat, species



naiveté and low range-size. Third, rare and original insular species may also be highly vulnerable to external threats such as climate or land-use changes because they have low dispersal abilities and no land at proximity to escape (Veron, Mouchet, Govaerts, Haevermans, & Pellens, 2019). Finally, evolutionary original species may be at higher extinction risks due to increased extinction risks with age (Warren et al., 2018). The most functionally original species were not rarer (geographic rarity) than other species, but they were distributed in a low number of islands. In addition, the rarest functionally and evolutionary original species generally did not occur on the same islands and regional or local conservation plans may not allow to protect rare species which are both distinct in their traits and evolutionary history. Understanding the diversity of current threats to original species and how they will respond to these pressures is key to assess their extinction status and to propose actions directed toward the preservation of these unique species (Forest et al., 2018), (Stein et al., 2018).

Finally, as a bridge between the conservation of biodiversity and human cultural heritage, and to conclude with originality, islands are not only places where the most original and threatened species occur but they also concentrate the most original and threatened human languages (Perrault, Farrell, & Davies, 2017).

## REFERENCES

- Bland, L. M., Collen, B., Orme, C. D. L., & Bielby, J. (2015). Predicting the conservation status of data-deficient species: Predicting Extinction Risk. *Conservation Biology*, *29*(1), 250–259. <https://doi.org/10.1111/cobi.12372>
- Borda-de-Água, L., Whittaker, R. J., Cardoso, P., Rigal, F., Santos, A. M. C., Amorim, I. R., ... Borges, P. A. V. (2017). Dispersal ability determines the scaling properties of species abundance distributions: a case study using arthropods from the Azores. *Scientific Reports*, *7*(1), 3899. <https://doi.org/10.1038/s41598-017-04126-5>
- Bracken, M. E. S., & Low, N. H. N. (2012). Realistic losses of rare species disproportionately impact higher trophic levels: Loss of rare ‘cornerstone’ species. *Ecology Letters*, *15*(5), 461–467. <https://doi.org/10.1111/j.1461-0248.2012.01758.x>
- Cadotte, M. W., Cardinale, B. J., & Oakley, T. H. (2008). Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences*, *105*(44), 17012–17017. <https://doi.org/10.1073/pnas.0805962105>
- Cadotte, Marc W., & Jonathan Davies, T. (2010). Rarest of the rare: advances in combining evolutionary distinctiveness and scarcity to inform conservation at biogeographical scales: Conservation phylo-biogeography. *Diversity and Distributions*, *16*(3), 376–385. <https://doi.org/10.1111/j.1472-4642.2010.00650.x>
- Carvajal-Endara, S., Hendry, A. P., Emery, N. C., & Davies, T. J. (2017). Habitat filtering not dispersal limitation shapes oceanic island floras: species assembly of the Galápagos archipelago. *Ecology Letters*, *20*(4), 495–504. <https://doi.org/10.1111/ele.12753>
- Cornwell, W. K., Westoby, M., Falster, D. S., FitzJohn, R. G., O’Meara, B. C., Pennell, M. W., ... Zanne, A. E. (2014). Functional distinctiveness of major plant lineages. *Journal of Ecology*, *102*(2), 345–356. <https://doi.org/10.1111/1365-2745.12208>
- Courchamp, F., Hoffmann, B. D., Russell, J. C., Leclerc, C., & Bellard, C. (2014). Climate change, sea-level rise, and conservation: keeping island biodiversity afloat. *Trends in Ecology & Evolution*, *29*(12), 179–186. <https://doi.org/10.1016/j.tree.2014.09.008>

- Evolution*, 29(3), 127–130. <https://doi.org/10.1016/j.tree.2014.01.001>
- Cronk, Q. C. B. (1997). Islands: stability, diversity, conservation. *Biodiversity & Conservation*, 6(3), 477–493. <https://doi.org/10.1023/A:1018372910025>
- Depraetere, C., & Dall, A. (2007). *IBPoW Database. A technical note on a global dataset of islands* (p. 58).
- Diamond, J. M. (1970). Ecological Consequences of Island Colonization by Southwest Pacific Birds, I. Types of Niche Shifts. *Proceedings of the National Academy of Sciences*, 67(2), 529–536. <https://doi.org/10.1073/pnas.67.2.529>
- Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., Dray, S., ... Gorné, L. D. (2015). The global spectrum of plant form and function. *Nature*, 529, 167.
- Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61(1), 1–10. [https://doi.org/10.1016/0006-3207\(92\)91201-3](https://doi.org/10.1016/0006-3207(92)91201-3)
- Faith, D. P. (2018). Phylogenetic Diversity and Conservation Evaluation: Perspectives on Multiple Values, Indices, and Scales of Application. In R. A. Scherson & D. P. Faith (Eds.), *Phylogenetic Diversity: Applications and Challenges in Biodiversity Science* (pp. 1–26). [https://doi.org/10.1007/978-3-319-93145-6\\_1](https://doi.org/10.1007/978-3-319-93145-6_1)
- Fonseca, C. R., & Ganade, G. (2001). Species functional redundancy, random extinctions and the stability of ecosystems. *Journal of Ecology*, 89(1), 118–125. <https://doi.org/10.1046/j.1365-2745.2001.00528.x>
- Forest, F., Moat, J., Baloch, E., Brummitt, N. A., Bachman, S. P., Ickert-Bond, S., ... Buerki, S. (2018). Gymnosperms on the EDGE. *Scientific Reports*, 8(1), 6053. <https://doi.org/10.1038/s41598-018-24365-4>
- 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). *Functional Ecology*, 29(5), 600–614. <https://doi.org/10.1111/1365-2435.12425>
- Gillespie, R. G., Baldwin, B. G., Waters, J. M., Fraser, C. I., Nikula, R., & Roderick, G. K. (2012). Long-distance dispersal: a framework for hypothesis testing. *Trends in Ecology & Evolution*,

- 27(1), 47–56. <https://doi.org/10.1016/j.tree.2011.08.009>
- Gillespie, R. G., & Roderick, G. K. (2002). ARTHROPODS ON ISLANDS: Colonization, Speciation, and Conservation. *Annual Review of Entomology*, 47(1), 595–632. <https://doi.org/10.1146/annurev.ento.47.091201.145244>
- Grandcolas, P., Nattier, R., & Trewick, S. (2014). Relict species: a relict concept? *Trends in Ecology & Evolution*, 29(12), 655–663. <https://doi.org/10.1016/j.tree.2014.10.002>
- Grenié, M., Mouillot, D., Villéger, S., Denelle, P., Tucker, C. M., Munoz, F., & Violle, C. (2018). Functional rarity of coral reef fishes at the global scale: Hotspots and challenges for conservation. *Biological Conservation*, 226, 288–299. <https://doi.org/10.1016/j.biocon.2018.08.011>
- Harnik, P. G., Simpson, C., & Payne, J. L. (2012). Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749), 4969–4976. <https://doi.org/10.1098/rspb.2012.1902>
- 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. *Current Biology*, 24(9), 919–930. <https://doi.org/10.1016/j.cub.2014.03.011>
- Jono, C. M. A., & Pavoine, S. (2012). Threat Diversity Will Erode Mammalian Phylogenetic Diversity in the Near Future. *PLoS ONE*, 7(9), e46235. <https://doi.org/10.1371/journal.pone.0046235>
- Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibisich, P. L., Nowicki, C., ... Barthlott, W. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences*, 106(23), 9322–9327. <https://doi.org/10.1073/pnas.0810306106>
- Knope, M. L., Morden, C. W., Funk, V. A., & Fukami, T. (2012). Area and the rapid radiation of Hawaiian *Bidens* (Asteraceae): Rapid radiation of Hawaiian *Bidens*. *Journal of Biogeography*, 39(7), 1206–1216. <https://doi.org/10.1111/j.1365-2699.2012.02687.x>

- Kunin, W. E., & Gaston, K. J. (1993). The biology of rarity: Patterns, causes and consequences. *Trends in Ecology & Evolution*, 8(8), 298–301. [https://doi.org/10.1016/0169-5347\(93\)90259-R](https://doi.org/10.1016/0169-5347(93)90259-R)
- Legendre, P., & Legendre, L. (2012). *Numerical ecology* (Third English edition). Amsterdam: Elsevier.
- Leitão, R. P., Zuanon, J., Villéger, S., Williams, S. E., Baraloto, C., Fortunel, C., ... Mouillot, D. (2016). Rare species contribute disproportionately to the functional structure of species assemblages. *Proceedings of the Royal Society B: Biological Sciences*, 283(1828), 20160084. <https://doi.org/10.1098/rspb.2016.0084>
- Losos, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, 11(10), 995–1003. <https://doi.org/10.1111/j.1461-0248.2008.01229.x>
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton, NJ: Princeton University Press.
- Mazel, F., Wüest, R. O., Gueguen, M., Renaud, J., Ficetola, G. F., Lavergne, S., & Thuiller, W. (2017). The Geography of Ecological Niche Evolution in Mammals. *Current Biology*, 27(9), 1369–1374. <https://doi.org/10.1016/j.cub.2017.03.046>
- Mouchet, M., Guilhaumon, F., Villéger, S., Mason, N. W. H., Tomasini, J.-A., & Mouillot, D. (2008). Towards a consensus for calculating dendrogram-based functional diversity indices. *Oikos*, 117(5), 794–800. <https://doi.org/10.1111/j.0030-1299.2008.16594.x>
- Mouillot, D., Culioli, J. M., Pelletier, D., & Tomasini, J. A. (2008). Do we protect biological originality in protected areas? A new index and an application to the Bonifacio Strait Natural Reserve. *Biological Conservation*, 141(6), 1569–1580. <https://doi.org/10.1016/j.biocon.2008.04.002>
- Mouillot, David, Bellwood, D. R., Baraloto, C., Chave, J., Galzin, R., Harmelin-Vivien, M., ... Thuiller, W. (2013). Rare Species Support Vulnerable Functions in High-Diversity Ecosystems. *PLoS Biology*, 11(5), e1001569. <https://doi.org/10.1371/journal.pbio.1001569>

- Mouillot, David, Graham, N. A. J., Villéger, S., Mason, N. W. H., & Bellwood, D. R. (2013). A functional approach reveals community responses to disturbances. *Trends in Ecology & Evolution*, *28*(3), 167–177. <https://doi.org/10.1016/j.tree.2012.10.004>
- Oliver, T. H., Heard, M. S., Isaac, N. J. B., Roy, D. B., Procter, D., Eigenbrod, F., ... Bullock, J. M. (2015). Biodiversity and Resilience of Ecosystem Functions. *Trends in Ecology & Evolution*, *30*(11), 673–684. <https://doi.org/10.1016/j.tree.2015.08.009>
- Pavoine, S. (2019). *Package "adiv."*
- Pavoine, Sandrine, Bonsall, M. B., Dupaix, A., Jacob, U., & Ricotta, C. (2017). From phylogenetic to functional originality: Guide through indices and new developments. *Ecological Indicators*, *82*, 196–205. <https://doi.org/10.1016/j.ecolind.2017.06.056>
- Pavoine, Sandrine, Ollier, S., & Dufour, A.-B. (2005). Is the originality of a species measurable? Originality of a species. *Ecology Letters*, *8*(6), 579–586. <https://doi.org/10.1111/j.1461-0248.2005.00752.x>
- Pendleton, R. M., Hoinghaus, D. J., Gomes, L. C., & Agostinho, A. A. (2014). Loss of Rare Fish Species from Tropical Floodplain Food Webs Affects Community Structure and Ecosystem Multifunctionality in a Mesocosm Experiment. *PLoS ONE*, *9*(1), e84568. <https://doi.org/10.1371/journal.pone.0084568>
- Perrault, N., Farrell, M. J., & Davies, T. J. (2017). Tongues on the EDGE: language preservation priorities based on threat and lexical distinctiveness. *Royal Society Open Science*, *4*(12), 171218. <https://doi.org/10.1098/rsos.171218>
- Petchey, O. L., Hector, A., & Gaston, K. J. (2004). HOW DO DIFFERENT MEASURES OF FUNCTIONAL DIVERSITY PERFORM? *Ecology*, *85*(3), 847–857. <https://doi.org/10.1890/03-0226>
- Pollock, L. J., Thuiller, W., & Jetz, W. (2017). Large conservation gains possible for global biodiversity facets. *Nature*, *546*(7656), 141–144. <https://doi.org/10.1038/nature22368>
- Qian, H., & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant

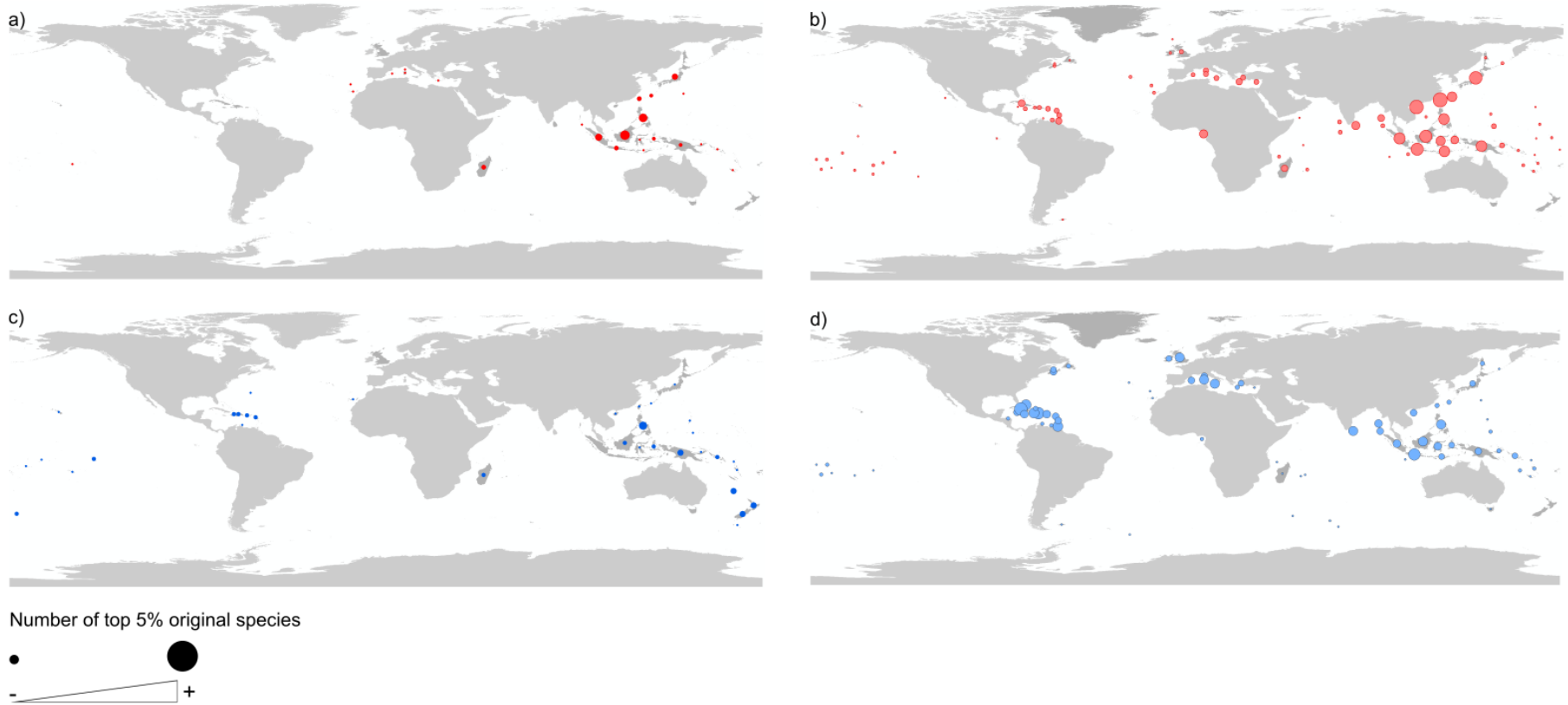
- phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology*, 9(2), 233–239. <https://doi.org/10.1093/jpe/rtv047>
- Redding, D. W., Hartmann, K., Mimoto, A., Bokal, D., DeVos, M., & Mooers, A. Ø. (2008). Evolutionarily distinctive species often capture more phylogenetic diversity than expected. *Journal of Theoretical Biology*, 251(4), 606–615. <https://doi.org/10.1016/j.jtbi.2007.12.006>
- Redding, D. W., Mazel, F., & Mooers, A. Ø. (2014). Measuring Evolutionary Isolation for Conservation. *PLoS ONE*, 9(12), e113490. <https://doi.org/10.1371/journal.pone.0113490>
- Revell, L. J. (2010). Phylogenetic signal and linear regression on species data: Phylogenetic regression. *Methods in Ecology and Evolution*, 1(4), 319–329. <https://doi.org/10.1111/j.2041-210X.2010.00044.x>
- Rodrigues, A., Pilgrim, J., Lamoreux, J., Hoffmann, M., & Brooks, T. (2006). The value of the IUCN Red List for conservation. *Trends in Ecology & Evolution*, 21(2), 71–76. <https://doi.org/10.1016/j.tree.2005.10.010>
- Rosauer, D., 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. *Molecular Ecology*, 18(19), 4061–4072. <https://doi.org/10.1111/j.1365-294X.2009.04311.x>
- Simberloff, D. S. (1974). Equilibrium Theory of Island Biogeography and Ecology. *Annual Review of Ecology and Systematics*, 5(1), 161–182. <https://doi.org/10.1146/annurev.es.05.110174.001113>
- Stein, R. W., Mull, C. G., Kuhn, T. S., Aschliman, N. C., Davidson, L. N. K., Joy, J. B., ... Mooers, A. O. (2018). Global priorities for conserving the evolutionary history of sharks, rays and chimaeras. *Nature Ecology & Evolution*, 2(2), 288–298. <https://doi.org/10.1038/s41559-017-0448-4>
- Thuiller, W., Maiorano, L., Mazel, F., Guilhaumon, F., Ficetola, G. F., Lavergne, S., ... Mouillot, D. (2015). Conserving the functional and phylogenetic trees of life of European tetrapods.

- Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1662), 20140005–20140005. <https://doi.org/10.1098/rstb.2014.0005>
- UNEP-WCM. (2013). *Global islands database*. United Nation's Environmental Program, Cambridge, United Kingdom.
- Veron, S., Haevermans, T., Govaerts, R., Mouchet, M., & Pellens, R. (2019). Distribution and relative age of endemism across islands worldwide. *Scientific Reports*, 9(1), 11693. <https://doi.org/10.1038/s41598-019-47951-6>
- Veron, S., Mouchet, M., Govaerts, R., Haevermans, T., & Pellens, R. (2019). Vulnerability to climate change of islands worldwide and its impact on the tree of life. *Scientific Reports*, 9(1), 14471. <https://doi.org/10.1038/s41598-019-51107-x>
- Veron, S., Mouchet, M., Grandcolas, P., Govaerts, R., Haevermans, T., & Pellens, R. (2019). *Tracking the origin of island diversity: insights from divergence with the continental pool in monocots* [Preprint]. <https://doi.org/10.1101/678300>
- Veron, S., Penone, C., Clergeau, P., Costa, G. C., Oliveira, B. F., São-Pedro, V. A., & Pavoine, S. (2016). Integrating data-deficient species in analyses of evolutionary history loss. *Ecology and Evolution*, 6(23), 8502–8514. <https://doi.org/10.1002/ece3.2390>
- Véron, S., Saito, V., Padilla-García, N., Forest, F., & Bertheau, Y. (2019). The Use of Phylogenetic Diversity in Conservation Biology and Community Ecology: A Common Base but Different Approaches. *The Quarterly Review of Biology*, 94(2), 123–148. <https://doi.org/10.1086/703580>
- Violle, C., Thuiller, W., Mouquet, N., Munoz, F., Kraft, N. J. B., Cadotte, M. W., ... Mouillot, D. (2017). Functional Rarity: The Ecology of Outliers. *Trends in Ecology & Evolution*, 32(5), 356–367. <https://doi.org/10.1016/j.tree.2017.02.002>
- Warren, B. H., Hagen, O., Gerber, F., Thébaud, C., Paradis, E., & Conti, E. (2018). Evaluating alternative explanations for an association of extinction risk and evolutionary uniqueness in multiple insular lineages: EXTINCTION AND EVOLUTIONARY UNIQUENESS.



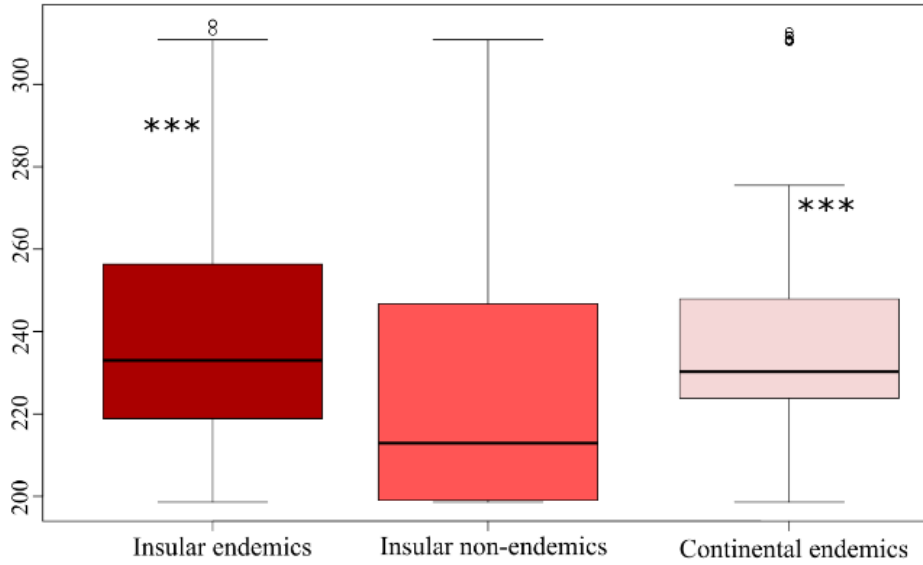
*Evolution*, 72(10), 2005–2024. <https://doi.org/10.1111/evo.13582>

- Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., ... Thébaud, C. (2015). Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters*, 18(2), 200–217. <https://doi.org/10.1111/ele.12398>
- Weigelt, P., Jetz, W., & Kreft, H. (2013). Bioclimatic and physical characterization of the world's islands. *Proceedings of the National Academy of Sciences*, 110(38), 15307–15312. <https://doi.org/10.1073/pnas.1306309110>
- Weigelt, Patrick, Daniel Kissling, W., Kisel, Y., Fritz, S. A., Karger, D. N., Kessler, M., ... Kreft, H. (2015). Global patterns and drivers of phylogenetic structure in island floras. *Scientific Reports*, 5(1). <https://doi.org/10.1038/srep12213>
- Whittaker, R. J., & Fernández-Palacios, J. M. (2007). *Island biogeography: ecology, evolution, and conservation* (Oxford University Press).
- Whittaker, R. J., Jones, S. H., & Partomihardjo, T. (1997). The rebuilding of an isolated rain forest assemblage: how disharmonic is the flora of Krakatau?. *Biodiversity and Conservation*, 6, 1671–1696.
- Whittaker, Robert J., Fernández-Palacios, J. M., Matthews, T. J., Borregaard, M. K., & Triantis, K. A. (2017). Island biogeography: Taking the long view of nature's laboratories. *Science*, 357(6354), eaam8326. <https://doi.org/10.1126/science.aam8326>
- Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., ... Beaulieu, J. M. (2013). Three keys to the radiation of angiosperms into freezing environments. *Nature*, 506, 89.

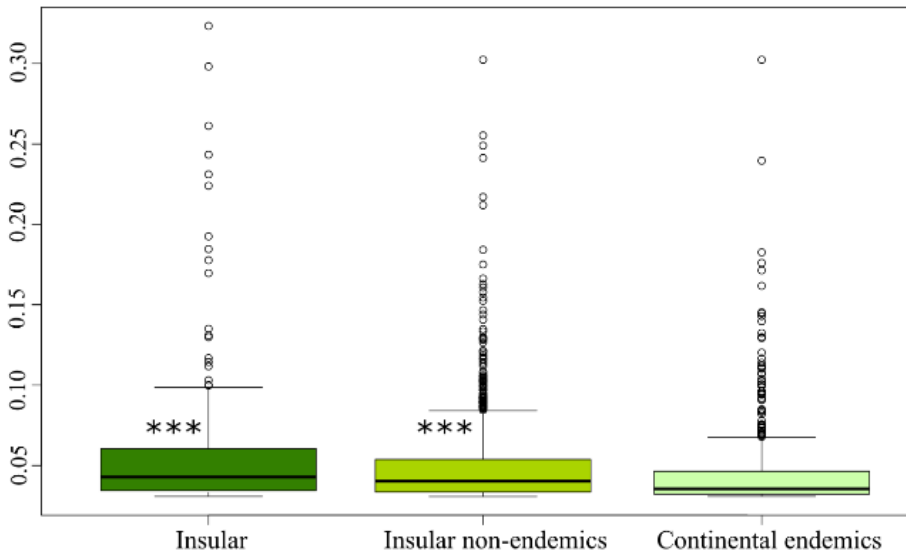


**Figure 1:** Number of top original species per island for a) Evolutionary Originality (EO) of endemic insular species, b) EO of non-endemic insular species, c) Functional Originality (FO) of endemic insular species, d) FO of non-endemic insular species.

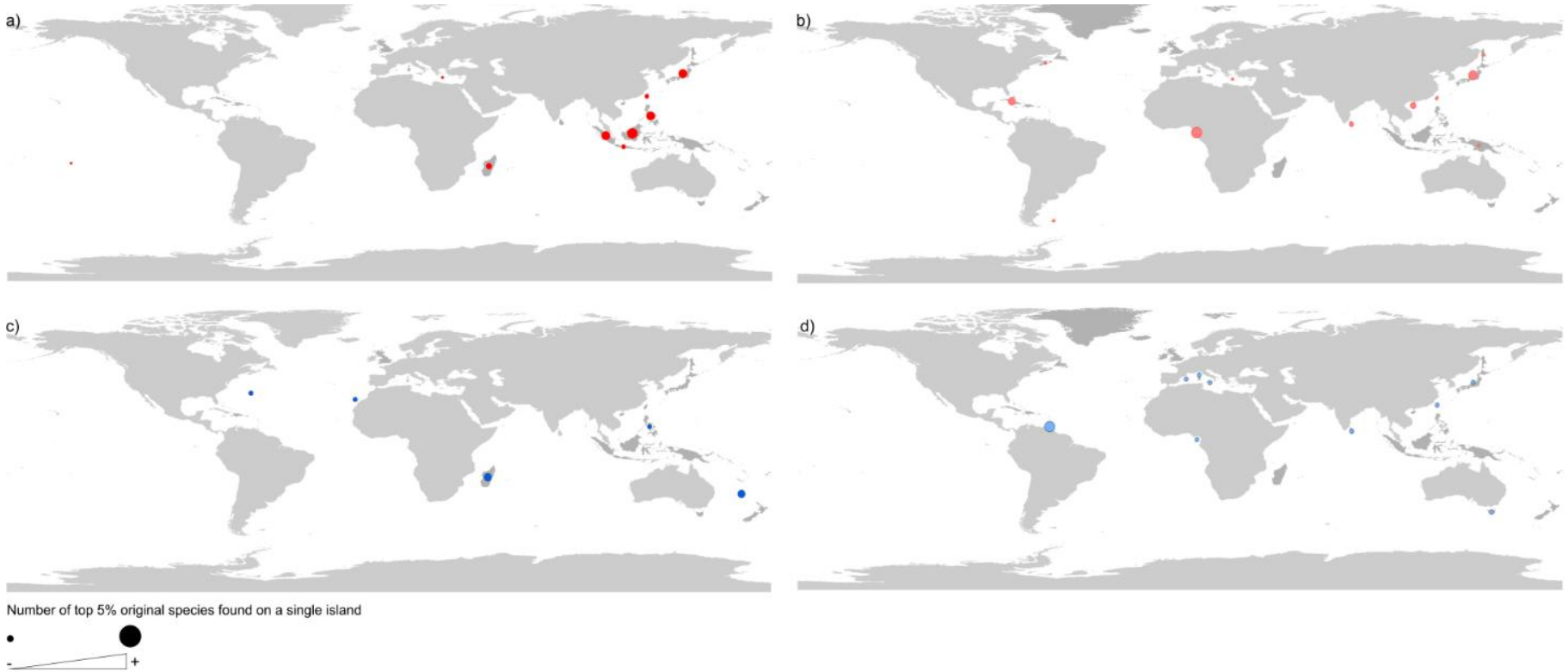
a) Evolutionary originality



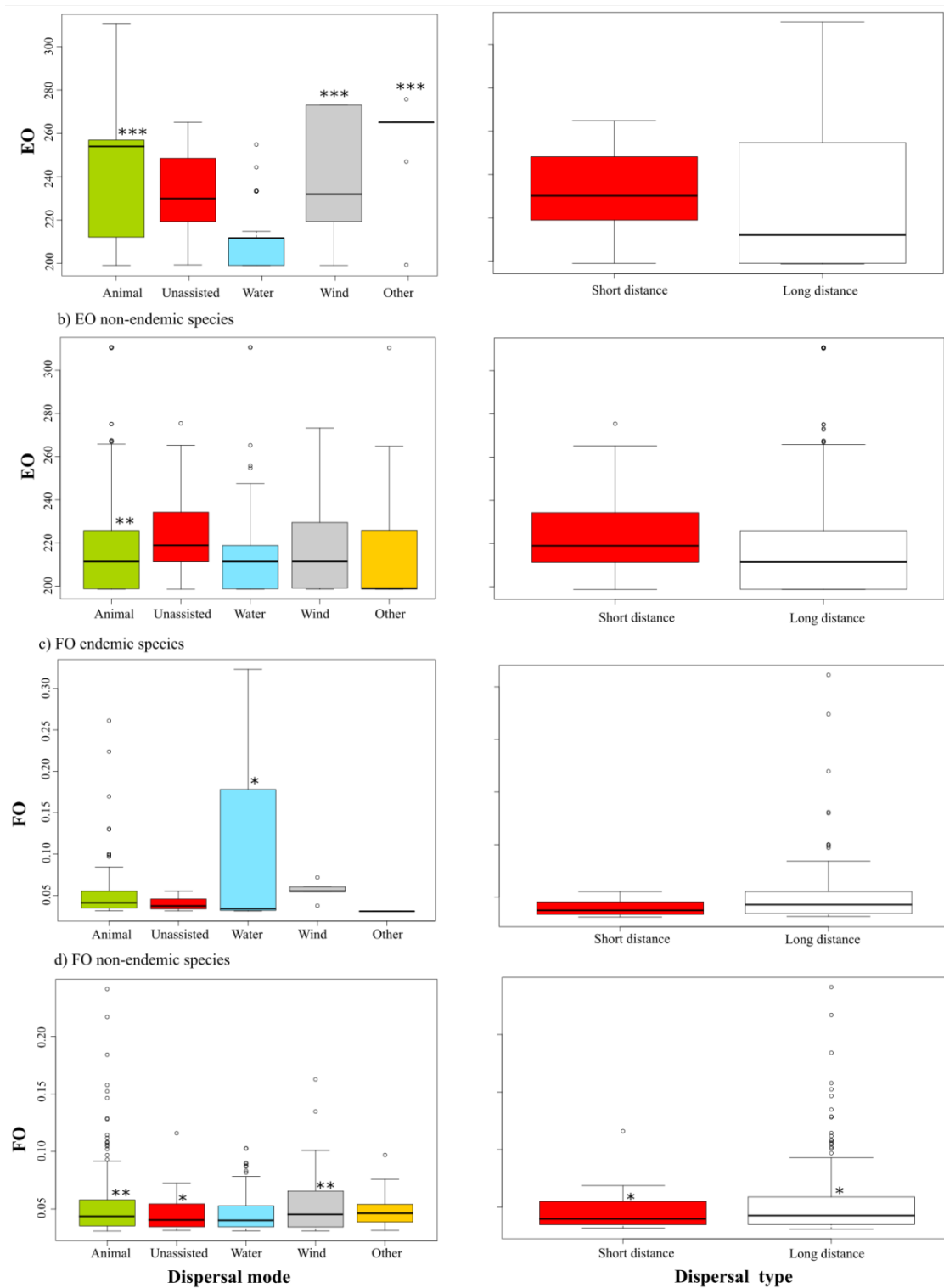
b) Functional originality



**Figure 2:** Species originality depending on its geographical origin. \*\*\* significance of a p-value < 0.001



**Figure 4:** Number of top 5% original species found on a single island for a) Evolutionary Originality (EO) of endemic insular species, b) EO of non-endemic insular species, c) Functional Originality (FO) of endemic insular species, d) FO of non-endemic insular species.



**Figure 4:** Originality scores depending on the dispersal mode and type. Significance level of p-value: \*\*\* <math>< 0.001</math>; \*\* <math>< 0.05</math>; \* <math>< 0.1</math>

## Appendix 1: Sensitivity analysis

### a) Missing trait values

#### i) Correlation between functional originality scores calculated when missing values were imputed or not

**Table 1.** Proportion of missing values per trait

Trait	Proportion of missing values
<i>Adult plant height</i>	46%
<i>Leaf size</i>	68%
<i>Stem specific density</i>	81%
<i>Leaf mass per area</i>	68%
<i>Leaf nitrogen content</i>	75%
<i>Diaspore mass</i>	23%

Missing values were imputed by performing random forest algorithm species (missForest package, Stekhoven and Bühlmann, 2012) a thousand times, and the mean of the 1000 imputations was estimated. To show how this imputation method may affect our originality results we tested the correlation between functional originality scores computed with trait data when missing values were considered as « Non-Attributed » and functional originality scores computed when missing values were imputed. We found that the correlation was positive but moderate (Spearman correlation coefficient  $r = 0.65$ ,  $p\text{-value} < 0.001$ ). This result was expected, because random forest method imputes missing values for a species according to the non-missing values of other traits.

#### ii) Correlation between the number of missing values and functional originality score

Second, we tested the correlation between the number of species' missing trait data and its functional originality score. We found that the number of species' missing trait values had a poor

influence on its functional originality score (Spearman’s correlation coefficient  $r = -0.14$ ,  $p$ -value  $< 0.001$ ). This contradicts the expectancy that random forest imputations would smooth trait values and therefore conduct to an under-estimation of the functional originality scores of species with many missing trait values.

iii) Correlation between functional originality scores calculated with a complete trait data and when missing trait values were created artificially

Third, we artificially created missing values for species with complete data for all traits (259 species). We then imputed these artificially missing values with a random forest method. We then calculated functional originality scores on the complete set of species and estimated their correlation with the functional originality scores coming from artificially created trait data. First, all values of a given trait were considered as missing. Second, we randomly created 2/3 of missing values among all traits and repeated the randomization a hundred times. Results showed high correlations between functional originality scores computed on a complete dataset and obtained after artificial deletion and imputation of trait values (Table 2).

**Table 2.** Correlation between observed functional originality scores of species with no missing values and functional originality scores obtained after artificially creating and imputing missing values in a particular trait or in all traits at random.

	Leaf area	Leaf Nitrogen content	Leaf Mass per area	Plant height	Diaspore mass	Stem specific density	2/3 of missing values randomly distributed
Pearson coefficient	0.999	0.907	0.993	0.941	0.999	0.879	0.41

iv) Relationship between functional and evolutionary originality scores in the set of species having no missing values

We calculated functional and evolutionary originality scores on the subset of 259 species for which all trait values were available. We found that the correlation estimated on this particular species subset was low (Spearman’s correlation coefficients  $r = 0.29$ ,  $p$ -values  $< 0.001$ ).

We finally estimated the phylogenetic signal of each trait from complete trait data subset using a Kstar method (Blomberg et al., 2003) in *adiv* package (Pavoine 2019). At the exception of the trait “Diaspore mass” none trait displayed a phylogenetic signal.

**Table 3.** Phylogenetic signal of each trait under study calculated with Kstar method.

	Leaf area	Leaf Nitrogen content	Leaf Mass per area	Plant height	Diaspore mass	Stem specific density
Kstar	0.093	0.0018	0.003	0.112	2.05	0.0036
	p=0.001	p=0.743	p=0.605	p=0.006	p=0.001	p=0.396

### b) Correlations between traits

**Table 4.** Pearson’s correlation coefficient between trait values

	<i>Leaf.Area</i>	<i>N.mass</i>	<i>LMA</i>	<i>Plant.Height</i>	<i>Diaspore.Mass (mg)</i>	<i>SSD.combined</i>
<i>Leaf.Area</i>	1	0.034	0.49	0.44	0.09	0.22
<i>N.mass</i>	0.034	1	0.47	0.039	0.001	0.65
<i>LMA</i>	0.49	0.47	1	0.24	0.081	0.43
<i>Plant.Height</i>	0.44	0.039	0.24	1	0.33	0.51
<i>Diaspore.Mass (mg)</i>	0.09	0.0013	0.08	0.33	1	0.16
<i>SSD.combined</i>	0.22	0.65	0.43	0.51	0.16	1

### c) Contribution of each trait to the functional originality score

To test the influence of each trait on species originality we performed a sensitivity analysis by excluding each trait separately and recalculating functional originality with the five remaining traits. Pearson’s correlation coefficient between functional originality calculated on 6 traits and calculated on 5 traits ranged from 0.91 to 0.99. We concluded that none of traits prevailed over others in the calculation of species functional originality.

**Table 5:** Correlations between the functional originality score calculated with all traits and obtained when a given trait was excluded. \*\*\* significance level for p-value < 0.001.



*Pearson's correlation with all-trait*

<i>Trait excluded</i>	<i>functional originality</i>
<i>Adult plant height</i>	0.98***
<i>Leaf size</i>	0.95***
<i>stem specific density</i>	0.98***
<i>Leaf mass per area</i>	0.91***
<i>Leaf nitrogen content</i>	0.99***
<i>Diaspore mass</i>	0.96***

**d) Sensitivity to the metric used to calculate functional originality**

To test the effect of choice of the originality metric we calculated functional originality with Fair Proportion index (FP, Isaac, 2007) on species differences represented by a functional dendrogram (hierarchical clustering method h-clust) and found a significant but moderate correlation with functional originality calculated with AVerage index (AV, Pavoine et al., 2017) on species trait-based dissimilarity matrix. We also calculated evolutionary originality with the same indices (FP and AV) respectively on phylogenetic tree and phylogenetic distances in dissimilarity matrix (Pearson's correlation for evolutionary originality  $r = 0.45$ ,  $p$ -value  $< 0.001$ ; for functional originality  $r = 0.40$ ,  $p$ -value  $< 0.001$ ).

## References for Appendix 1 in Veron et al., in prep.

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.

Isaac, N. J. B., Turvey, S. T., Collen, B., Waterman, C., and Baillie, J. E. M. (2007). Mammals on the EDGE: Conservation priorities based on threat and phylogeny. *PLoS One* 2, e296. doi:10.1371/journal.pone.0000296

Pavoine, S. (2019). *adiv: Analysis of Diversity*. R package version 1.3. <https://CRAN.R-project.org/package=adiv>

Pavoine, S., Bonsall, M. B., Dupaix, A., Jacob, U., and Ricotta, C. (2017). From phylogenetic to functional originality: Guide through indices and new developments. *Ecol. Indic.* 82, 196–205. doi:10.1016/j.ecolind.2017.06.056.

Stekhoven, D.J., Bühlmann, P. MissForest—non-parametric missing value imputation for mixed-type data, *Bioinformatics*, Volume 28, Issue 1, 1 January 2012, Pages 112–118, <https://doi.org/10.1093/bioinformatics/btr597>

## Appendix 2 : Phylogenetically corrected models

### 1. *Species endemic to islands are more original*

Method: Phylogenetic Analysis of variance (package phytools, Revell, 2018)

#### 1.1 *Evolutionary Originality*

#### ANOVA table: Phylogenetic ANOVA

Pairwise posthoc test using method = "holm"

Pairwise t-values:

	Insular endemics	Continental endemics	Insular non-endemics
Insular endemics		2.14	12.72
Continental endemics	-2.14		16.26
Insular non-endemics	-12.72	-16.26	

Pairwise corrected P-values:

	Insular endemics	Continental endemics	Insular non-endemics
Insular endemics		0.38	0.003
Continental endemics	0.38		0.003
Insular non-endemics	0.003	0.003	

#### 1.2 *Functional Originality*

Pairwise t-values:

	Insular endemics	Continental endemics	Insular non-endemics
Insular endemics		7.95	6.24
Continental endemics	-7.95		-3.9
Insular non-endemics	-6.24	3.9	

Pairwise corrected P-values:

	Insular endemics	Continental endemics	Insular non-endemics
Insular endemics		0.012	0.07
Continental endemics	0.012		0.647
Insular non-endemics	0.07	0.647	

### 2. *Original species are more range-restricted*

#### 2.1 Relationship rarity and originality

Method: Generalized least squares under phylogenetic constraint

##### 2.1.1 Insular endemics

##### 2.1.1.1 Evolutionary Originality

AIC	BIC	logLik
2094	2107	-1044.15

	Value	Std.Err	t-value	P_val
Intercept	277.13	1.49	169.23	0.00

<b>Number of islands</b>	-0.00038	0.008	-0.04	0.97
--------------------------	----------	-------	-------	------

2.1.1.2 Functional Originality

	<b>Value</b>	<b>Std.Err</b>	<b>t-value</b>	<b>P_val</b>
<b>Intercept</b>	277.13	1.49	169.23	0.00
<b>Number of islands</b>	-0.00038	0.008	-0.04	0.97

2.1.2 Insular non-endemics

2.1.2.1 Evolutionary originality

	<b>Value</b>	<b>Std.Err</b>	<b>t-value</b>	<b>P_val</b>
<b>Intercept</b>	0.000	0.0002	-0.0025	0.99
<b>Number of islands</b>	-0.00024	0.007	-0.03	0.97

2.1.2.2 Functional originality

	<b>Value</b>	<b>Std.Err</b>	<b>t-value</b>	<b>P_val</b>
<b>Intercept</b>	0.023	0.084	0.27	0.78
<b>Number of islands</b>	0.014	0.0012	11.48	0.0

2.2 Rarity depending on the category of evolutionary originality

Method: Phylogenetic Analysis of variance

2.2.1 Insular endemics

2.2.1.1 Evolutionary Originality

Pairwise posthoc test using method = "holm"

Pairwise t-values:

	High originality	Low originality	Moderate originality	Very high originality	Very low originality
High originality	0.0	-1.32	-1.69	0.72	-0.08
Low originality	1.32	0.0	-0.39	1.65	1.39
Moderate originality	1.68	0.39	0.0	1.89	1.80
Very high originality	-0.72	-1.65	-1.89	0.0	-0.81
Very low originality	0.08	-1.39	-1.80	0.81	0.0

lity					
------	--	--	--	--	--

Pairwise corrected P-values:

	High originality	Low originality	Moderate originality	Very high originality	Very low originality
High originality	1	1	1	1	1
Low originality	1	1	1	1	1
Moderate originality	1	1	1	1	1
Very high originality	1	1	1	1	1
Very low originality	1	1	1	1	1

### 2.2.1.2 Functional Originality

	High originality	Low originality	Moderate originality	Very high originality	Very low originality
High originality		1.18	1.50	-2.56	1.06
Low originality	-1.18		0.32	-3.48	-0.06
Moderate originality	-1.50	-0.32		-3.7	0.37
Very high originality	2.56	3.48	3.74		3.32
Very low originality	-1.06	0.061	0.37	-3.31	

Pairwise corrected P-values:

	High originality	Low originality	Moderate originality	Very high originality	Very low originality
High originality	1	1	1	1	1
Low originality	1	1	1	1	1
Moderate originality	1	1	1	1	1
Very high originality	1	1	1	1	1
Very low originality	1	1	1	1	1

## 2.2.2 Insular non-endemics

### 2.2.2.1 Evolutionary Originality

Pairwise posthoc test using method = "holm"

Pairwise t-values:

	High originality	Low originality	Moderate originality	Very high originality	Very low originality
High originality	0.0	-7.6	-4.24	-2.14	-6.22
Low originality	7.60	0.0	3.72	2.24	1.42
Moderate originality	4.24	-3.72	0.0	0.28	-2.24
Very high originality	2.14	-2.28	-0.28	0.0	-1.49

ality					
Very low originality	6.22	-1.44	2.24	1.49	0.0

Pairwise corrected P-values:

	High originality	Low originality	Moderate originality	Very high originality	Very low originality
High originality	1	1	1	1	1
Low originality	1	1	1	1	1
Moderate originality	1	1	1	1	1
Very high originality	1	1	1	1	1
Very low originality	1	1	1	1	1

### 2.2.2.1 Functional Originality

Pairwise t-values:

	High originality	Low originality	Moderate originality	Very high originality	Very low originality
High originality		0.15	-1.7	1.25	-0.41
Low originality	-0.15		-1.95	1.19	0.57
Moderate originality	1.69	1.95		2.18	1.22
Very high originality	-1.25	-1.19	-2.18		-1.47
Very low originality	0.41	0.58	-1.22	1.48	

Pairwise corrected P-values:

	High originality	Low originality	Moderate originality	Very high originality	Very low originality
High originality	1	1	1	1	1
Low originality	1	1	1	1	1
Moderate originality	1	1	1	1	1
Very high originality	1	1	1	1	1
Very low originality	1	1	1	1	1

## 3. *Dispersal mode*

Method: Phylogenetic Analysis of variance

### 3.1 *Type of dispersal*

#### 3.1.1 Insular endemics

##### 3.1.1.1 Evolutionary originality

Pairwise t-values

	<i>Short distance</i>	<i>Long distance</i>
--	-----------------------	----------------------

<i>Short distance</i>		-0.47
<i>Long distance</i>	0.47	

Pairwise corrected P-values:

	<i>Short distance</i>	<i>Long distance</i>
<i>Short distance</i>		0.76
<i>Long distance</i>	0.76	

### 3.1.1.2 Functional originality

	<i>Short distance</i>	<i>Long distance</i>
<i>Short distance</i>		-.115
<i>Long distance</i>	1.15	

Pairwise corrected P-values:

	<i>Short distance</i>	<i>Long distance</i>
<i>Short distance</i>		0.31
<i>Long distance</i>	0.31	

## 3.1.2 Insular non-endemics

### 3.1.2.1 Evolutionary Originality

Pairwise t-values

	<i>Short distance</i>	<i>Long distance</i>
<i>Short distance</i>		0.64
<i>Long distance</i>	-0.64	

Pairwise corrected P-values:

	<i>Short distance</i>	<i>Long distance</i>
<i>Short distance</i>		0.84
<i>Long distance</i>	0.84	

### 3.1.2.2 Functional Originality

Pairwise t-values

	<i>Short distance</i>	<i>Long distance</i>
<i>Short distance</i>		-1.44
<i>Long distance</i>	1.44	

Pairwise corrected P-values:

	<i>Short distance</i>	<i>Long distance</i>
<i>Short distance</i>		0.41

<i>Long distance</i>	0.41	
----------------------	------	--

### 3.2 Mode of dispersal

Method: Phylogenetic Analysis of variance

#### 3.2.1 Insular endemics

##### 3.2.1.1 Evolutionary originality

Pairwise t-values

	<i>Animal</i>	<i>Other</i>	<i>Unassisted</i>	<i>Water</i>	<i>Wind</i>
<i>Animal</i>	0	2.36	0.92	1.64	1.45
<i>Other</i>	-2.36	0	-1.52	-0.98	-1.31
<i>Unassisted</i>	-0.92	1.52	0	0.63	0.33
<i>Water</i>	-1.64	0.98	-0.63	0	-0.34
<i>Wind</i>	-1.45	1.31	-0.33	0.34	0

Pairwise corrected P-values

	<i>Animal</i>	<i>Other</i>	<i>Unassisted</i>	<i>Water</i>	<i>Wind</i>
<i>Animal</i>	0	1	1	0.48	1
<i>Other</i>	1	0	1	1	1
<i>Unassisted</i>	1	1	0	1	1
<i>Water</i>	0.48	1	1	0	1
<i>Wind</i>	1	1	1	1	0

##### 3.2.1.1 Functional originality

Pairwise t-values

	<i>Animal</i>	<i>Other</i>	<i>Unassisted</i>	<i>Water</i>	<i>Wind</i>
<i>Animal</i>	0	1.37	1.14	-0.69	0.56
<i>Other</i>	-1.38	0	-0.19	-1.63	-0.62
<i>Unassisted</i>	-1.13	0.19	0	-1.43	-0.43
<i>Water</i>	0.69	1.64	1.44	0	0.98
<i>Wind</i>	-0.57	0.62	0.44	-0.97	0

Pairwise corrected P-values

	<i>Animal</i>	<i>Other</i>	<i>Unassisted</i>	<i>Water</i>	<i>Wind</i>
<i>Animal</i>	0	1	1	1	1
<i>Other</i>	1	0	1	1	1
<i>Unassisted</i>	1	1	0	1	1
<i>Water</i>	1	1	1	0	1
<i>Wind</i>	1	1	1	1	0

#### 3.2.2 Insular non-endemics



### 3.2.2.1 Evolutionary Originality

Pairwise t-values

	<i>Animal</i>	<i>Other</i>	<i>Unassisted</i>	<i>Water</i>	<i>Wind</i>
<i>Animal</i>	0	1.40	-0.32	1.08	0.50
<i>Other</i>	-1.40	0	-1.35	-0.5	-0.82
<i>Unassisted</i>	0.32	1.35	0	1.048	0.65
<i>Water</i>	-1.08	0.54	-1.04	0	-0.38
<i>Wind</i>	-0.50	0.82	-0.65	0.38	0

Pairwise corrected P-values

	<i>Animal</i>	<i>Other</i>	<i>Unassisted</i>	<i>Water</i>	<i>Wind</i>
<i>Animal</i>	0	1	1	0.48	1
<i>Other</i>	1	0	1	1	1
<i>Unassisted</i>	1	1	0	1	1
<i>Water</i>	0.48	1	1	0	1
<i>Wind</i>	1	1	1	1	0

### 3.2.2.2 Functional Originality

Pairwise t-values

	<i>Animal</i>	<i>Other</i>	<i>Unassisted</i>	<i>Water</i>	<i>Wind</i>
<i>Animal</i>	0	-0.016	1.84	1.93	0.28
<i>Other</i>	0.016	0	1.48	1.38	0.22
<i>Unassisted</i>	-1.84	-1.48	0	-0.35	-1.36
<i>Water</i>	-1.93	-1.38	0.35	0	-1.25
<i>Wind</i>	-0.28	-0.22	1.36	1.25	0

Pairwise corrected P-values

	<i>Animal</i>	<i>Other</i>	<i>Unassisted</i>	<i>Water</i>	<i>Wind</i>
<i>Animal</i>	0	1	1	1	1
<i>Other</i>	1	0	1	1	1
<i>Unassisted</i>	1	1	0	1	1
<i>Water</i>	1	1	1	0	1

<i>Wind</i>	1	1	1	1	0
-------------	---	---	---	---	---

#### 4 Features of islands where original species are found

Method: generalized least squares under phylogenetic constraint

##### 4.1 Insular endemics

##### 4.1.1 Evolutionary Originality

*Coefficients*

	<i>Value</i>	<i>Std.Error</i>	<i>t-value</i>	<i>p-value</i>
<i>Intercept</i>	275.81	1.83	150.52	0,00
<i>Dist</i>	-0.002	0.017	-0.12	0.89
<i>Area</i>	0.006	0.042	0.15	0.87
<i>Latitude</i>	-0.007	0.068	-0.10	0.91
<i>Longitude</i>	0.0067	0.053	0.12	0.89
<i>SLMP</i>	0.0055	0.080	0.0686	0.94
<i>GMMC</i>	0.0060	0.037	0.16	0.87
<i>Elev</i>	-0.0089	0.045	-0.19	0.84
<i>Agemax</i>	0.0026	0.065	0.041	0.96
<i>Glaciated</i>	0.00074	0.015	0.047	0.96

##### 4.1.2 Functional originality

	<i>Value</i>	<i>Std.Error</i>	<i>t-value</i>	<i>p-value</i>
<i>Intercept</i>	0.05372446	0.094	0.56	0.57
<i>Distance</i>	-0.014	0.0075	-1.98	0.049
<i>Area</i>	0.042	0.010	4.05	0.0001
<i>Latitude</i>	0.016	0.010	1.63	0.10
<i>Longitude</i>	-0.0032	0.011	-0.28	0.77
<i>SLMP</i>	-0.042	0.020	-2.02	0.045
<i>GMMC</i>	-0.0018	0.012	-0.14	0.88
<i>Elevation</i>	-0.042	0.014	-2.89	0.0047
<i>Age</i>	-0.042	0.017	-2.4	0.015

<i>Glaciated</i>	-0.011	0.0031	-3.75	0.0003
------------------	--------	--------	-------	--------

## 4.2 Insular non-endemics

### 4.2.1 Evolutionary originality

#### *Coefficients*

	<i>Value</i>	<i>Std.Error</i>	<i>t-value</i>	<i>p-value</i>
<i>Intercept</i>	269.35	1.02	263.37	0
<i>Distance</i>	0.00004	0.0039	0.010	0.99
<i>Area</i>	0.00004	0.0037	0.0098	0.99
<i>Latitude</i>	-0.0002	0.0074	-0.026	0.97
<i>Longitude</i>	0.00012	0.0062	0.019	0.98
<i>SLMP</i>	0.00009	0.0075	0.011	0.99
<i>GMMC</i>	0.00007	0.0052	0.014	0.98
<i>Elevation</i>	0.00005	0.0043	0.011	0.99
<i>Age</i>	-0.00005	0.0042	-0.012	0.99
<i>Glaciated</i>	-0.00011	0.0072	-0.015	0.98

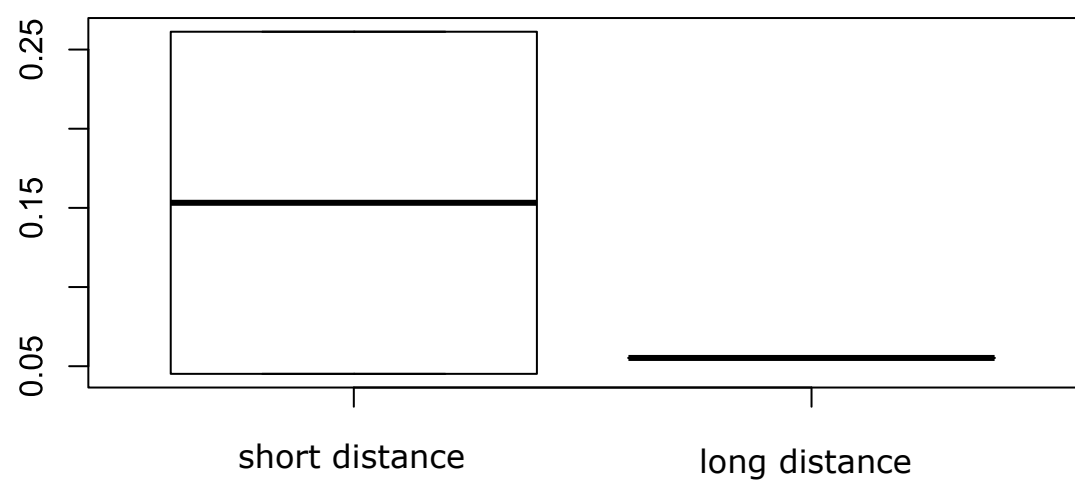
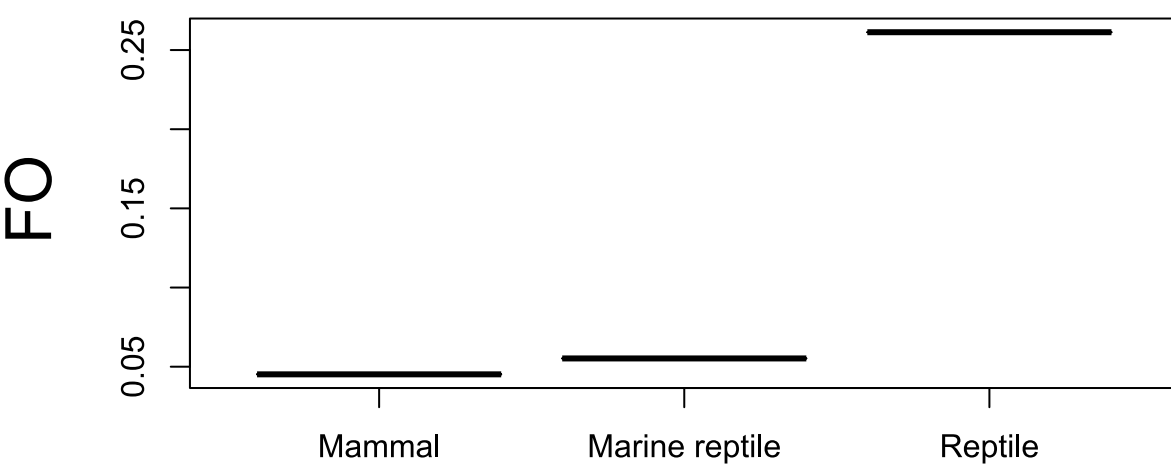
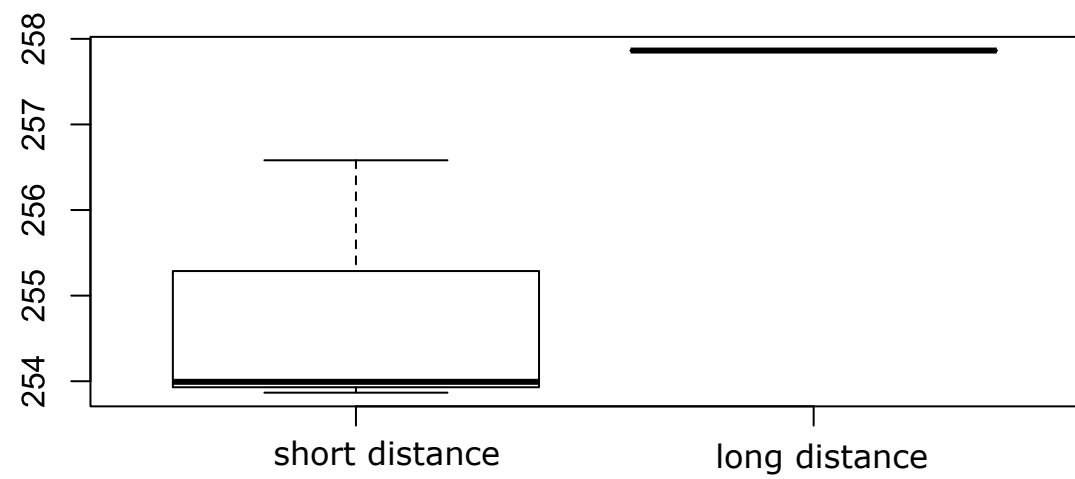
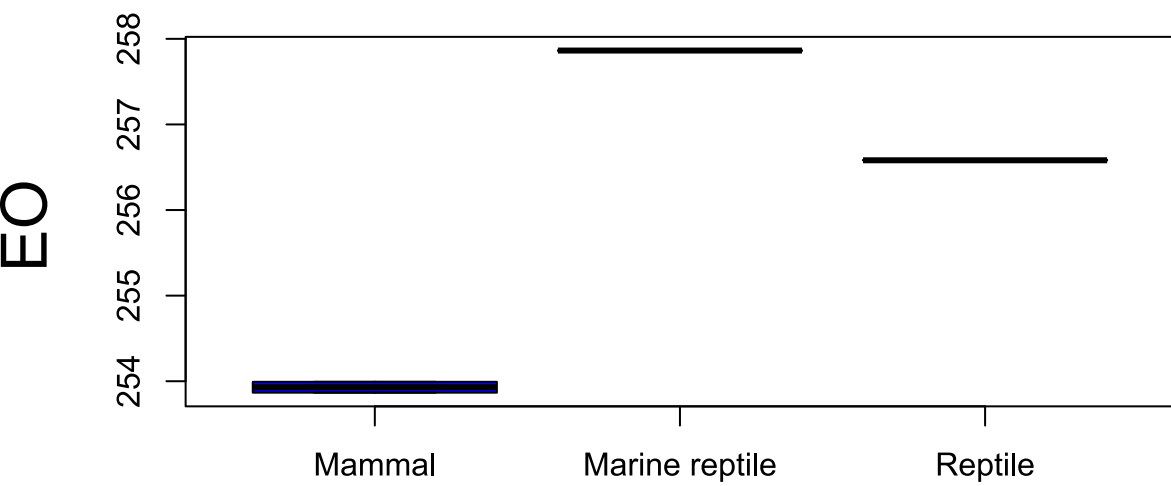
### 4.2.2 Functional originality

	<i>Value</i>	<i>Std.Error</i>	<i>t-value</i>	<i>p-value</i>
<i>Intercept</i>	0.062	0.10	0.60	0.54
<i>Distance</i>	0.0035	0.00045	7.80	0.00
<i>Area</i>	-0.0019	0.00049	-3.98	0.0001
<i>Latitude</i>	0.0041	0.00090	4.56	0.00
<i>Longitude</i>	0.00129	0.00073	1.69	0.090
<i>SLMP</i>	0.0094	0.00087	10.83	0.00
<i>GMMC</i>	0.00082	0.00062	1.31	0.18
<i>Elevation</i>	0.0027	0.00048	5.6	0.00
<i>Age</i>	-0.00013	0.00049	-0.27	0.78
<i>Glaciated</i>	-0.0042	0.00084	-5.09	0.00

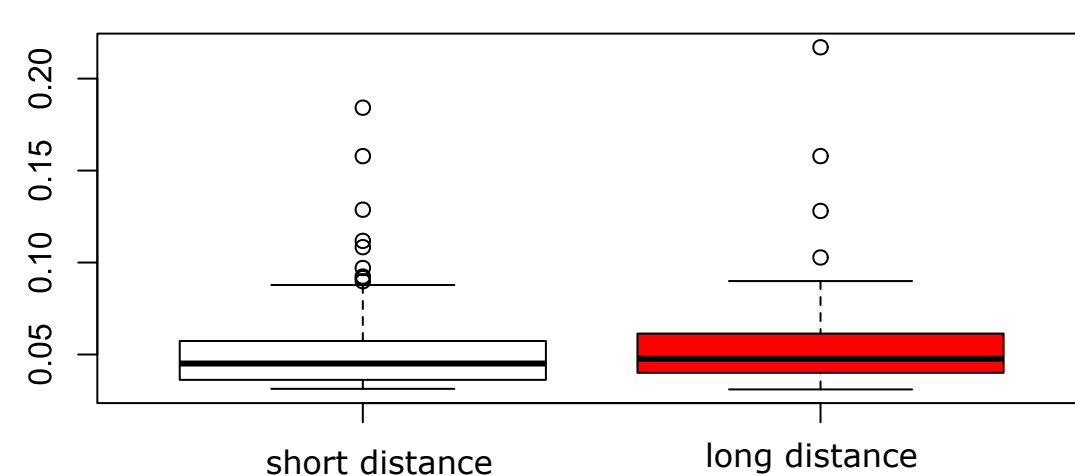
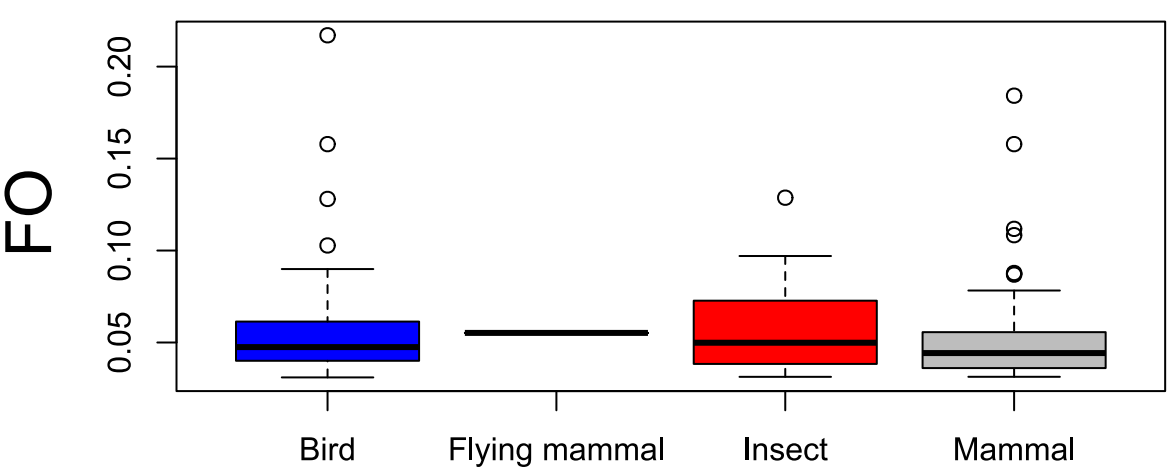
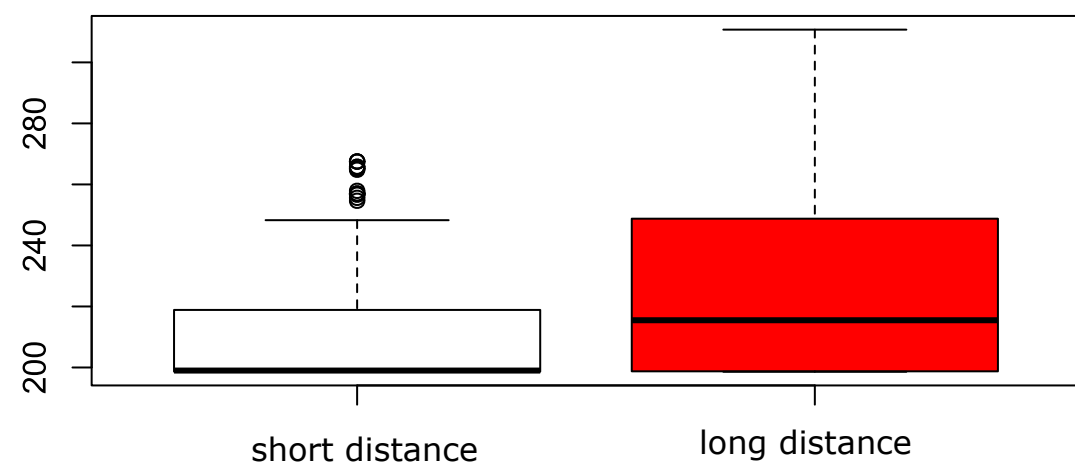
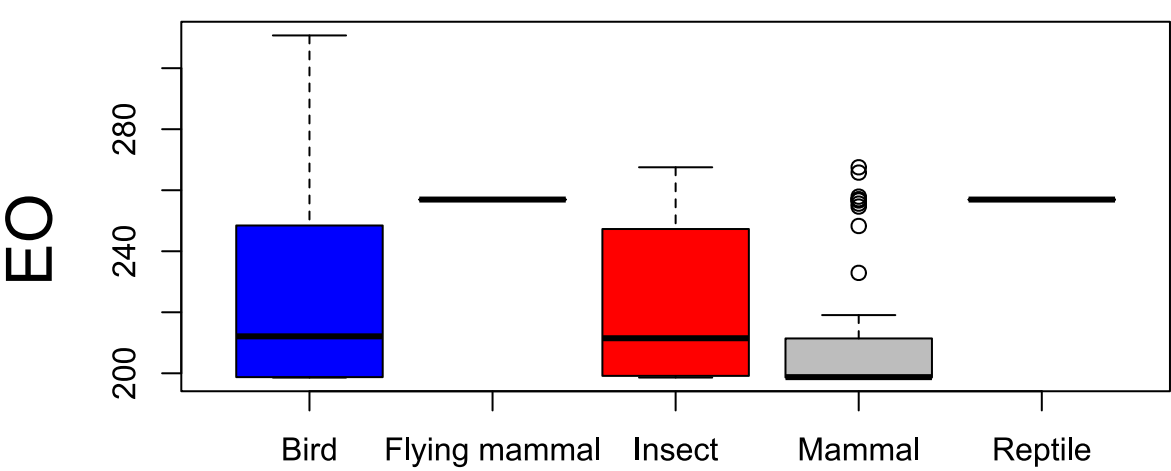
# Appendix 3

## Originality scores depending on dispersal vectors of zoochor species

### a) Endemic species



### b) Non endemic species





# VIGIE-FLORE PROTOCOL

---



## Protocole du programme Vigie-flore Observatoire de la flore commune



*Protocole défini collectivement à l'UMR 7204,  
Conservation des Espèces, Restauration et Suivi des Populations  
du Muséum national d'histoire naturelle, Paris  
et testé depuis 2005 en Ile-de-France.*

**Mars 2010.**

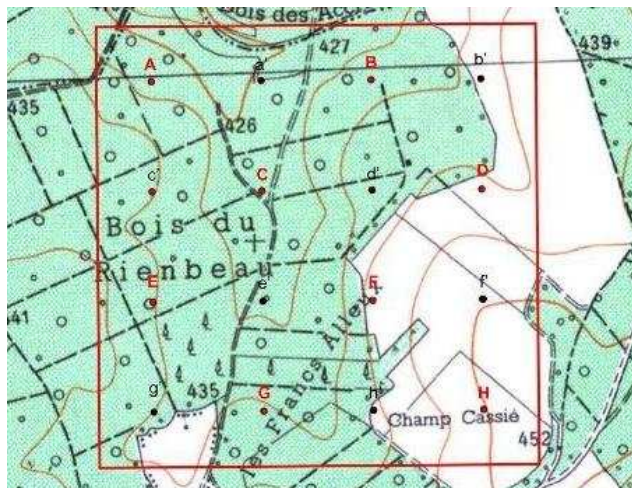
*Correspondantes : Nathalie Machon et Emmanuelle Porcher  
vigie-flore@mnhn.fr*

### *Attribution de maille(s) à l'observateur*

Chaque observateur bénévole doit choisir la maille dans laquelle il désire faire son relevé. Pour ce faire, vous devez vous inscrire sur le site [www.vigie-flore.fr](http://www.vigie-flore.fr) rubrique « devenir observateur ». Durant l'inscription, vous aurez accès à une carte de France des mailles Vigie-flore. Les points rouges correspondent aux mailles déjà sélectionnées. Les mailles vertes sont celles que vous pourrez choisir. Passez la flèche de votre souris sur la maille qui vous intéresse, son numéro s'affichera. C'est ce numéro que vous devrez entrer dans le formulaire d'inscription pour qu'elle vous soit attribuée. Vous pouvez renouveler cette opération si vous désirez inventorier 2 mailles. Ensuite, vous devez remplir le formulaire d'inscription. Une fois votre inscription faite, vous aurez la possibilité de charger la carte, la photo aérienne de votre maille et tous les documents nécessaires pour faire votre suivi via l'interface de saisie dans la rubrique « saisir vos données » du site internet.

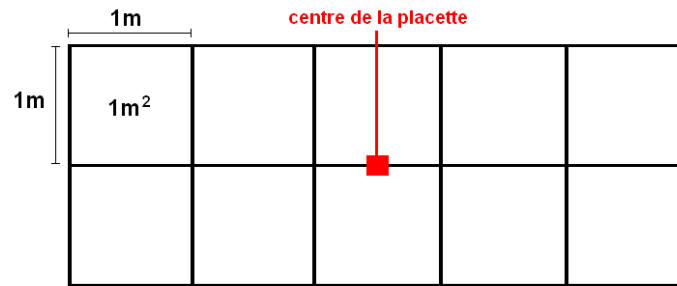
### *Échantillonnage de la maille*

La maille d'un kilomètre carré est échantillonnée selon un dispositif systématique. 8 placettes fixes sont disposées selon une configuration pré-établie (Figure 1) désignées sur la carte qui est envoyée. L'observateur va alors chercher à se rendre sur chaque point afin d'effectuer ses inventaires. Il doit échantillonner au moins les quatre points de la diagonale (A, C, F, H). Dans l'idéal, il échantillonnera également les 4 autres points (B, D, E et G). Si seuls 1, 2 ou 3 points ont été échantillonnés, les données nous intéressent quand même. Vous pourrez nous les envoyer.



**Figure 1 :** Disposition des 8 placettes dans la maille. Chacune est indiquée par une lettre (en rouge) de A à H. Pour chaque point se trouve un point de rechange (indiqué en noir) de a' à h'. Ces points serviront de substitution lorsque les principaux seront inaccessibles.

## *Inventaire des placettes*



**Figure 2 :** Forme et dimension d'une placette

La zone à inventorier est délimitée par une placette de forme rectangulaire, recouvrant 10 m<sup>2</sup> (Figure 2) à l'intérieur de laquelle 10 quadrats de 1 m<sup>2</sup> sont réalisés. Dans la pratique cela revient à utiliser quatre piquets reliés par des cordes de 1 m de long afin de repérer les limites d'un quadrat (Figure 3) et de répéter cette zone d'échantillonnage 10 fois. Les 10 quadrats doivent être placés de façon contigüe afin que la localisation des quadrats ne soit pas orientée par un choix de l'observateur.



**Figure 3 :** un quadrat

Seules les plantes vivantes ayant leur pied dans la placette doivent être prises en compte. Ainsi, un arbre dont la couronne surplombe la placette mais dont le pied est en dehors de celle-ci ne doit pas être inventorié.

## *Positionnement des placettes*

Afin de garantir la représentativité des données, il est important que l'observateur se rende aussi précisément que possible à l'endroit indiqué sur la carte. Quelques exceptions sont envisagées ; lorsque l'observateur manque de points de repères, ou n'a pas la possibilité de marquer l'emplacement de son relevé, ce dernier peut choisir de se placer à proximité d'un point de repère (arbre isolé, pancarte, rocher, etc.), à condition toutefois que la placette soit positionnée dans le même type de milieu que celui désigné par le point. Si par exemple un des points d'échantillonnage tombe dans une pelouse, l'observateur est libre de se déplacer de plusieurs dizaines de mètres jusqu'à trouver un endroit facilement relocalisable dans cette pelouse. En revanche le fait de se déplacer, même sur une plus courte distance hors de cette pelouse pour aller échantillonner un milieu jugé plus intéressant, entraînerait un biais dans les données recueillies. Enfin, dans ce cas, il est souhaitable que la placette se trouve séparée d'au moins plusieurs mètres



du point de repère afin d'éviter des effets dus à la présence de ce dernier (traitement au pied d'un panneau, flore particulière au pied d'un arbre, etc.).

- *Cas des champs cultivés :*

Afin de pouvoir relocaliser plus facilement la placette et d'éviter de trop piétiner les cultures, les points tombant dans les parcelles agricoles devront être échantillonnés dans le champ mais seulement à 5m de la bordure comme indiqué sur la Figure 4.

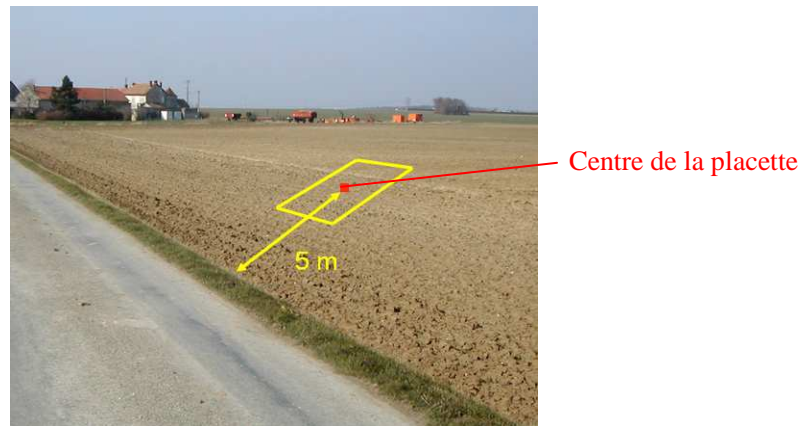


Figure 4 : Exemple de relevé effectué dans une parcelle agricole

- *Points inaccessibles :*

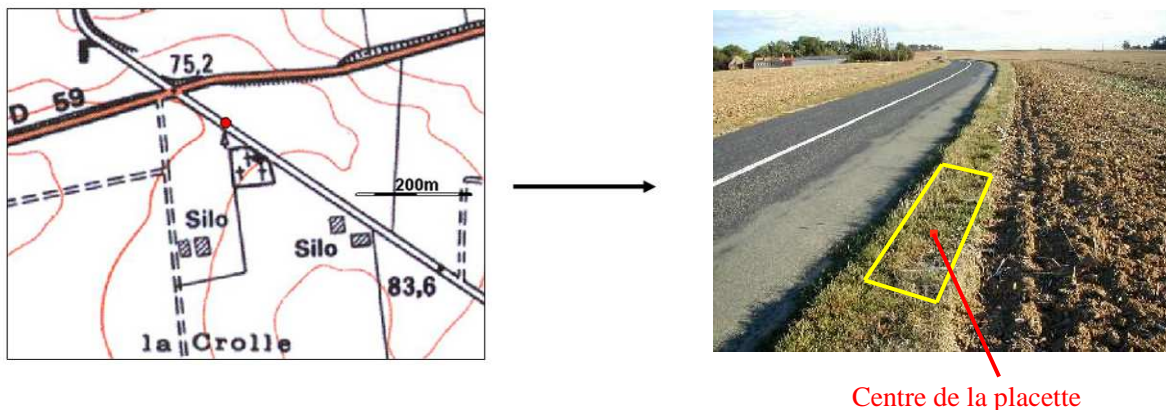
Lorsqu'un point à échantillonner s'avère inaccessible (accès non autorisé par le propriétaire, toit, falaise, végétation trop dense etc.) l'observateur utilisera un point de rechange (Figure 2) prédéfini. Si l'accès à ce dernier est également impossible, aucun relevé n'est effectué et la donnée est considérée comme manquante.

- *Zones sans végétation :*

Il peut arriver que le point tombe sur une zone non végétalisée (parking, etc.). Dans ce cas le relevé ne doit pas être déplacé, il est considéré comme vide (absence de plante) mais doit tout de même faire l'objet d'un bordereau.

- *Structures linéaires :*

Lorsque le point se situe sur un élément linéaire (route, rivière, haie, etc.) ou dans son voisinage immédiat, la placette sera positionnée systématiquement sur cet élément comme indiqué sur la Figure 5, sans déborder sur plusieurs habitats différents.



**Figure 5** : Un point peut tomber à proximité d'un élément linéaire. Ainsi selon l'interprétation de l'observateur la placette pourra se situer sur la route (relevé nul) ou sur la bordure.

### ***Périodicité des relevés***

Il est recommandé que chaque maille soit échantillonnée une fois par an.

- En juin ou juillet pour l'ensemble des régions atlantiques et continentales,
- En avril ou mai pour la région méditerranéenne
- En juillet ou août pour les mailles situées à plus de 1000 m d'altitude.

Le passage sur chaque maille sera ensuite renouvelé tous les ans, si possible par le même observateur.

### ***Retour sur les placettes***

Autant que possible, l'observateur devra se repositionner au même endroit l'année suivante. Afin de retrouver facilement le positionnement des placettes d'un passage à l'autre, le relevé des coordonnées GPS avec un appareil standard pour itinéraire routier d'une précision de 1 à 10m est suffisant et peut s'avérer utile. La photographie du lieu peut aussi être efficace. Si malgré ces précautions, la relocalisation précise est impossible, l'observateur devra se placer dans le même habitat que le relevé précédent pour effectuer le relevé. Dans le cas où le milieu est fortement perturbé/modifié mais que la relocalisation est possible, le relevé doit se faire au même endroit que le relevé précédent.

### ***Cas des espèces de détermination délicate***

Les botanistes participant au programme devront viser l'exhaustivité dans l'identification des plantes se trouvant dans leurs placettes. Néanmoins, la détermination de certaines familles de plantes (Poacées, Cypéracées, etc....) peut poser problème à certains observateurs. Dans ce cas il est préférable que l'observateur en reste à une détermination au niveau du genre (par exemple « *Carex sp* »), voire de la famille (par exemple « Poacée »), et signale toute donnée dont il n'est pas certain comme « douteuse ». Une donnée imprécise est en effet préférable à une donnée fausse.

### ***Paramètres du milieu à relever***

Un certain nombre de variables devront être notées au cours du relevé :

- Type d'habitat selon une typologie simplifiée fournie à l'observateur (d'après CORINE BIOTOPE)
- Pente (en degrés) selon une échelle fournie à l'observateur
- Exposition (N/S/E/O)
- Ombrage : oui (placette toujours à l'ombre)/non (placette toujours à la lumière)/partiel : placette à l'ombre sur une partie de sa surface ou bien une partie de la journée
- Type de sol selon une typologie fournie à l'observateur
- Signes de dégradation éventuels

Hormis le type d'habitat, le relevé des autres variables est facultatif. Néanmoins, les observateurs pourront relever ces données au fur et à mesure de leurs différents passages sur la maille.

### **Avant les premiers relevés**

Avant de faire vos tout premiers relevés, lisez attentivement le protocole. Si vous avez des questions, contactez-nous à l'adresse [vigie-flore@mnhn.fr](mailto:vigie-flore@mnhn.fr).

Construisez votre quadrat : utilisez, par exemple, quatre petits piquets en bois de 30 à 40 cm suffisamment solides pour être enfoncés dans le sol. Reliez-les avec des morceaux de cordes de 1 m de long. Vous pouvez aussi utiliser deux mètres de carrelleur de 2m chacun que vous pliez en deux pour former un quadrat d'1m<sup>2</sup> (voir Figure 6).



**Figure 6** : Quadrat fabriqué avec deux mètres de carrelleur de 2 m.

### ***Dernières recommandations***

#### **Avant de partir**

Avant de vous rendre sur une maille, étudiez attentivement la carte qui vous est fournie afin de localiser précisément l'endroit où vous devez vous rendre.

Liste du matériel nécessaire :

- le protocole
- quatre à huit exemplaires du bordereau (vous devez remplir un bordereau par placette) ainsi que des feuillets supplémentaires dans le cas où certaines placettes compteraient plus de 19 espèces.
- Un crayon de papier et une gomme
- Un support pour écrire
- Votre quadrat
- De quoi récolter les individus que vous n'arriverez pas à identifier
- La lettre de demande d'accès aux propriétés privées
- Une boussole pour déterminer l'orientation de vos placettes
- Un récepteur GPS si vous en possédez un
- Un appareil photo (facultatif)

#### **Sur place**

##### **- premier passage sur une maille**

Lors de votre premier passage sur une maille, positionnez vos placettes le plus près possible de l'endroit indiqué sur la carte et dans le même habitat. Prenez des repères précis et pérennes pour retrouver la placette lors du prochain passage. Eventuellement prenez des coordonnées GPS de la placette ou prenez une photo du lieu. Un GPS classique pour itinéraire routier a une précision de 1 à 3 m en milieu ouvert et de 5 à 10 m

en forêt, cette précision sera suffisante lors du prochain passage pour que vous sachiez que vous êtes proche de votre placette, les autres repères visuels que vous aurez notés vous permettront alors de la trouver facilement.

Si une des placettes se trouve sur une propriété privée, demandez l'accès au propriétaire en expliquant le but du programme (la lettre type fournie peut vous y aider). Si le propriétaire vous a refusé l'accès ou s'il est injoignable et que la propriété est fermée, utilisez le point de rechange prédéfini.

- **passages suivants**

Lors des passages suivants sur une maille, aidez vous des repères pris lors du premier passage pour replacez le plus précisément possible vos placettes (une précision de quelques mètres est suffisante). Si vos repères sont perdus, replacez-vous dans le même habitat.

- **dans tous les cas**

Remplissez un bordereau par placette comme indiqué dans la notice.

Récoltez les individus que vous n'arrivez pas à déterminer sur place en notant avec attention la maille, la placette et le quadrat dans lequel(le)s il(s) étai(en)t présent(s). Evidemment, ne prélevez pas un des individus situés dans vos quadrats, choisissez le autour.

**De retour**

Déterminez à l'aide de flore les individus que vous n'avez pas pu déterminez sur le terrain. Si certains individus restent indéterminables, mettez-les en herbier en notant avec attention la maille, la placette et le quadrat dans lequel(le)s ils étaient présents afin de vous faire aider ultérieurement pour ces déterminations. (Voir avec le site de Tela botanica par exemple ou avec les autres bénévoles Vigie-flore de votre région)

Saisissez vos données en ligne sur la base de données SPAF (Suivi des Plantes A Fleur), voir notice de saisie.



**TEACHING THEORY OF EVOLUTION  
AT PRIMARY SCHOOL**

---



Formation d'Ecole Doctorale 227

Accompagnement scientifique du Muséum

## « Classification de l'arbre du vivant et son évolution »



Remis le 27/07/2017

Doctorante Ière année:  <b>Anna Kondratyeva</b> Muséum National d'Histoire Naturelle Département Homme et environnement UMR 7204 CESCO 61, rue Buffon, 75005, Paris	Enseignante :  <b>Isabelle Bénéfice</b> Ecole élémentaire publique Parmentier, 109, av. Parmentier XIIème  classe de CM2 (25 élèves)
---	--



## Introduction

*« Rien n'a de sens en biologie si ce n'est à la lumière de l'évolution »*

Theodosius Dobzansky, généticien allemand, 1973

Les sciences devraient être accessibles à tous. De nombreux faits scientifiques se voient déformer le visage suite à des fausses interprétations des médias faites sans aucune consultation des spécialistes concernés. Partant de ce constat je me suis engagée depuis l'année 2010 dans l'animation scientifique auprès des enfants de 3 à 14 ans. Dans les colonies de vacances à thèmes naturaliste, les ateliers scientifiques en temps périscolaire et un service civique en enseignement à l'environnement et au développement durable. Je pense que la sensibilisation des jeunes à l'environnement qui les entoure à travers de nombreuses expériences ludiques est une action importante dans leur formation extra-scolaire, car en dehors des bancs de l'école l'enfant perçoit et appréhende les informations d'une manière différente, selon ses propres choix et intérêts (l'éducation populaire n'est pas obligatoire mais elle est ouverte à tous).

Le doctorat ne me laisse pas suffisamment de temps pour m'occuper pleinement des activités avec les enfants. Cependant, grâce à une proposition de formation doctorale qui s'intitule « L'accompagnement scientifique du Muséum » j'ai pu intervenir dans une classe d'école primaire sur une des thématiques de ma recherche : l'évolution et la classification du vivant. Ce dispositif est mis en place par une collaboration entre le Muséum National d'Histoire Naturelle et la Fondation « La Main à La Pâte » (LAMAP). Le but est d'associer l'enseignant, le scientifique (un doctorant ou un chercheur) et les élèves dans une logique d'enrichissement mutuel et de partage de compétences. Cet accompagnement se distingue d'autres formes d'interventions car l'enseignant n'est plus tout seul pour tenir une séance en classe. La présence d'un scientifique et ses apports dans la préparation des séquences aident l'enseignant à conduire avec confiance les séquences dont il n'est pas forcément spécialiste et d'avoir une autre vision sur sa pratique d'enseignement. Pour le scientifique, cet engagement permet de développer des compétences de médiation scientifique

auprès d'un public non expert. Enfin les élèves, qui sont les éléments principaux pour moi dans ce dispositif, développent leur questionnement du monde ainsi qu'une ouverture d'esprit et satisfont leur curiosité par l'observation et l'expérimentation.

L'évolution : un enseignement à risque ?

La thématique choisie, celle du concept de l'évolution et de la classification du vivant n'est pas un savoir neutre comme en témoignent les polémiques actuelles et passées. L'évolution est un phénomène non spectaculaire qui s'inscrit dans une fenêtre de temps et d'espace qui échappe pour une bonne part à la perception humaine. L'évolution du vivant se déroule à une échelle de temps géologiques trop longs et aussi dans l'infiniment petit<sup>10</sup>. Mais c'est également un principe organisateur de toutes les disciplines biologiques qui permet de comprendre l'unité et la diversité du vivant. Ce sujet d'étude par la nature même des savoirs en jeu, est particulièrement intéressant et délicat à expliquer car il est confronté aux idées présentes dans le domaine public où se mêlent sciences, philosophies, idéologies et religions. C'était donc pour moi un véritable défi de prendre en compte dans mes explications la perception personnelle du monde des élèves. Ainsi, je voudrais non seulement partager mes connaissances et vulgariser mon travail au muséum (j'effectue une thèse dans un domaine d'écologie évolutive) mais aussi retrouver le plaisir du travail avec les jeunes. Avant de commencer j'ai décidé de me renseigner sur le niveau d'enseignement des concepts d'évolution et de la classification du vivant dans les écoles primaires en France. L'évolution apparaît de façon explicite dans les programmes d'écoles primaires qu'à partir de 1985. Ce sujet était alors abordé de façon succincte en s'appuyant principalement sur l'observation des fossiles et l'analyse de documents. Avec les programmes scolaires renouvelés depuis 1995 nous observons une progression dans les objectifs de connaissances enseignées. De « L'évolution du vivant » en 1985, on passe à « Première approche de la notion d'évolution à partir de l'unité du vivant, caractérisée par la mise en évidence de quelques grands traits communs, puis de sa diversité, illustrée par

---

<sup>10</sup> Guillaume Lecointre, 2002. *Classification phylogénétique du vivant*, deuxième édition, avec Hervé Le Guyader. Belin

l'observation de différences, le tout conduisant aux notions d'espèce et d'évolution ; première approche de la notion d'évolution des êtres vivants à partir de quelques fossiles typiques » en 2007.

En 2001 le comité scientifique de la fondation « LAMAP » propose de diffuser au niveau des écoles primaires un protocole pédagogique adapté consacré aux classifications du monde vivant. La même année la classification phylogénétique devient un savoir obligatoire pour passer le CAPES et l'agrégation. En 2004 la classification du vivant fait partie du programme du cycle 2 (CP, CE1 et CE2) et de cycle 3 (CM1, CM2 et 6e) au sein du chapitre sur l'évolution du vivant.

Etant donné que j'allais travailler avec une classe de CM2 je me suis particulièrement intéressée à des programmes du cycle 3. Le dernier Bulletin Officiel d'Enseignement National de 2016 dit que dans un cadre de la matière « Sciences et techniques » du cycle 3, une des thématiques à enseigner est « Le vivant, sa diversité et les fonctions qui le caractérisent ». Un des attendus à la fin du cycle 3 précise que les élèves doivent savoir « classer les organismes, exploiter les liens de parenté pour comprendre et expliquer l'évolution de ces organismes ». En même temps, sur le site « EduScol », mis en place pour informer et accompagner les professionnels de l'éducation, on retrouve la clef pour la mise en œuvre et la progressivité des enseignements. Concernant la thématique du vivant il est précisé que « il n'est pas attendu au cours du cycle 3 une quelconque explication de la théorie de l'évolution, mais simplement de poser les bases qui permettront d'aborder les mécanismes explicatifs développés au cycle 4 ». Il est également conseillé de ne pas aborder la construction des arbres de parentés et les arguments moléculaires en faveur d'une parenté. N'étant pas d'accord avec ce dernier point, j'ai pu expliquer à la classe l'utilité d'un arbre phylogénétique, de faire la différence avec un arbre généalogique et d'en construire un avec les élèves. J'expliquerai ma démarche en détails dans la partie suivante.

Avant de me lancer dans une construction de séance, je défini d'abord les deux concepts principaux autour desquels je travaille. Il existe certainement plusieurs façons de les définir et le lecteur peut ne pas être entièrement d'accord avec les définitions synthétiques que je propose :

L'évolution (en termes de Biologie) : processus continu par lequel les espèces se transforment au cours du temps et des générations avec la sélection naturelle comme moteur et la variabilité génétique et phénotypique comme carburant.

La classification du vivant (la Systématique) : le fait de regrouper des êtres vivants selon les caractères qu'ils possèdent. La classification établie doit refléter l'ordre phylogénétique et donc la parenté entre ces êtres vivants.

## Le projet

### *Première proposition du sujet*

Avant de rencontrer un enseignant intéressé j'avais proposé mon projet à LAMAP que j'avais intitulé « L'arbre du vivant et son évolution ». Il a été diffusé dans l'Académie de Paris auprès des professeurs des écoles. En plus des concepts de l'évolution et de la classification du vivant qui resteront des notions centrales de ce projet, j'avais inclus des notions de la biodiversité, des relations des êtres vivants avec leur milieu et donc de la diversité des habitats, ainsi que des relations trophiques entre les êtres vivants au sein de ces habitats. Personnellement, je trouvais qu'il n'était pas possible de parler des espèces vivantes sans prendre en compte le contexte écosystémique.

### *Première rencontre avec l'enseignante*

Ma proposition de séances a intéressé Isabelle Bénéfice, une enseignante stagiaire d'une classe de CM2 de l'école primaire Parmentier. Elle a été très enthousiaste à propos de la thématique de l'évolution et m'avait paru tout de suite très impliqué dans ce projet. Venant d'une formation pluridisciplinaire en économie et relations publiques elle a préféré se consacrer à l'enseignement en secteur privé d'abord, puis cette année elle a obtenu le statut de titulaire d'un professeur des écoles. Isabelle m'avait confié que les sciences en biologie n'étaient pas son « point fort » et qu'elle espérait que notre collaboration soit enrichissante pour elle.

### *Mise en place des séances*

Nous avons revu ensemble avec Isabelle le programme des séances que j'avais proposé. Sachant que j'avais écrit la première version sans consulter les manuels scolaires proposés par l'Enseignement National, j'ai été contente d'apprendre que ces séances suivaient plutôt bien leurs contenus. Isabelle m'avait expliqué les thèmes que les élèves avaient déjà vus dans la classe de « Sciences et techniques » cette année : la constitution de la matière et les différentes sources d'énergie. Elle prévoyait également de faire un cours sur la reproduction et le développement des êtres vivants, ce qui fait partie du programme obligatoire de CM2. Je pouvais donc m'appuyer sur

les connaissances que les élèves avaient déjà abordées cette année, ainsi que les années précédentes en classe de sciences. Celles-ci sont notamment les notions du vivant / non-vivant, de la diversité des êtres vivants et des milieux, et de l'espace-temps (une notion importante pour comprendre l'échelle du temps sur laquelle se déroule l'évolution)<sup>11</sup>. Nous avons également ajouté une séance consacrée entièrement à l'histoire passée du vivant, la partie que je n'avais pas prévu d'aborder de façon explicite.

Nous nous sommes mises d'accord sur le déroulement général des séances. Chaque séance commencerait par une confrontation des conceptions premières des élèves pour mettre en route des investigations personnelles et une remise en cause des idées. En deuxième partie un recueil des idées nouvelles et une mise en commun des résultats permettrait de faire un bilan avec toute la classe. A la fin une évaluation pourrait éventuellement avoir lieu. Isabelle avait mis un accent sur la démarche d'investigation, de recherche personnelle et de questionnement à suivre. Enfin, nous avons convenu de préparer les séances en détail au fur-et-à-mesure du déroulement du projet en fonction de la vitesse d'avancement et des remarques des élèves. Pour cela, nous restions dans la classe après chaque séance pour préparer une séance suivante.

Pour développer au mieux les séances je me suis surtout servie des ressources pédagogiques sur le site de LAMAP (« A l'école de la biodiversité ») et de l'ouvrage « Comprendre et enseigner la classification du vivant » sous la direction de Guillaume Lecointre. La totalité de ressources utilisées se trouve dans la partie Bibliographie.

---

<sup>11</sup> Bulletin Officiel de l'Education Nationale 2016, Ministère de l'Education Nationale

Tableau 1. Le contenu des séances.

Séance	Thème	Méthode / mots clé	Matériel	Commentaires
24/02	Unité et diversité du vivant	Vivant/non-vivant; espèce; biodiversité	Planches d'images des espèces à regrouper	Premier contact avec la classe
03/03	Classification du vivant : décrire	Trier/Ranger/Classer Caractères partagés	Tableau de caractères Images des espèces à décrire	Observation et description des espèces
10/03	Classification du vivant : classer	Classer, emboîter, nommer	Tableau de caractères Images des espèces à décrire	Classification emboîtée
17/03	Diversité actuelle et passée des espèces	Fossile, ancêtre commun, évolution	Vidéo, diaporama	Première définition d'évolution
24/03	Parenté des espèces	Construction d'un arbre de parenté, notion du temps lent	Diaporama	Insister sur le terme de changement des caractères
31/03	Espèces, habitats, chaînes trophiques	Habitat, adaptation, interaction Jeu de la chaîne trophique	Images de 6 habitats Liste d'espèces à replacer sur l'arbre	Adaptations aux différents habitats

26/06	Jeu de piste dans la GGE	Jeu de piste	Feuille de piste Tableau de caractères	Découverte des espèces du Muséum
30/06	Bilan	Discussion collective	Diaporama	Recueil des impressions générales sur les séances



## Déroulement des séances

Tout le matériel nécessaire était disponible sur place (un vidéoprojecteur et un PC, des tableaux, une imprimante et une photocopieuse). Les séances se déroulaient chaque vendredi après-midi, de 13h30 à 15h. Etant donné que les élèves sortaient de la cantine et de la cour de récréation les premières minutes des séances se passaient dans le brouhaha avant le retour du calme. Lors des séances Isabelle préférait rester au fond de la classe, à surveiller le retour au calme et à écouter en même temps ce que je racontais.



## Séance 1 - Unité et diversité du vivant

Cette première séance était également une première confrontation avec la classe. En arrivant un peu en avance, j'ai pu m'installer afin de me repérer dans la classe. Isabelle avait même fait un schéma des tables d'élèves avec leurs prénoms dessus. Les élèves étaient déjà au courant d'une nouvelle intervenante et sont restés attentifs lors de la présentation du sujet. Ma promesse d'une sortie nature (qui malheureusement n'a pas pu avoir lieu) et d'un jeu de piste au Muséum ont certainement attiré leur attention et renforcé leur motivation pour participer lors des séances.

### Recueil des conceptions premières des élèves

Pour introduire un vaste sujet de la diversité du vivant il fallait mettre au clair ce que les élèves sous-entendaient comme le « vivant » et le « non-vivant ». J'avais demandé ce que leur évoquaient ces deux mots. Un par un, chacun à son tour, les élèves ont donné des exemples que j'avais noté sur le tableau. Les élèves ont surtout eus du mal à dire que les végétaux sont aussi des êtres vivants. Il a fallu alors faire un cheminement logique de la graine qui pousse (une expérience

que la plupart des élèves ont déjà pratiqué) et qui ensuite donne une plante. D'autres exemples, non seulement des noms mais aussi des concepts, des verbes et des actions (paysage, respirer, vent, animal...) ont été donnés.

Après une brève analyse des mots notés nous avons défini en trace écrite :

Le vivant : tout être organisé qui naît, se développe (respire et se nourrit), se reproduit et meurt.

Pour terminer j'avais expliqué que nous pouvons rassembler l'énorme quantité des êtres vivants que nous connaissons aujourd'hui sous un seul mot qui est la biodiversité. Je me suis attardé sur le « -bio » car parfois ça peut être associé à des produits bio, biologiques, il fallait donc mettre ces choses au clair.

### ***Mise en route des investigations***

Le but de la séance étant de préparer les élèves à l'action de la classification du vivant, je voulais faire avec eux une classification « test ». Nous avons séparé les élèves en groupes de 4. Chaque groupe a reçu un jeu de photos d'organismes vivants (Annexes). J'avais demandé alors d'observer attentivement ces photos et de mettre ensemble les êtres vivants qui se ressemblent le plus. J'avais aussi demandé de noter des raisons pour lesquels les élèves avaient mis ensemble certains organismes et pas d'autres.

### ***Recueil et mise en commun des résultats***

Après une quinzaine de minutes, chaque groupe avaient révélé les ensembles d'organismes qu'ils avaient formés et expliqué quels étaient les indices utilisés pour ce classement. Les raisons de classements étaient diverses : la couleur, la forme ou encore le « type » d'un être vivant. Certains avaient regroupé les espèces vivant dans l'eau et sur terre. En posant la question à la classe « Pourquoi les êtres vivants de chaque groupe se ressemblent-ils autant ? » la classe avait proposé : « Ils se ressemblent parce qu'ils sont de la même famille ; de la même espèce ». Ce dernier terme est sorti assez rapidement chez certains élèves ce qui ne pouvait que m'encourager à poser la question suivante : « Comment pourrait-on appeler des individus regroupés ensemble parce qu'ils se ressemblent ? ». Nous avons alors nommé chaque groupe formé : des fleurs ou des végétaux (ou

des marguerites et des bleuets), des humains, des poissons (poissons-clowns et poissons rouges) et des grenouilles (grenouille dendrobates et des grenouilles rousses).

Pour élargir les investigations des élèves j'avais demandé si les individus au sein des ensembles formés pouvaient se reproduire et donner des petits ? Et entre les ensembles ? La réponse correcte est venue rapidement, même s'ils hésitaient pour les 2 espèces de grenouilles et pour les plantes. Pour l'homme la réponse était catégorique : pas de reproduction en dehors de ce groupe. Donc chaque groupe correctement formé représentait une espèce différente.

Enfin nous avons regardé plus attentivement chaque ensemble. Le constat était le même : au sein de chacun, les individus avaient des petites différences. A la fin de la séance les élèves ont gardé une trace écrite de la définition d'une espèce :

Une espèce est un ensemble d'êtres vivants ayant des caractères en commun et qui sont susceptibles de se reproduire entre eux. À l'intérieur d'une espèce, les individus possèdent de petites différences qui les rendent uniques.

## **Séance 2.1 – Classification du vivant : décrire**

La science de la classification s'appelle la Systématique. Son but est de mettre de l'ordre et de donner du sens à l'immense diversité du vivant en se basant sur l'identification des espèces après une observation des caractères qu'ils présentent. A l'école primaire nous utiliserons uniquement des caractères morphologiques, de préférences externes (à l'exception du squelette).

### ***Recueil des conceptions premières des élèves***

En rappelant la séance précédente j'avais demandé aux élèves combien d'espèces différentes nous avions observé. Après avoir réfléchi, les élèves ont convenu qu'il existait une pléthore d'espèces animales et végétales sur notre planète. Mais personne ne pouvait les nommer toutes. En plus j'avais fait remarquer que ces espèces se retrouvaient comme dans un énorme tas en désordre et si je leur demandais de m'en sortir une en particulier ils passeraient beaucoup de temps à la chercher. Ainsi j'avais introduit le besoin de mettre de l'ordre dans la diversité du vivant,

c'est-à-dire de classer des espèces. Pour mieux comprendre le monde qui l'entoure, l'homme a besoin de classer les êtres vivants, d'organiser cette variété pour y voir plus clair.

Avant de commencer la classification proprement dite, il fallait s'assurer que les élèves comprennent bien le sens du mot « classer » et le distinguent avec les actions de « trier » et de « ranger ». Au titre d'exemple j'avais pris une des trousse et sorti tout ce qu'il y avait dedans sur une table. Ensemble nous avons trié les objets de la trousse : les crayons papier et les stylos, en objets pour écrire et les autres en objets par couleurs, etc. Donc le tri s'effectue sur une présence ou absence d'un seul critère.

Ensuite j'avais demandé de ranger ces objets. Les élèves avaient du mal à comprendre ce que je leur demandais. Pour les aider j'avais rangé les stylos du plus petit au plus grand. Donc, le rangement s'effectuait sur un critère continu.

Enfin, j'avais remis les objets ensemble et commencé par les séparer en stylos et en crayons. Au sein de chaque groupe nous les avons séparés par couleur, puis dans chaque sous-groupe par taille. Donc pour classer il faut déterminer des critères communs aux objets qui permettent d'établir une hiérarchie relationnelle entre ces objets.

Nous avons donc rappelé que ce qu'ils ont fait avec les espèces vivantes en séance 1 n'était pas du classement mais du tri car il n'y avait pas de relations entre les groupes formés, tout était basé sur des informations ponctuelles et exclusives. Le classement au contraire doit toujours se construire à partir de ce que les espèces possèdent comme critère et non de ce que les espèces font (nagent, volent, mangent de la viande, etc.).

### ***Mise en route des investigations***

J'avais proposé aux élèves d'effectuer une vraie classification scientifique. Les élèves avaient été séparés en binômes, chaque binôme ayant reçu une collection d'animaux à classer (Annexes). Nous n'avions pas fait de classification des végétaux car celle-ci est trop complexe à comprendre et il n'y a pas assez de caractères facilement observables.

Tout d'abord j'avais demandé aux élèves d'observer quelques animaux et de noter des

caractères en commun que ces espèces partageaient. Les yeux ou les poils ont été mentionnés le plus souvent. Mais certains avaient encore listés ce que les espèces font et où elles habitent. J'avais ensuite expliqué que les scientifiques utilisent des listes de critères bien définis parce que nous ne pouvons pas utiliser n'importe quoi, qu'il y avait des critères qui changeaient trop au sein de la même espèce etc. Nous avons ensuite distribué à chaque groupe d'élèves un exemplaire de la « Fiche d'observation de caractères » prédéfinis (Annexes). J'avais demandé d'observer encore une fois tous les animaux de la collection et de cocher dans le tableau les critères. Comme certains caractères n'étaient pas familiers une « Fiche de définition de caractères » avait été distribuée (Annexes) que nous avons lues ensemble.

## Séance 2.2 – Classification du vivant : classer

Dans la suite de la séance précédente nous avons repris le tableau des caractères que j'avais projeté sur le tableau que nous avons rempli tous ensemble. À partir de la « Fiche de définition de caractères » corrigé j'avais demandé à repérer des caractères que ces animaux ont en commun. Les élèves ont alors remarqué que par exemple l'humain et le chat ont des poils et 4 pattes, que la mouche et le coléoptère ont 6 pattes et un squelette externe ou encore que le pigeon et le canard ont tous les deux des plumes.

La classe a donc été mise au défi : mettre des espèces dans des boîtes selon des critères qu'ils partagent et ainsi dresser une classification emboîtée. Au titre d'exemple pour expliquer comment procéder j'avais utilisé les objets de la fig.1.

Pour aider la classe les consignes suivantes ont été données :

Les animaux sont classés à partir de ce qu'ils ont et pas de ce qu'ils n'ont pas, ni de ce qu'ils font.

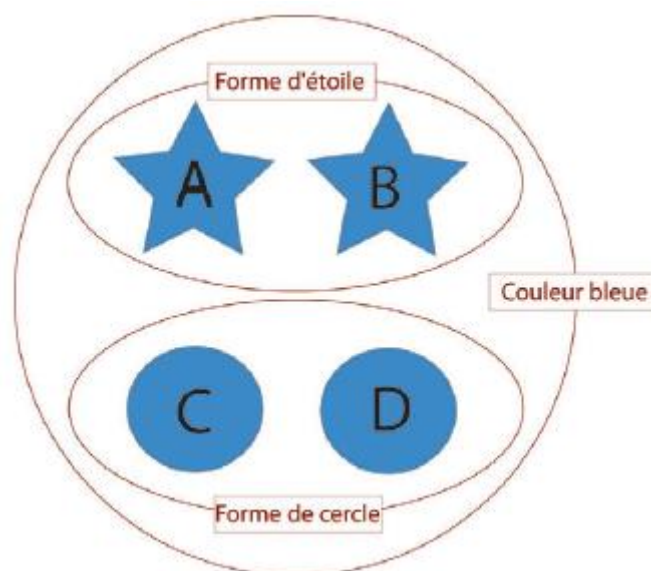


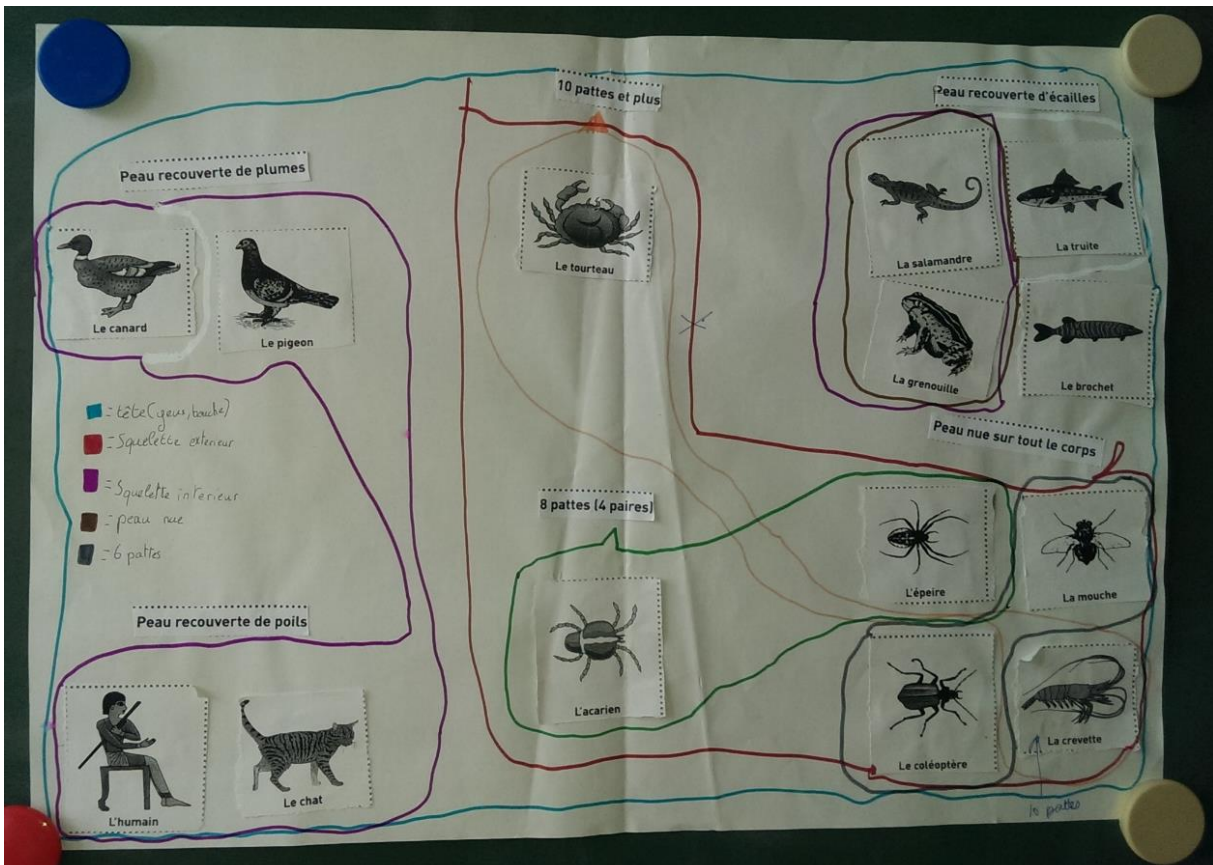
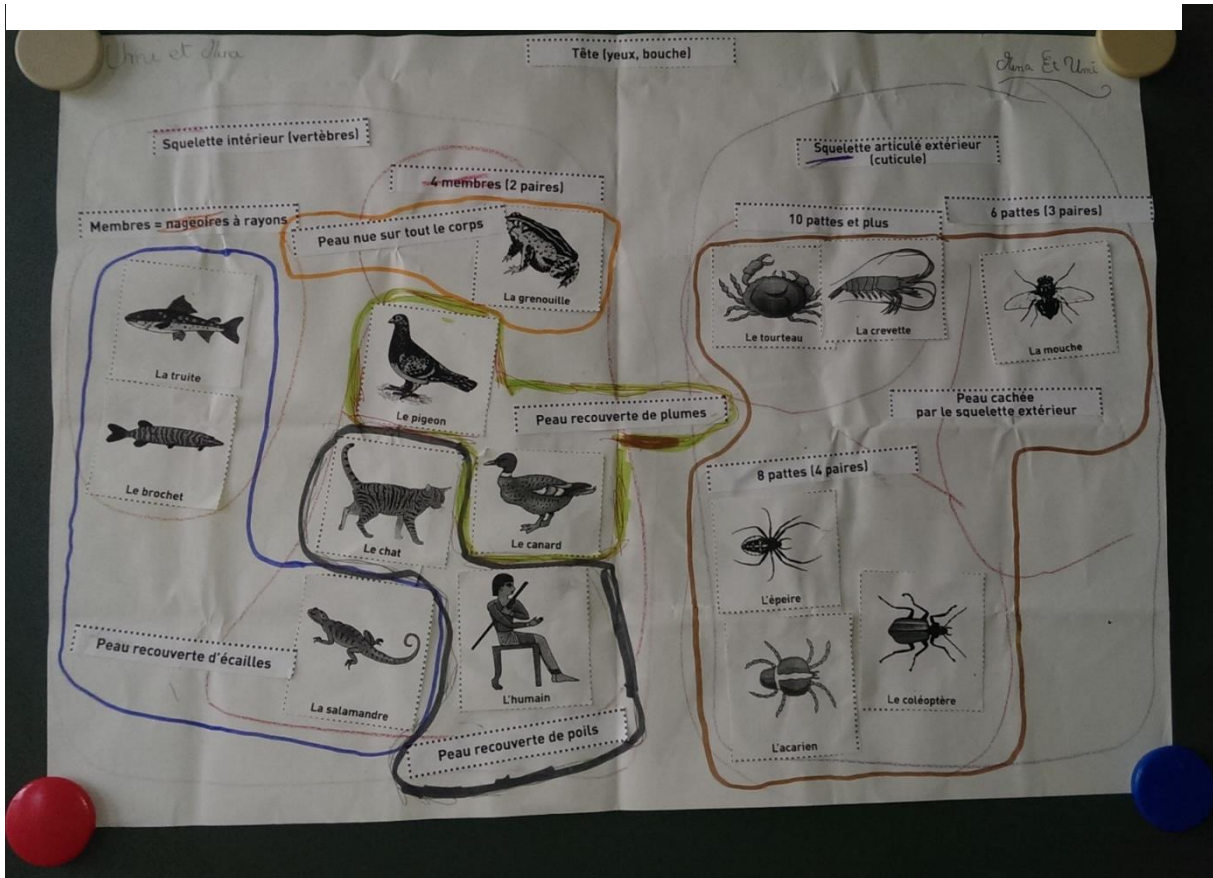
Figure 1. Exemple d'une classification emboîtée

Les boîtes construites doivent porter un seul nom – le nom d'un caractère partagé par toutes les espèces dans cette boîte. On partirait d'un caractère le plus partagé (la tête et les yeux) vers les caractères les moins partagés.

### ***Recueil et mise en commun des résultats***

Après une vingtaine minutes de découpage (images des espèces et des étiquettes avec des critères), de réflexion et de collage par groupe de 4 nous avons affiché les classements sur le tableau, mais la séance avait déjà dépassé le temps dédié, et nous avons dit que la correction se ferait la semaine suivante.

Figure 2. Exemple d'une classification emboîtée effectué par les groupes d'élèves





## Séance 4 – Diversité actuelle et passée du vivant

Avant d'aborder une nouvelle thématique nous avons fait une correction des classements emboîtés. Certains groupes ont eu du mal ou n'ont pas été jusqu'au bout de leur classement. J'avais alors demandé : « Quels noms peut-t-on donner à chaque "boite" » ? L'exercice s'est avéré plutôt simple, les élèves ont vite trouvé des noms à la plupart des groupes. Il a fallu juste aider avec les Métazoaires (la tête et les yeux), les Arthropodes et les Vertébrés. Le groupe de « Poissons » et de « Reptiles » n'existent plus dans la classification actuelle, ce sont des groupes para phylétiques (qui n'ont pas un unique ancêtre en commun). Mais au niveau de classe de primaire il est difficile (et pour moi inutile) de contredire les élèves en expliquant que dans les reptiles on devrait inclure les oiseaux et que les poissons s'appellent en vrai des Actinoptérygiens (on a exclue les raies et les requins qui ont un squelette entièrement cartilagineux). J'avais donc utilisé ces termes pour définir des groupes d'espèces, le plus important étant la compréhension du principe de la classification emboîtée.

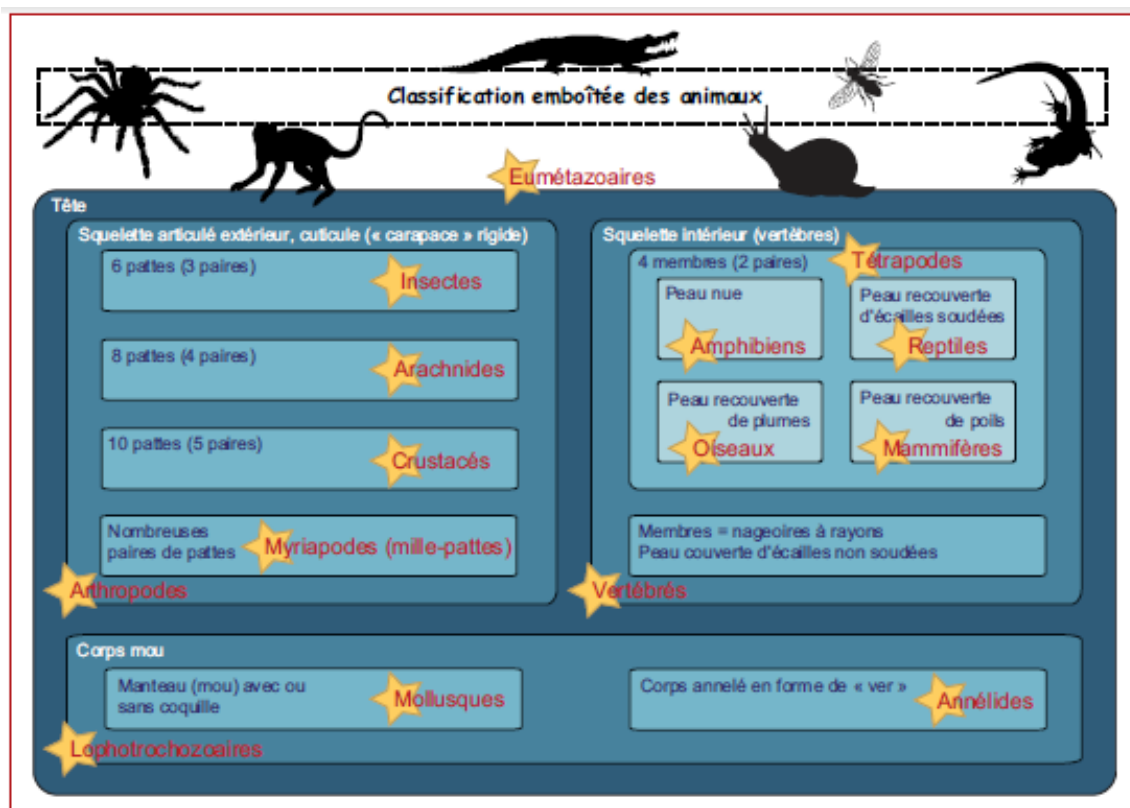


Figure 3. Corrigé de la classification emboîtée des animaux avec des noms de boîtes



Maintenant nous pouvons passer à la thématique de la séance 4. C'était une séance de transition entre la classification du vivant et son évolution. Le but était de montrer aux élèves l'existence dans le passé d'un nombre extraordinaire d'espèces vivantes qui ont disparues et qui peuvent souvent ressembler aux espèces actuelles. Sur la base de ce constat j'introduis une notion d'un ancêtre commun et de l'évolution des espèces.

### ***Recueil des conceptions premières des élèves***

J'avais commencé par demander : « Combien d'espèces existe-t-il sur la Terre aujourd'hui ? ». Les élèves ont donné des chiffres très différents allant de quelques milliers jusqu'à quelques milliards. Il est difficile d'imaginer de tels chiffres, même pour un adulte, alors j'avais comparé le nombre d'espèces estimé sur terre à un nombre de gouttes dans un océan (ce qui est en vérité beaucoup plus grand). Ensuite nous avons regardé un extrait d'un documentaire de Denis Van Waerebeke « Espèces d'espèces », pour avoir une réponse correcte : actuellement on estime le nombre total d'espèces vivantes entre 5 et 10 millions, avec seulement 1.7 millions d'espèces connues.

« Et est-ce qu'il y a toujours eu les mêmes espèces présentes sur la planète ? ». La réponse de la classe était unanime : « non, il existait d'autres animaux, par exemple des dinosaures. » J'avais

donc passé un diaporama en montrant à la classe des images de fossiles (un nouveau mot à définir) et des images d'espèces animales



et végétales en disant qu'elles n'existaient plus. J'avais fait remarquer qu'elles ressemblaient parfois souvent aux espèces actuelles. Par exemple l'Archéoptéryx avait des plumes et des ailes, 4 membres, une tête et des yeux, mais aussi des dents.

Il partageait des caractères en commun avec des oiseaux et des « reptiles » d'aujourd'hui (qui font partie d'un même groupe monophylétique). Ils sont donc des proches parents et ils possèdent un ancêtre en commun dans le passé. Il était important d'insister sur le fait que nous ne connaissons jamais à quoi ressemblaient les ancêtres communs des espèces actuelles, que ce sont des ancêtres hypothétiques, disparues aujourd'hui. Elles ont évolué, changé et sont devenues telles qu'on les connaît aujourd'hui. Ce processus continue toujours.

### ***Recueil et mise en commun des résultats***

Cette séance m'a permis également de préparer le terrain pour des explications de concept d'un arbre phylogénétique en introduisant la notion de parenté entre les espèces et de leurs évolutions. Lors de cette séance les élèves se sont montrés très attentifs, certainement captivés par les images des espèces du passé qui rassemblent à des monstres (des trilobites par exemple) et ont posé beaucoup de questions sur les causes de leur disparition et leurs modes de vies.



Figure 4. Faune de l'océan Cambrien : il y a 540 000 000 d'années

## Séance 5 – Parenté des espèces

Comme mentionné précédemment, je trouvais qu'il est impossible de parler de la diversité du vivant sans toucher un mot sur la représentation de la parenté sous forme d'un arbre. En effet, une classification emboîtée représente un arbre phylogénétique vu de dessus. Avec une telle représentation on introduit la notion du temps, et le concept d'évolution dans le temps.

### *Recueil des conceptions premières des élèves*

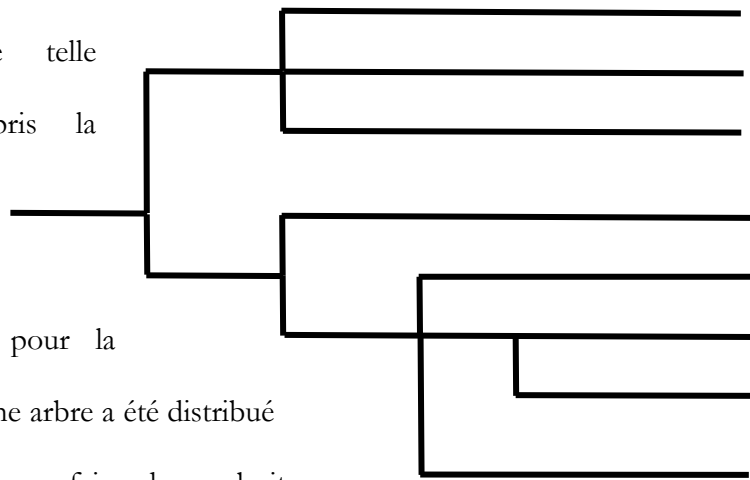
Les élèves de l'école primaire n'ont pas forcément la facilité de se projeter dans une troisième dimension, celle du temps et encore moins de la voir sous forme d'un arbre. J'avais donc commencé par rappeler, en séance précédente, que nous avons déjà parlé des ancêtres communs qui existaient par le passé. Des parents des espèces actuelles. J'ai alors demandé si les élèves connaissent tous les membres de leur famille. Forcément, les élèves ont pensé aux personnes encore vivantes : des parents, des frères, des sœurs, des cousins, des grands-parents. « Mais vos grands-parents avaient eux aussi des parents ? ». Ah, oui, on les appelle des arrière grands-parents. « Vous connaissez alors un arbre généalogique de vos familles ? ». La classe savait à quoi correspondait un tel arbre : une histoire d'une famille où on connaît chaque ancêtre dans le passé.

La question était « Est-ce qu'on peut représenter un arbre généalogique des toutes les espèces vivantes sur Terre ? ». Les élèves ont hésité à répondre positivement, vu l'énorme quantité des espèces. Et bien on peut, mais on appelle un tel arbre « un arbre phylogénétique » puisque dans ce cas les ancêtres communs ne sont que hypothétiques, on ne connaît pas leur identité exacte. C'est là où se trouve une différence avec la généalogie, où dans chaque nœud de l'arbre il y a des êtres vivants connus.

### ***Mise en route des investigations***

Pour mieux comprendre une telle représentation nous avons repris la classification emboîtée des séances

2. J'avais projeté un arbre phylogénétique vide, fait exprès pour la collection d'espèces classées. Le même arbre a été distribué aux d'élèves. On s'est rappelé que pour faire chaque boîte nous



avons utilisé un seul caractère. « Quel caractère est partagé par toutes les espèces ? ». La tête et les yeux, c'est donc un caractère qui était possédé par un ancêtre de toutes les espèces de la collection et c'est ainsi un caractère le plus vieux. Pour illustrer ce constat j'ai dessiné une flèche au-dessus de la phylogénie représentant le temps (les élèves étaient déjà familiers avec une telle représentation du temps après des cours d'Histoire). Nous avons donc placé le caractère « tête » à la racine de notre arbre, là où le temps était le plus « ancien ». J'avais demandé ensuite à la classe, par groupe de 4, de placer les espèces de leur collection sur l'arbre, en disant que les espèces qui partagent le plus de caractères vont se trouver sur les branches plus proches et que chaque caractère utilisé pour désigner une boîte doit se retrouver dans un des nœuds de l'arbre.

**Figure 5.** Arbre phylogénétique à remplir

### ***Recueil et mise en commun des résultats***

La classe a eu des difficultés à placer correctement les espèces sur l'arbre, alors rapidement nous avons procédé à une correction collective. Après avoir replacé toutes les espèces et tous les caractères sur notre phylogénie nous avons remonté le temps en partant des feuilles de l'arbre. Par exemple, en prenant le pigeon on a lu dans l'arbre qu'il avait des plumes (sur les ailes), 4 pattes, un squelette interne et une tête. En s'arrêtant au nœud « plumes » j'avais montré une autre branche qui en parlait, celle du canard. Ça voulait dire que le pigeon et le canard, qui font partie du même groupe, celui des oiseaux, avaient un même ancêtre commun dans le passé et cet ancêtre possédait des plumes. Ensuite il a évolué, changé et donné naissance à plein d'autres espèces d'oiseaux.

A la fin j'avais montré un diaporama pour illustrer la parenté des espèces. Je voulais surtout demander : « Est-ce que vous pensez que l'être humain est un descendant des singes ou est-ce que nous avons des ancêtres communs ? ». Les réponses étaient étonnamment correctes, non, on n'est pas des singes (et puis quoi encore !), on doit partager des parents en commun. J'avais donc expliqué que représenter la parenté des hominidés de façon linéaire était fausse et qu'il fallait le faire plutôt avec un arbre phylogénétique.

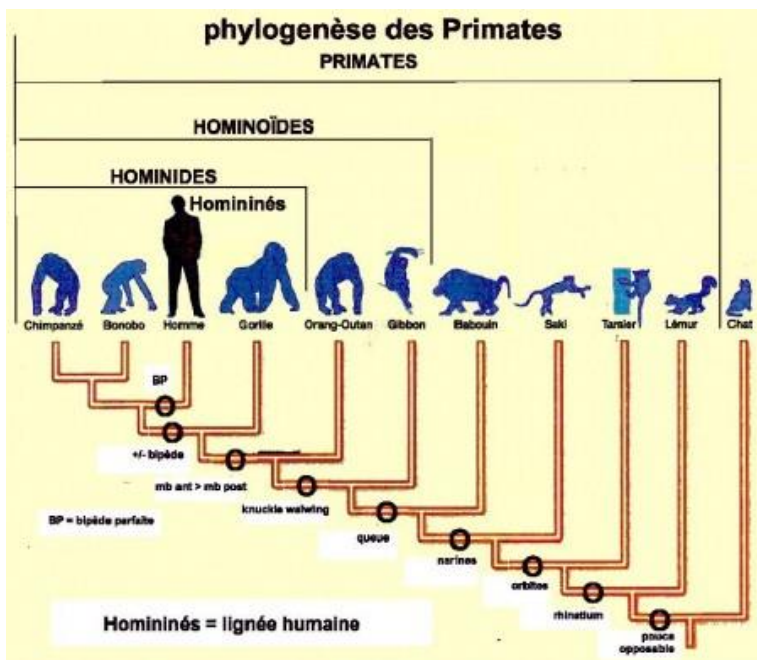


Figure 5. Phylogénie des Primates

Pour conclure, nous avons défini en quoi un arbre phylogénétique rend-il compte de l'histoire du vivant ? <sup>12</sup>

- 1) Il montre un ordre d'apparition de différents caractères : plus un caractère est partagé par les espèces plus il a apparu tôt dans le temps de l'histoire de la vie.
- 2) Il montre des liens de parentés des espèces : plus elles sont proches dans l'arbre, plus grande est leur parenté.
- 3) Il montre la séparation de lignées d'espèces avec ancêtres communs hypothétiques aux nœuds et donc permet d'introduire l'idée de l'évolution.

<sup>12</sup> Guide Belin. « Comprendre et enseigner la classification du vivant ». Sous la direction de Guillaume Lecointre.

## **Séance 6 – Habitats des espèces et chaînes trophiques**

Dans cette séance je voulais replacer les espèces vivantes, dont on parlait isolement sans aucun contexte environnemental, dans leurs milieux naturels.

### ***Recueil des conceptions premières des élèves***

J'avais demandé aux élèves d'imaginer des milieux, des endroits typiques où on pourrait trouver des espèces différentes. Les élèves ont cité plutôt des composantes que des habitats : l'eau, la terre, l'air, ils associaient donc un habitat avec les éléments naturels. Pour mettre un peu de contexte dans les différences entre les espèces j'avais affiché des images des habitats principaux sur Terre : un récif corallien, un désert, un désert polaire, une forêt tropicale, une forêt boréale et une prairie.

### ***Mise en route des investigations***

J'ai demandé à la classe de décrire les images avec les habitats. Ensemble nous avons séparé leurs propositions en vivant (végétaux, animaux) et non-vivant (chaleur, vent, température, neige, rochers, eau, etc.). Ensuite j'ai demandé quels animaux pourraient vivre dans chacun de ces habitats et pourquoi ? Ainsi nous avons évoqué la notion d'adaptation des espèces à leur milieu déjà effleuré dans les séances précédentes. Au début j'avais surtout insisté sur l'interaction entre les espèces vivantes et les composantes non-vivantes des habitats : besoin de respirer, d'avoir de la lumière du soleil, de se cacher dans des rochers, etc. Ensuite, en prenant un exemple d'une forêt et de toutes les espèces que la classe a placé dedans, j'avais demandé : « est-ce qu'il y avait des relations entre ces espèces ? » Comme la réponse avait du mal à venir j'avais demandé : « Est-ce que vous connaissez des chaînes alimentaires entre les espèces ? ». A partir de là les élèves m'ont donné des exemples de « qui mange qui » dans la forêt.

### ***Recueil et mise en commun des résultats***

J'ai présenté un diaporama des 6 habitats et des schémas de chaînes trophiques placés dedans. Il était important de ne pas présenter des chaînes alimentaires en s'appuyant sur le sens de la prédation : le lapin mange l'herbe, le renard mange le lapin. Il fallait indiquer le sens de circulation de la matière et de l'énergie : l'herbe est mangée par le lapin, le lapin est mangé par le renard. De plus les



élèves ont déjà vu les concepts de la matière et de l'énergie, ce qui facilitait la compréhension.

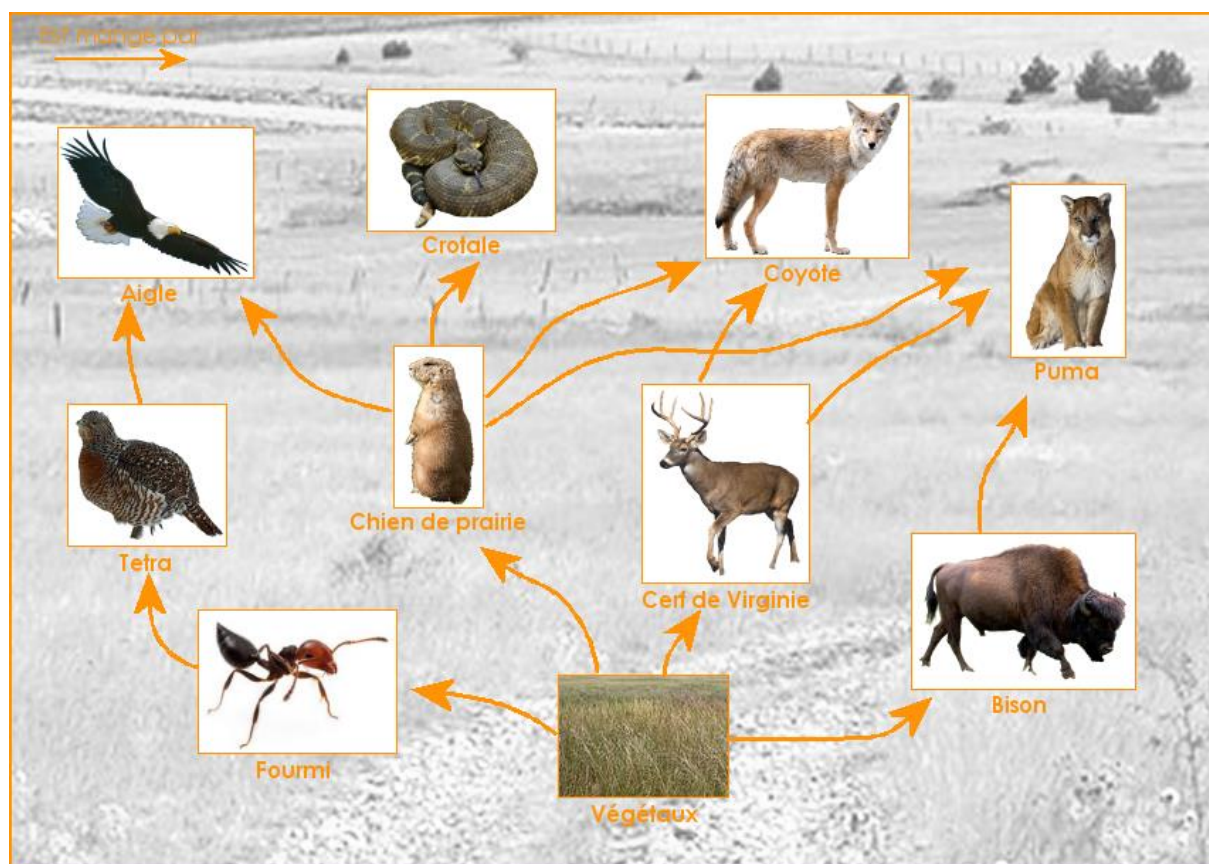


Figure 6. Chaîne trophique d'une prairie d'Amérique du Nord

Pour terminer la séance de façon plus ludique j'avais proposé un jeu sur les chaînes alimentaires que j'avais déjà testé en étant animatrice à l'époque. Nous avons collé le nom d'un animal, d'un végétal ou d'une matière (eau, terre, air) sur le dos de chaque élève. La consigne était simple : créer des chaînes alimentaires en se prenant la main entre ceux qui étaient en interaction trophique. Avec Isabelle on répondait à certaines questions (on faisait partie du jeu également), surtout à celle de « est-ce qu'un tel peut manger un tel ? ». Certains élèves ont passé plus du temps à rigoler du rôle des autres plutôt que de jouer le jeu. Mais au total on a pu construire quelques chaînes trophiques. Chaque chaîne trophique s'est présentée devant le tableau et nous avons discuté des variations et des erreurs possibles.

A la fin de la séance nous avons regardé des images d'espèces des chaînes en les séparant en :

végétaux, qui sont à la base de chaque chaîne alimentaire ; animaux herbivores ; animaux carnivores ; animaux détritivores. La discussion était vive et pleine de questions.

### **Séance 7 – Jeu de piste dans la Grande Galerie de l'Évolution**

Fin juin nous avons organisé une sortie au Muséum sous forme d'un jeu de piste. Lors de la préparation de la visite j'ai été étonné d'apprendre qu'il n'existait qu'un petit jeu de piste rédigé par l'équipe d'animation du Muséum. Je le trouvais trop léger pour le niveau de CM2, de plus je voulais que la visite renforce les connaissances et les compétences que les élèves ont pu acquérir lors des séances en classe : l'observation, la description et la classification des espèces. Il existe également des parcours plus spécifiques sur les dents et les os, mais cela ne correspondait pas non plus à mes besoins.

Avec Isabelle nous nous sommes alors donné rendez-vous à la galerie de l'évolution afin de préparer notre propre parcours. Nous avons repéré 15 espèces à trouver et à classer, ainsi que nous avons noté quelques questions « bonus » pour animer la visite. Le parcours complet se trouve dans les annexes.

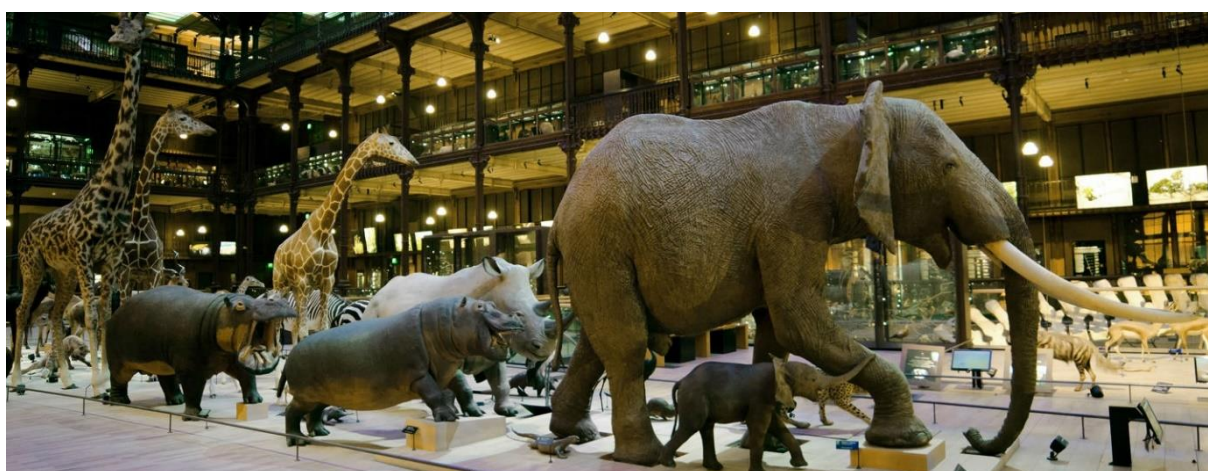


Figure 7. Grande Galerie de l'Évolution, MNHN

Le jour de la visite 3 parents et 2 services civiques de l'école ont accompagné la classe. Les élèves ont été séparés en groupes de 4-5 avec un accompagnateur adulte par groupe (Isabelle et moi compris). Avant d'entrer dans la galerie j'avais expliqué les règles : trouver des espèces représentées



dans la feuille de route, les observer et remplir le tableau de caractères que les élèves connaissaient déjà. Un rappel sur l'importance d'un bon comportement (on n'est pas tout seul au musée) et les groupes sont partis à la découverte de la galerie

Globalement, tous les groupes ont pu finir le parcours à temps (on n'avait que 1h15, les élèves devaient être de retour à l'école pour la cantine). J'avais récupéré les feuilles avec des réponses qu'on allait corriger ensemble à la dernière séance.

Des choses à améliorer

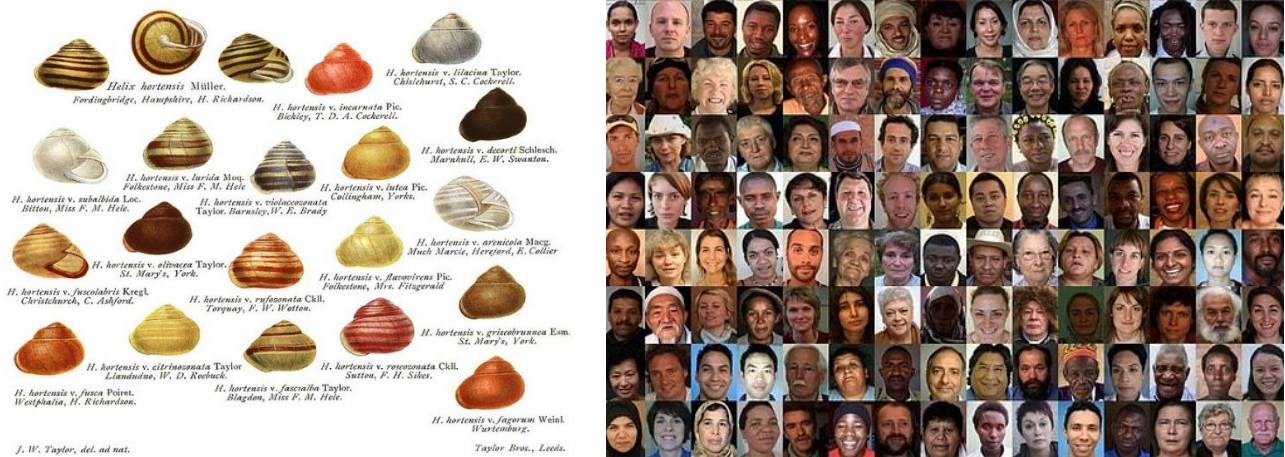
Je suis restées avec mon groupe pendant tout le jeu de piste, mais d'après les remarques des autres accompagnateurs les élèves se sont globalement bien investis dedans, mais préféraient poser des questions à un adulte plutôt que d'aller chercher la réponse seuls. Il fallait donc les relancer à chaque étape.

La collection d'espèces choisies n'était pas forcément la plus simple à classer. Nous nous sommes basés sur les espèces faciles à trouver et pas sur leurs caractères. Parfois il était compliqué d'observer tel ou tel caractère, comme par exemple le recouvrement de la peau d'un requin (des écailles osseuses) ou encore les 4 pattes chez les baleines qui ont des vestiges des membres postérieurs.

### **Séance 8 – Correction du jeu de piste et bilan global**

Tous ensemble nous avons procédé à une description des espèces observées dans la galerie de l'évolution à l'aide d'un diaporama. Nous avons débattu chaque espèce : ces caractères observables, ces particularités et adaptations à son milieu naturel. Finalement je n'ai pas eu le temps de construire une classification emboîtée de ces espèces ce qui n'était pas le plus mal, étant donné les ambiguïtés dans les caractères utilisés.

J'avais terminé la séance sur un rappel de notion d'une espèce en montrant des images des humains et des coquilles d'une espèce d'escargots. Le but était de montrer que même au sein d'une même



espèce les individus peuvent être très différents. Nous avons également parlé des espèces hybrides, ces exceptions de la définition d'une espèce où les individus des espèces différentes ne pouvaient pas se reproduire entre eux.

## Conclusion

Introduire auprès des jeunes la classification du vivant n'a pas seulement pour objectif d'expliquer ce qu'est la systématique qui rend intelligible la diversité du vivant. En passant par une phase de description des caractères des espèces vivantes, les élèves apprennent à être des observateurs curieux de leur environnement. Cela leur permet d'ouvrir les yeux sur la nature qui nous entoure et de s'en émerveiller. Malheureusement, une sortie naturaliste, prévue initialement dans un parc de Belleville n'a pas pu avoir lieu à cause de problèmes de sécurité qui n'était pas suffisante dans un parc public selon le directeur de l'école.

L'évolution est un savoir complexe à expliquer, surtout au niveau de l'école primaire. Un important travail de vulgarisation pédagogique est nécessaire. C'est là-dessus qu'Isabelle avait besoin d'aide, pour ne pas commettre des erreurs classiques de l'enseignement sur ce sujet :

La notion d'un fossile vivant

Certaines espèces actuelles, comme les cœlacanthes ont subi très peu de modifications visibles à l'œil nu ce qui suggère à tort que ces espèces n'auraient pas évolué depuis les temps « fossiles » des espèces qui leur sont très semblables. Une forte ressemblance n'exprime pas une absence de changements génétiques au cours des générations, mais l'expression d'adaptations similaires face à des conditions de vie similaires. On appelle ces espèces des espèces panchroniques.



Figure 2 Fossile d'un cœlacanthe de Jurassique; Individu d'une espèce actuelle de cœlacanthe *Latimeria chalumnae*

### ***La notion du chaînon manquant***

En présentant l'histoire évolutive du vivant il est facile de faire une erreur en disant que des fossiles que nous retrouvons aujourd'hui représentent des ancêtres communs des espèces actuelles et peuvent donc servir de preuve d'une forme du vivant qui relie le passé et le présent. J'avais illustré ce problème par le fait que nous ne disposons pas d'une machine à remonter le temps, et donc que nous ne pouvons pas savoir à quel moment une espèce du passé a changé et donné naissance à un nouvel embranchement d'un arbre phylogénétique et donc à une nouvelle espèce. Il est également possible qu'elle s'est éteinte sans laisser de descendants. Car la question dans la systématique n'est pas « Qui descend de qui » mais « Qui est plus proche de qui? ». Donc nous ne savons pas si le fossile en question est vraiment un ancêtre d'une espèce étant plus apparentée actuellement.

Il est possible d'expliquer ceci en demandant aux élèves de décrire les caractères d'un fossile et de

le classer avec des espèces actuelles, un exercice que je n'ai malheureusement pas eu le temps de faire.

### ***La notion du progrès dans l'évolution***

Le progrès est une notion très anthropique, reflétant un point de vue assez subjectif de l'homme sur les changements subis par des objets en général. Dans un contexte d'évolution cette notion suggère une amélioration et des perfectionnements constants des espèces vivantes. Or les espèces évoluent toutes en même temps mais dans des conditions très différentes, ce qui font d'elles une mosaïque de caractères « primitifs » et de caractères « dérivés ». J'avais donné un exemple d'une taupe qui perd sa vue ou des Cétacés qui n'ont plus de membres postérieurs, ce qui peut être vu comme une dégression et pourtant fait partie de leur chemin d'évolution. Darwin lui-même écrivait « survival of the fittest » et non « of the best » ou « of the strongest ». Ainsi, il n'existe pas de caractères supérieurs et inférieurs.



### **Des suggestions pour l'enseignement de l'évolution**

Suite à mon expérience personnelle et une lecture de plusieurs ouvrages lors de cette formation je voudrais proposer les idées suivantes :

- Intégrer l'explication du concept de l'évolution tout au long de l'enseignement des sciences et techniques plutôt que de l'isoler comme un sujet à part.

- Combiner l'histoire évolutive et la classification de la diversité du vivant au lieu de les séparer.
- Prendre en compte des obstacles liés à la notion des temps géologiques et à la polysémie des mots « évolution » et « classification ».
- Prendre en compte la pluralité du rapport personnel des élèves (social, culturel, religieux).
- Expliquer aux élèves l'intérêt d'étudier de tels concepts.

### **Bilan personnel du projet**

C'était un grand plaisir de se retrouver de nouveau face à un groupe de jeunes esprits, d'éveiller leur curiosité et de partager avec eux des siècles de connaissances scientifiques. C'était aussi une expérience différente de toutes celles que j'ai pu avoir avec des groupes d'enfants, car cette fois je collaborais avec une enseignante, une chose qui manque cruellement aux animateurs périscolaires pour savoir ce que les élèves apprennent en classe afin d'adapter le contenu de leurs ateliers.

J'aimerais que cette formation soit mise plus en valeur parmi les doctorants qui n'osent pas s'y investir par manque de temps ou par la peur d'une perspective de travail avec un public trop jeune et donc impatient et bruyant. Pour y remédier j'avais organisé un « Café scientifique », une conférence hebdomadaire au sein de mon UMR, où j'avais présenté ce travail. J'espère que ce partage d'expérience donnera envie à mes collègues du Muséum de partager leurs propres connaissances scientifiques avec le grand public qui n'y a malheureusement pas forcément accès.

## **Bibliographie**

### **Ouvrages**

Guide Belin de l'enseignement. « Comprendre et enseigner la classification du vivant », sous la direction de Guillaume Lecointre. Editions Belin

Guillaume Lecointre. « L'évolution, question d'actualité ? ». Editions Quae

Corinne Jégou. L'enseignement de l'évolution des espèces vivantes à l'école primaire française.

Rapports au savoir d'enseignants et d'élèves de cycle 3. Education. Université de Provence -Aix-Marseille I, 2009.

Corinne Fortin. "Evolutionary Theory in Secondary Schools: Some Teaching Issues". T. Heams et al. (eds.), Handbook of Evolutionary Thinking in the Sciences, 2015.

### **Sites internet**

La fondation « La Main A La Pâte » : <http://www.fondation-lamap.org/fr>

Dispositif de LAMAP « Accompagnement en sciences et technologie à l'école primaire :

<http://www.fondation-lamap.org/fr/astep>

<http://eduscol.education.fr/>

Evolutionary Knowledge for Everyone : <https://evokeproject.org/>







Humanity strongly depends on biodiversity and services it provides. To prevent the biodiversity loss and to establish sustainable relations with nature humanity has to efficiently manage and protect natural resources. The problem of "what to protect" is not new but became more important than ever and could be resolved by an appropriate use of biodiversity measures. Many indices of biodiversity have been developed in the last four decades, with species being one of the central units. However, evolutionary and ecological studies need a precise description of species' characteristics to best quantify inter-species diversity, as species are not equivalent and exchangeable. First measures taking into account species biological differences were based on species phylogenetic relations and trait values. However, many of them measure a diversity of a set of species, and does not indicate the respective contribution of each species to the diversity of the set. To find a remedy to this issue, other type of measures appeared in early 90's, comparing species through the shared amount of characteristics, but were put aside, erroneously classified as diversity measures too. In this thesis we refer to these measures as species originality indices. A species is original if it possesses unusual trait values compared to all others in a community or if it is distantly related with other species in a community. Thus, the most original species have the greatest contribution to the diversity of that community. In this thesis we sought to demonstrate the benefits of originality metrics, particularly in conservation biology and community ecology. First we review the relation of species originality with concepts of species' diversity and rarity and we compare their related measures. Following theoretical links between originality and diversity measures we propose a practical application of a two-step (and two-scale) originality framework to a real plant species data. Finally, we discuss main pitfalls and advantages related to species data, spatial scale of a study and the choice of an originality measure. Future studies could use originality measures with other entities than species, such as genes or habitats, and therefore broad the extent of biodiversity assessment and conservation.

L'humanité dépend fortement de la biodiversité et des services qu'elle nous fournit. Pour prévenir la perte de biodiversité et établir des relations durables avec la nature, l'humanité doit gérer et protéger des ressources naturelles. Le problème de "what to protect" n'est pas nouveau, mais il est devenu aujourd'hui plus important que jamais et pourrait être résolu par une utilisation appropriée des mesures de la biodiversité. De nombreuses mesures de biodiversité ont été élaborées au cours des quatre dernières décennies avec l'abondance des espèces comme l'une des unités centrales. Cependant, les études en écologie et évolution nécessitent une description précise des caractéristiques des espèces pour quantifier au mieux la diversité interspécifique, car en effet les espèces ne sont pas équivalentes. Les premières mesures tenant compte des différences biologiques entre les espèces étaient fondées sur les relations phylogénétiques et les valeurs de traits des espèces. Cependant, beaucoup d'entre elles mesurent la diversité d'un ensemble d'espèces et n'indiquent pas la contribution de chaque espèce à la diversité de l'ensemble. Comme une solution à ce problème, d'autres types de mesures sont apparus au début des années 90's, comparant les espèces en fonction de ce qu'elles ont en commun, mais elles ont été mises de côté, classées à tort comme mesures de la diversité. Néanmoins, ces mesures donnent une valeur à chaque espèce et non à l'ensemble des espèces. Dans cette thèse, nous appelons ces mesures des indices d'originalité des espèces. Une espèce est originale si elle possède des valeurs de traits inhabituels par rapport à toutes les autres espèces dans une communauté ou si elle est phylogénétiquement éloignée des autres. Ainsi, les espèces les plus originales sont celles qui contribuent le plus à la diversité de cette communauté. Dans cette thèse, nous avons cherché à démontrer des avantages des mesures d'originalité, en particulier en biologie de la conservation et en écologie des communautés. Tout d'abord, nous examinons la relation entre les concepts de l'originalité, de la diversité et de la rareté des espèces et nous comparons leurs mesures associées. Poursuivant des liens théoriques entre les mesures d'originalité et de diversité, nous proposons une application pratique d'indices d'originalité en deux étapes (et à deux échelles) à un jeu de données réel des espèces végétales. Enfin, nous discutons des principaux points forts et faibles liés aux données sur les espèces, à l'échelle spatiale des études et au choix des mesures d'originalité, impliqués dans l'analyse d'originalité. Un outil prometteur, les mesures d'originalité pourraient être utilisés avec d'autres entités que des espèces, tels que les gènes ou les habitats, et donc élargir notre compréhension et la conservation de la biodiversité.