



**HAL**  
open science

# Understand and predict the geographical spread of alpine invasive plant species

Laure Gallien

► **To cite this version:**

Laure Gallien. Understand and predict the geographical spread of alpine invasive plant species. Biodiversity. Université de Grenoble, 2013. English. NNT : 2013GRENV062 . tel-01561063

**HAL Id: tel-01561063**

**<https://theses.hal.science/tel-01561063>**

Submitted on 12 Jul 2017

**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.

## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biodiversité Ecologie Environnement**

Arrêté ministériel : 7 août 2006

Présentée par

**Laure GALLIEN**

Thèse dirigée par **Wilfried THUILLER** et **Sébastien LAVERGNE**

préparée au sein du **Laboratoire d'Ecologie Alpine**  
dans l'**École Doctorale de Chimie et Sciences du Vivant**

# Comprendre et prédire l'expansion géographique des espèces végétales invasives dans les Alpes

Thèse soutenue publiquement le **25 juin 2013**,  
devant le jury composé de :

**M. Stefan DULLINGER**

Professeur à l'Université de Vienne (Autriche), Rapporteur

**M. Jonathan LEVINE**

Professeur à l'ETH de Zürich (Suisse), Rapporteur

**M. John, THOMPSON**

Directeur de recherches au CEFE de Montpellier, Président

**M. Jake ALEXANDER**

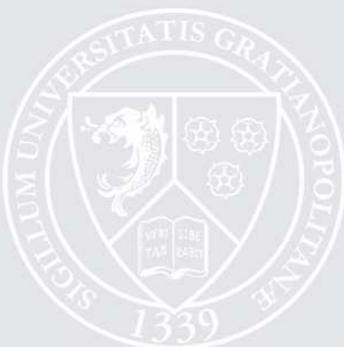
Chargé de recherche à l'ETH de Zürich (Suisse), Examineur

**M. Wilfried THUILLER**

Directeur de recherche au LECA de Grenoble, Directeur de thèse

**M. Sébastien LAVERGNE**

Chargé de recherche au LECA de Grenoble, Directeur de thèse





## REMERCIEMENTS

Pour commencer, je souhaiterais avant tout remercier **Wilfried**. Merci Wil pour m'avoir fait confiance il y a 5 ans, quand tu as accepté de m'encadrer pendant mon master 2. Cette première expérience du travail de recherche a été le déclencheur de ma mono-mania pour l'étude des espèces invasives... Et ensuite avec **Séb**, vous vous êtes impliqué et m'avez appuyé pour que j'obtienne une bourse de thèse, et c'est là que tout a commencé... Merci à vous deux pour avoir cru en moi! Tous les deux, avec vos différences, vous avez su m'encadrer, me guider et... me supporter ☺ Tous les deux vous avez su être patient au quotidien, canaliser ma curiosité et mes instants de dispersion pour me permettre d'extraire les points essentiels de mon travail et m'enseigner la rigueur scientifique. Et puis il y a aussi cet insatiable plaisir que vous avez à parler de science en tout lieu et en toute heure qui a animé ma curiosité sur toujours plus de sujets d'écologie et d'évolution... Merci ! Finalement, vous n'avez pas seulement été des encadrants, mais aussi des amis. Et ce sont ces moments de détente partagés autour d'un mur d'escalade ou d'un barbecue qui ont rendu ce doctorat si agréable. Merci pour ça aussi !

Bien sur, le travail effectué pendant ces années de thèses n'aurait pas eu la même couleur, la même richesse si il n'y avait pas eu toute une équipe de personnes extraordinaires avec qui échanger au quotidien. Il y a bien sur les personnes avec qui j'ai directement collaboré comme **Tamara** (compagne de complexification des idées), **Cécile** (la grimpeuse), **Marta** (« sorry »), **Florent M.**, **Florian A.**, **Loïc C.**, **Martí B.**, **Julien** (l'asticot), **Isabelle B.**, **Delphine R.**... Mais les développement d'une pensée scientifique ne se fait pas seulement dans le cadre de travaux bien précis, elle se fait aussi autour d'un café ou d'une bière quelque soit l'heure de la journée. Et ces échanges n'auraient pas été les mêmes sans **Florian B.**, **Laure** (Mam, Robert et son tutu), **Katja S.**, **Cristina R.**, **Damien G.**, **Laura M.**, **Isabel C.R.**, **Charline Gigi**, **Mathieu F.**, **Julien R.**, **Nico L.**, **Pierre G.**, **Marco A.**, et tant d'autre. C'est réellement grâce à vous tous que je me suis sentie murir au fil du temps, et que j'ai finalement pu prendre confiance en moi. Merci, encore à vous tous de m'avoir soutenu durant les mois critiques de finalisation de la thèse, de l'avoir relu et amélioré, de l'avoir écouté en prime time pour m'entraîner et de l'avoir

réécouté le jour J! Merci encore pour tous ces instants passés à vos côtés, j'espère qu'il y en aura plein d'autre !

En plus du quotidien heureux au LECA, j'ai eu la chance de collaborer avec le conservatoire de Gap (CBNA) et notre cher botaniste de la station alpine (SAJF) **Rolland** (et ses séries Z!) C'est grâce à ces collaborations que de nombreux projets ont pu aboutir. Merci à vous ! Et tout spécialement, je voudrais souligner l'enthousiasme qui a entouré le projet Ambroisie à Gap, grâce à l'implication et la bonne humeur (même dans l'urgence !) de **Noémie, Sébastien, Françoise, Bernard, Juliette, Jérémie et Luc**.

Je souhaiterais remercier plus généralement tous ceux qui ont travaillé ou qui travaillent encore au labo, non-permanents et permanents. Depuis **Irène** et ses câlins du matin, à la bonne humeur du team **Florence-Kim-Nancy**, en passant par les **poulettes de FYSE**, les **fanatiques des plantes alpines** de la SAJF, les **fous des moustiques**, et les **geeks du troisième** ! Merci à tous pour vos petites attentions de tous les jours, qui rendent la vie si agréable !

Finalement, j'aimerais remercier mon compagnon **Florian** qui a toujours su trouver les mots pour me soutenir tout au long de ces années de thèse. C'était un bonheur de t'avoir à mes côtés autant pour nos discussions scientifiques au sommet d'une montagne, que pour tous ces moments de plaisir non scientifique !

# ***TABLE OF CONTENTS***

<b><i>REMERCIEMENTS</i></b> .....	<b><i>i</i></b>
<b><i>TABLE OF CONTENTS</i></b> .....	<b><i>iii</i></b>
<b><i>CONTRIBUTIONS</i></b> .....	<b><i>v</i></b>

## ***INTRODUCTION***

1. Ecological & scientific implications of biological invasion .....	<b>3</b>
1.1. Biological invasion, a major agent of global change .....	<b>3</b>
1.2. Biological invasion, a long lasting global change .....	<b>5</b>
1.3. Biological invasion, an experiment on a global scale .....	<b>7</b>
1.4. Terminology used in this thesis .....	<b>9</b>
2. Understanding invasions from a historical perspective .....	<b>11</b>
2.1. The pioneers of invasion ecology .....	<b>11</b>
2.2. From impact studies towards explanatory science .....	<b>13</b>
2.3. Modern syntheses & future challenges .....	<b>15</b>
3. The role of the environmental filter in plant invasion ecology .....	<b>17</b>
4. The role of the biotic filter in plant invasion ecology .....	<b>23</b>
4.1. Two competitive interaction strategies .....	<b>25</b>
4.2. Methodological issues .....	<b>27</b>
5. The role of the evolutionary filter in plant invasion ecology .....	<b>31</b>
5.1. Why study niche evolution? .....	<b>31</b>
5.2. What drives adaptive evolution during invasion? .....	<b>33</b>
6. Organisation of my work and the study system. ....	<b>37</b>
6.1. Study system: the French Alps .....	<b>37</b>
6.2. Organisation of the work .....	<b>39</b>

## **CHAPTERS**

Chapter 1. Modelling invasive species distribution – where are we now? .....	42
Chapter 2. Invasive species distribution models – taking advantage of their limitations .....	61
Chapter 3. Are biotic interactions influencing invasion success? .....	78
Chapter 4. Disentangling the relative importance of different community assembly rules to understand invasion success in alpine plant communities .....	127
Chapter 5. Evidence for rapid – genetically-based – climatic niche expansion in an invasive weed: <i>ambrosia artemisiifolia</i> L. ....	146

## **SYNTHESIS**

1. Ecological and methodological findings .....	173
1.1. Environmental filtering .....	173
1.2. Biotic filtering .....	176
1.3. Evolutionary filtering .....	180
2. General limitations & perspectives .....	181
2.1. Historical contingencies .....	182
2.2. The missing 4 <sup>th</sup> musketeer: dispersal .....	183
2.3. Impacts of invasions .....	184
3. Personal conclusions .....	186

<b>REFERENCES</b> .....	187
-------------------------	-----

## **APPENDICES**

Appendix 1. Chalmandrier et al. 2013 .....	201
Appendix 2. Münkemüller et al. Submitted .....	205

# CONTRIBUTIONS

## Chapters

- § 1 { **Gallien, L.**, Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) *Predicting potential distributions of invasive species: where to go from here?* Diversity and Distributions, **16**, 331-342.
- §2 { **Gallien, L.**, Douzet, R., Pratte, S., Zimmermann, N. & Thuiller, W. (2012) *Invasive species distribution model - How violating the equilibrium assumption can create new insights.* Global Ecology and Biogeography, **21**, 1126-1136.
- §3 { Thuiller W., **Gallien, L.**, Boulangeat, I., de Bello, F., Münkemüller, T., Roquet-Ruiz, C. & Lavergne, S. (2010) *Resolving Darwin's naturalization conundrum: a quest for evidence.* Diversity and Distributions, **16**, 461-475.
- §3 { Carboni, M., Münkemüller, T., **Gallien, L.**, Lavergne, S., Acosta, A & Thuiller, W. *Darwin's naturalization hypothesis: scale matters in coastal plant communities.* Ecography, **35**, 01-09.
- §3 { **Gallien, L.**, Münkemüller, T., Carboni, M., Chalmandrier, L., Lavergne, S., Thuiller, W. *Choosing the appropriate indices and null model in invasion community ecology – A simulation study.* In prep.
- §4 { **Gallien, L.**, Mazel, F., Lavergne, S., Roquet, C., Renaud, J., Douzet, R. & Thuiller, W. *Disentangling the relative importance of different community assembly rules to understand invasion success in Alpine plant communities.* Submitted.
- §5 { **Gallien, L.**, Thuiller, W., Lainé, J., Boleda, M., Alberto, F., Rioux, D., CBNA, Lavergne, S. et al. *Evidence for rapid, genetically-based, climatic niche expansion in an invasive (annual) weed.* Submitted.

## Appendices

Chalmandrier, L., Münkemüller, T., **Gallien, L.**, de Bello, F., Mazel, F., Lavergne, S. & Thuiller, W. *A family of null models to disentangle habitat filtering vs. biotic interactions from functional diversity patterns.* Journal of Vegetation Science.

Münkemüller, T., **Gallien, L.**, Lavergne, S., Renaud, J., Roquet, C., Dullinger, S., Guisan, A., Lenoir, J., Svenning, J.C., Vittoz, P., Willner, W., Wohlgemuth, T., Zimmermann, N. & Thuiller, W. *Scale decisions can reverse conclusions on community assembly processes.* Submitted to Global Ecology and Biogeography.



# *Introduction*



## **1. ECOLOGICAL & SCIENTIFIC IMPLICATIONS OF BIOLOGICAL INVASION**

*“The chess-board is the world; the pieces are the phenomena of the universe; the rules of the game are what we call the laws of Nature. The player on the other side is hidden from us. We know that his play is always fair, and patient. But also we know, to our cost, that he never overlooks a mistake, or makes the smallest allowance for ignorance.”* Thomas Henry Huxley

### **1.1. BIOLOGICAL INVASION, A MAJOR AGENT OF GLOBAL CHANGE**

A biological invasion is the transportation of an organism to a new and often distant area where its descendants can survive and spread (Elton 1958). It is a well-established fact that biological invasions are a major threat to biodiversity, ecosystem functioning and associated goods and services (Vitousek *et al.* 1997, Mack *et al.* 2000, Walther *et al.* 2009, Stohlgren *et al.* 2011). Furthermore, invasive species can reduce ranges of native species, sometimes to the point of their extinction (notably on island ecosystems; Sax *et al.* 2002, Clavero & Garcia-Berthou 2005, Medina *et al.* 2011), and, thus, ultimately lead to the homogenization of the Earth's biota (Olden *et al.* 2004, Hobbs *et al.* 2006, Vila *et al.* 2011). For instance, among the 680 known recent animal extinctions, the causes of extinction are known for 170 cases, of which 54% were partly caused by invasive species (Clavero & Garcia-Berthou 2005). Additionally, the invasion of ecosystem engineers can result in strong perturbations of ecosystem functioning; for example, invasive pine species providing favourable conditions for intense fires thus indirectly causing increased soil erosion and modified vegetation structure (Richardson & Bond 1991). Economically, the annual losses caused by invasive species are tremendous. They have been estimated at around €300 billion in Australia, Brazil, India, South Africa, United Kingdom and United States (Pimentel *et al.* 2001, 2005) and at €12 billion in Europe (Kettunen *et al.* 2009).

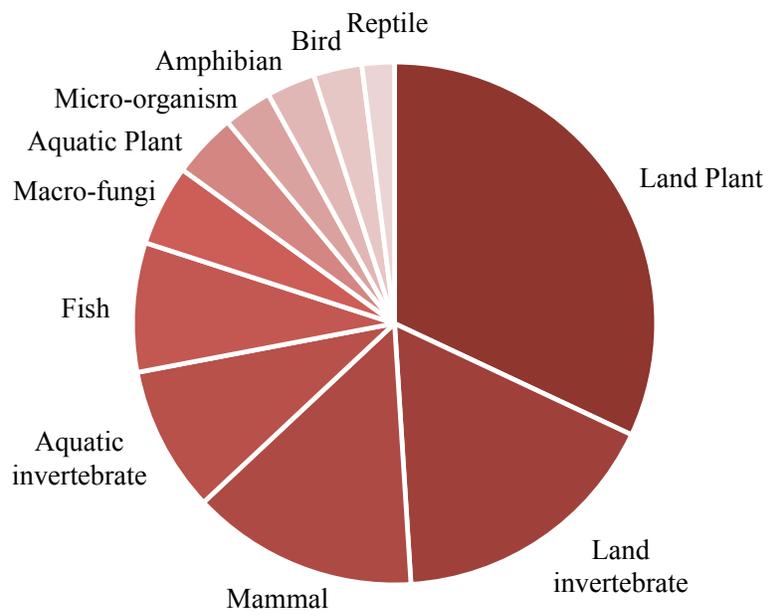


Figure 1. Proportion of the main species groups among 100 of the most invasive species globally according to Lowe *et al.* (2004).

Although invasions are numerous, not all species become invasive when introduced to a new area. Some groups such as plants, invertebrates and mammals contain particularly high numbers of invaders, while others like reptiles and amphibians hold few invaders (Fig. 1). Similarly, some ecosystems, such as islands and Mediterranean coasts, are highly susceptible to invasion, while mountaintops and deserts are more resistant. Among the reasons that make some places more prone to be invaded than others is the fact that in some ecosystems environmental conditions are so extreme that few potential invaders possess the pre-adaptation necessary to survive these conditions. Another often-cited explanation is that some ecosystems have been isolated from particular types of predators or competitors for so long that they are thus particularly vulnerable to them.

Today, due to ever-increasing human transportation and global trade, very few habitats on earth remain free of human-caused biological invasions. Additionally, the establishment and persistence of invaders benefit strongly from landscape fragmentation and land-use changes as human-altered ecosystems generally provide the primary foci for introduced organisms before their further spread (D'antonio & Vitousek 1992, Gonzales-Moreno *et al.* 2013). The impacts of invasive species on native ecosystems can be further enhanced by the third major global threat to biodiversity: climate change. Indeed, rapid climate change can increase the vulnerability of native ecosystems to perturbations, including biological invasions (Walther *et al.* 2009).

## **1.2. BIOLOGICAL INVASION, A LONG LASTING GLOBAL CHANGE**

Natural invasions, or more simply species range changes, have always occurred in the history of life (Sax *et al.* 2005). Large migrations in earlier geological times have occurred as a result of lithospheric movements that have opened land-bridge connections, long distance dispersal that have led to island colonisation or the tracking of shifting favourable climates (see examples in Box 1). During the last 20 million years, many episodes of land mass biota interchanges occurred, essentially as a result of tectonic activity. Over the last



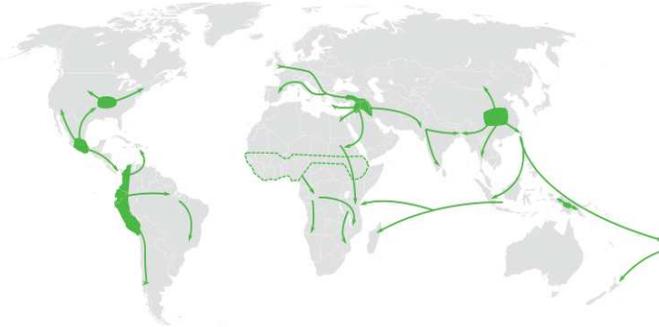
10,000 years, with the end of the Pleistocene ice age, climatic conditions, and thus species' ranges, have changed everywhere on earth (Lodge 1993); although many species have recolonized areas they previously occupied (i.e. before glaciation).

In contrast to these natural invasions, ancient and recent human activities are promoting new invasions at an unprecedented rate, notably through human movement and ecosystem modification. Plant domestication, exportation, or accidental introductions provide good examples of human-mediated invasions that can have strong positive or negative influences on native ecosystems (see examples in Box 2). However, it is not always simple to identify whether the invasion has been human-mediated or to define what can be considered as human intervention (e.g. land use or human induced climate change), and to establish for how far back in time invasion should be considered (e.g. Neolithic age and the widespread crops, 1492 and the discovery of the Americas; Rikli 1903).

### **1.3. BIOLOGICAL INVASION, AN EXPERIMENT ON A GLOBAL SCALE**

Interestingly, besides being a global threat to biodiversity, biological invasions have the potential to help us understand fundamental research questions in ecology and evolution (Hierro *et al.* 2005, Holt 2005). They actually provide a unique opportunity to observe large-scale ecological and evolutionary changes following the introduction of new species to natural ecosystems (e.g. Callaway & Maron 2006). It has also led to wide recognition that rapid evolutionary processes occur and alter the speed of species range changes and their integration into natural communities (Prentis *et al.* 2008).

## BOX 2. EXAMPLES OF HUMAN MEDIATED INVASIONS



*Legend: Map of the world showing approximate centres of origin of agriculture and its spread in prehistory: the Fertile Crescent (11,000 BP), the Yangtze and Yellow River basins (9,000 BP), the New Guinea Highlands (7,500 BP), Central Mexico (4,500 BP), Northern South America (4500 BP), sub-Saharan Africa (4500 BP), eastern USA (3500 BP).*

### *Neolithic founder crops.*

10,000 years ago in the Fertile Crescent, the Neolithic agricultural revolution rose with the domestication of 8 plant species (3 cereals, 4 pulses, 1 flax). These domesticated species combined with specialised food-crop cultivation allowed surplus food-production and thus an increase in the density of human populations. These species are now cultivated all over the world and are essential to food production (Diamond & Bellwood 2003).

### *Rabbit, fox, & myxomatosis.*

Among the most famous examples of dramatic consequences following uncontrolled introduction is the history of rabbits in Australia (*Oryctolagus cuniculus*). Rabbits were first introduced from Europe in 1788 as food resource and bred in cages. In 1859, 24 rabbits were released for hunting purpose. Following a huge demographic explosion, rabbits expanded their range at about 130km per years, growing their population to 104 billions individuals in 1944. Their spread caused devastating ecological damages: by inhibiting plant regeneration, it has led to serious erosion problems and thus threatened many native mammals. To control their population, foxes and cats (natural predators of rabbits in their native range) were introduced. But these two species dramatically changed the predator-prey balance, and exerted a strong influence on native prey species that were at low densities. Following the failure of such a strategy, in 1950 the *Myxoma* virus was deliberately released in wild rabbit populations, causing the death of about 500 billions of individuals. However, genetic resistance quickly evolved, allowing the population to recover. Today, the virus has been accidentally introduced in the rest of the world, affecting many pets (only a vaccine exists now, but no cure yet).

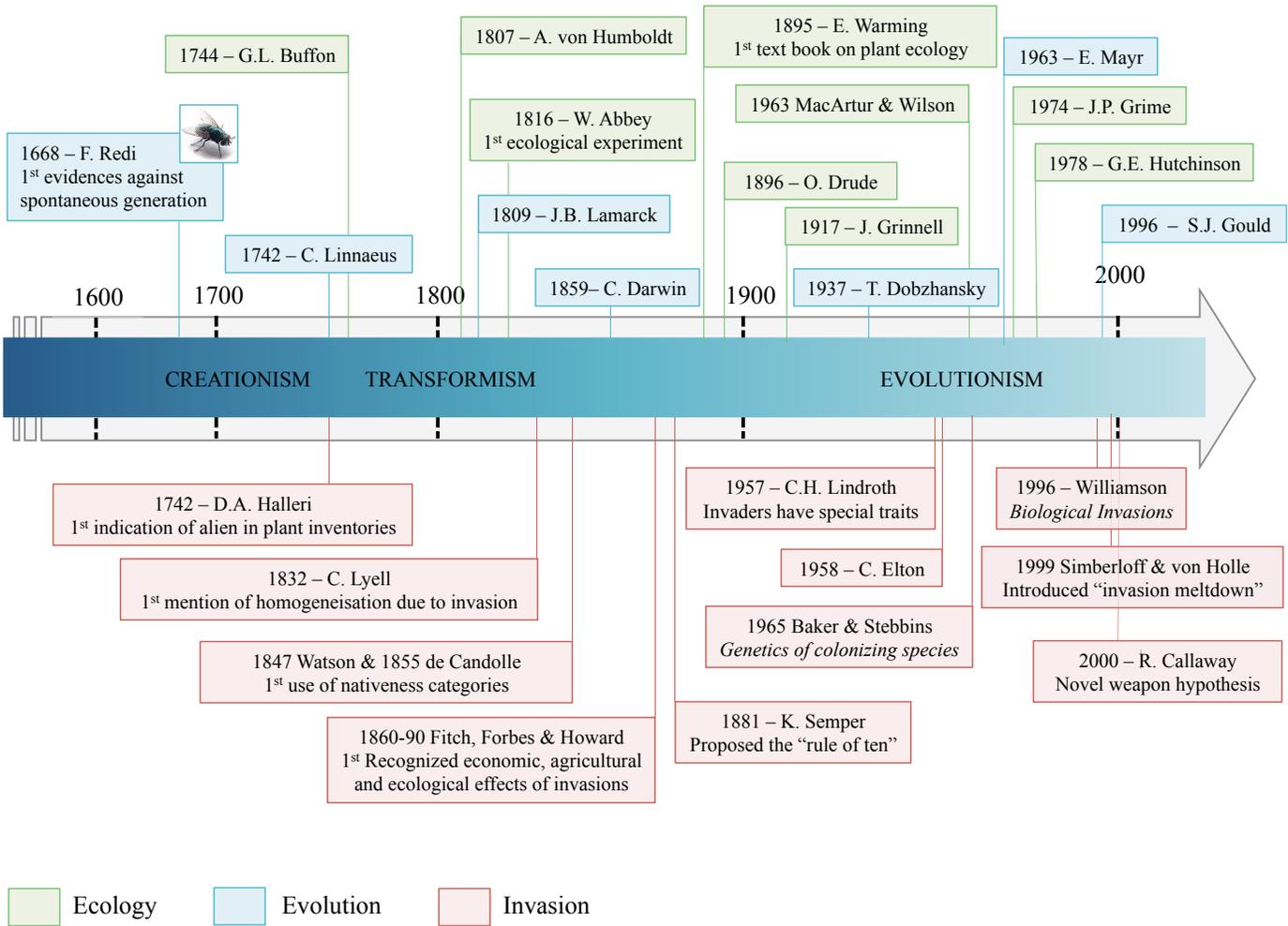


Species invasions thus constitute an unprecedented number of natural experiments that are replicated across taxa (different species that are introduced into one region), across space (species that are introduced into multiple regions), and across ecological systems (species that are introduced into different habitats). Invasions thus offer a promising alternative to complex and cost-intensive experiments and help us to answer crucial questions, such as: *What drives species distribution in space and time? What drives species coexistence? What drives adaptive evolution?*

#### **1.4. TERMINOLOGY USED IN THIS THESIS**

In the context of my PhD, I have focused on recent plant invasions (after 1492) that are both “natural” and “human-mediated”. To describe the different stages of invasion, I will use the terminology of Richardson *et al.* (2000): casual, naturalised and invasive (regardless of whether they are pests). The term ‘casual’ is used to refer to species or populations that have been introduced but are unable to sustain viable populations without new introductions (i.e. sink populations). Naturalised species or populations are able to maintain population sizes without new immigrants, but do not spread from ruderal areas to natural ecosystems. Invasive species or populations are the ones able to colonise natural communities. This terminology can obviously be applied to both species and population levels of organisation. Indeed, within a species, different populations are not necessarily at the same stage of invasion due to varying colonisation velocity in heterogeneous environments and resistance of native communities (Theoharides & Dukes 2007).

BOX 3. A SHORT HISTORY OF INVASION ECOLOGY.



This figure presents key concepts as well as some key scientists in ecology (green boxes) and evolution (blue boxes), who influenced the development of ecological and evolutionary theories in the context of biological invasions (red boxes). For simplicity's sake, only key contributions to invasion ecology and evolution are indicated in the boxes

## **2. UNDERSTANDING INVASIONS FROM A HISTORICAL PERSPECTIVE**

*Nowadays we live in a very explosive world, and while we may not know where or when the next outburst will be, we might hope to find ways of stopping it or at any rate damping down its force. It is not just nuclear bombs and wars that threaten us [...]. An ecological explosion means the enormous increase in numbers of some kind of living organism—it may be an infectious virus like influenza, or a bacterium like bubonic plague, or a fungus like that of the potato disease, a green plant like the prickly pear, or an animal like the grey squirrel.* Charles S. Elton

### **2.1. THE PIONEERS OF INVASION ECOLOGY**

In Europe, the notion of biological invasion has existed since millenaries. Early reports date back to the Roman Empire when natural philosophers, such as Pliny the Elder, were already writing about the severe problem of rabbit invasion in the Balearic Islands (EEA 2012). More recently, in the 17<sup>th</sup> century, a renewed interest in biological invasions was triggered by the accumulation of biogeographic studies and checklists of plant and animal species, in which human-introduced species were commonly marked with an asterisk (e.g. Halleri 1742, see also Chew & Hamilton 2011). However, over the following 200 years and until Elton's work (1958), the study of the ecological and evolutionary processes driving invasion was neglected (see Box 3). The tardiness of the theoretical developments in invasion ecology and evolution can be explained by three historical components. First, invasive species had to be spread via repeated human-mediated long distance movements that were not frequent during this 200-years period. Second, in order to understand why there could be a sudden invasion and replacement of native biota by a species that was originally absent, researchers had to step back from the established and fixed Christian vision (i.e. that species distribution and creation are of divine origin),



and accept that the Earth's biota had a much longer history than previously established. Such an evolution in thinking was notably set in motion by Linnaeus (1707-78) who proposed that species within a genus might have arisen after the creation of earth through hybridization, and by Buffon (1788) who estimated the Earth's age to be 75,000 years old (instead of the 6,000 years taught by the Bible). This estimated age, made possible fossil interpretations and the recognition of extinct species. Third, the main focus of concerns about invasive species was their ecological and economical impacts. For example, Bartam (1669-1777) noticed that some invasive plants had negative impacts on the native biota and that some were difficult to control. Lyell (1832) also noticed that native species extinctions could be caused by habitat destruction, species invasions and hunting (cited in Wilkinson *et al.* 2002). Hooker (1867) wrote that European plants were rapidly replacing New Zealand native flora. Finally, the economic threat that invasive species posed to agriculture was also a major motivation for studying biological invasions and developing the first management plans for invasive species (e.g. Fitch 1861, Forbes 1883, Howard 1898, see also Inderjit *et al.* 2005). Interestingly, even though the original concern about invasions was their impacts on native species, the question of how far what invaders are able to disturb native ecosystems, remains largely unsolved even today (Sax *et al.* 2007). For some (or all) of the reasons listed above, ecological and evolutionary mechanisms have only been a centre of interest for the last 50 years or so.

## **2.2. FROM IMPACT STUDIES TOWARDS EXPLANATORY SCIENCE**

At the end of the 19<sup>th</sup> century, developments in ecology and evolution theories permitted scientific interests to be redirected towards processes and a better understanding of the spatio-temporal dynamics of invasions. The new theory of evolution developed by Darwin (1859) and Wallace (1870) allowed for more comprehensive investigations of the influence of evolution on species' invasion (e.g. Darwin's Naturalisation hypothesis, Diez *et al.* 2008). This, together with



the development of biogeography (Humboldt 1807, de Candolle 1839, Wallace 1876), landscape and community ecology (Warming 1909), and plant succession (Clements 1916), has led to a better understanding of the interactions between dispersal, abiotic and biotic drivers of invasion.

At the beginning of the 20<sup>th</sup> century, Rickly (1903) and Thellung (1912) sketched the first classification of invasion stages (based on the seminal work of Watson 1847 and de Candolle 1855). The development of the abiotic niche concept followed (Grinnell 1917), and then the discovery that invaders had specific life-history characteristics (Lindroth 1957) attracted more scientific interests towards invasion studies. Elton (1958) was among the first scientists who noticed that species that were introduced long ago could suddenly show population outbreaks, and thus needed to be studied (e.g. the Colorado potato beetle demographic explosion 300 years after the introduction of potatoes).

### **2.3. MODERN SYNTHESSES & FUTURE CHALLENGES**

Nowadays, it is widely recognised that four major processes influence the success of introduced species (sometimes termed filters to invasion): dispersal, environmental filtering, biotic interactions, and rapid evolution. Long distance dispersal is required for the colonisation of new areas, where the species can survive and maintain viable populations if both the abiotic (Grinnell 1917) and biotic (Elton 1958) conditions make the local habitat suitable. These two latter components define the ecological niche of the species (Hutchinson 1957). However, as all individuals of a species are not genetically identical, this ecological niche can evolve through time and space (Baker & Stebbins 1965).

Traditionally, the four major filters of invasions have been studied independently from each other using either experimentation or by analysing field data. Although complementary, these two approaches have not usually been combined. Experimentation was mainly used to study biotic interactions between invasive and native species, because of the difficulty of taking dispersal or a large range of abiotic conditions into account (e.g. Levine 2000,



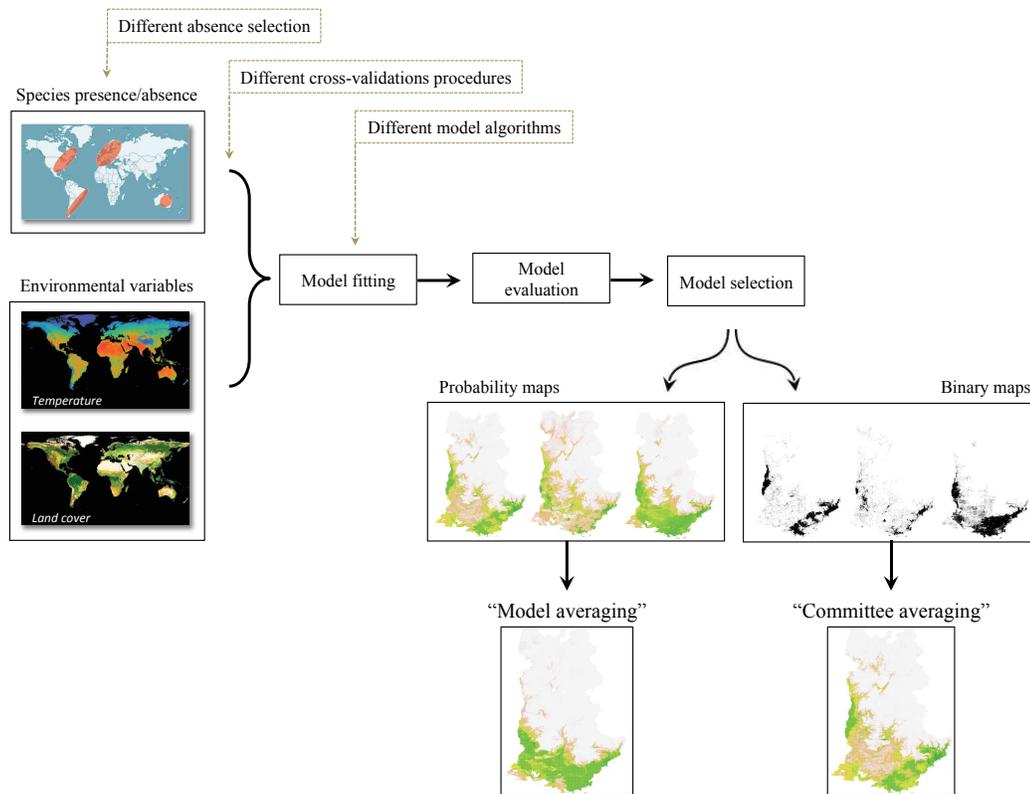
Fargione *et al.* 2003). Instead, field data was used to understand the dispersal and abiotic filters of invasion, because biotic interactions are difficult to measure directly in the field (see Shea & Chesson 2002). However, in order to understand the overall process of invasion, we need to investigate how these filters interact and a combined approach using both experiments and field data is crucial.

In my PhD thesis the objective was to overcome these limitations by explicitly considering the filters that take place in the adventive region, after the introduction of invasive species (i.e. environmental, biotic, and evolutionary filters) and in particular, by considering their interactions. In order to achieve this aim, experimental approaches were combined with the analysis of field data, predictive biogeography and process-based simulation modelling. The role of these filters in invasion ecology is examined in more detail below.

### **3. THE ROLE OF THE ENVIRONMENTAL FILTER IN PLANT INVASION ECOLOGY**

Once a species has been introduced to a new location, its survival depends first on whether the environmental conditions are favourable or not. The so-called environmental filter relies directly on the concept of the *fundamental niche*, which can be defined as all states of the environment that enable positive population growth rates (Hutchinson 1957, Pulliam 2000). Today, finding appropriate and efficient ways to identify this fundamental niche for invasive species is crucial when addressing the following major ecological questions: Do invaders have particular niches that allow them to be especially successful and do these niches match the local conditions at their sites of introduction? If so, which species are at high risk of becoming invasive? Where are the potential hotspots of invasion today? And where will they be in the future?

## BOX 4. ENSEMBLE MODELLING METHODS

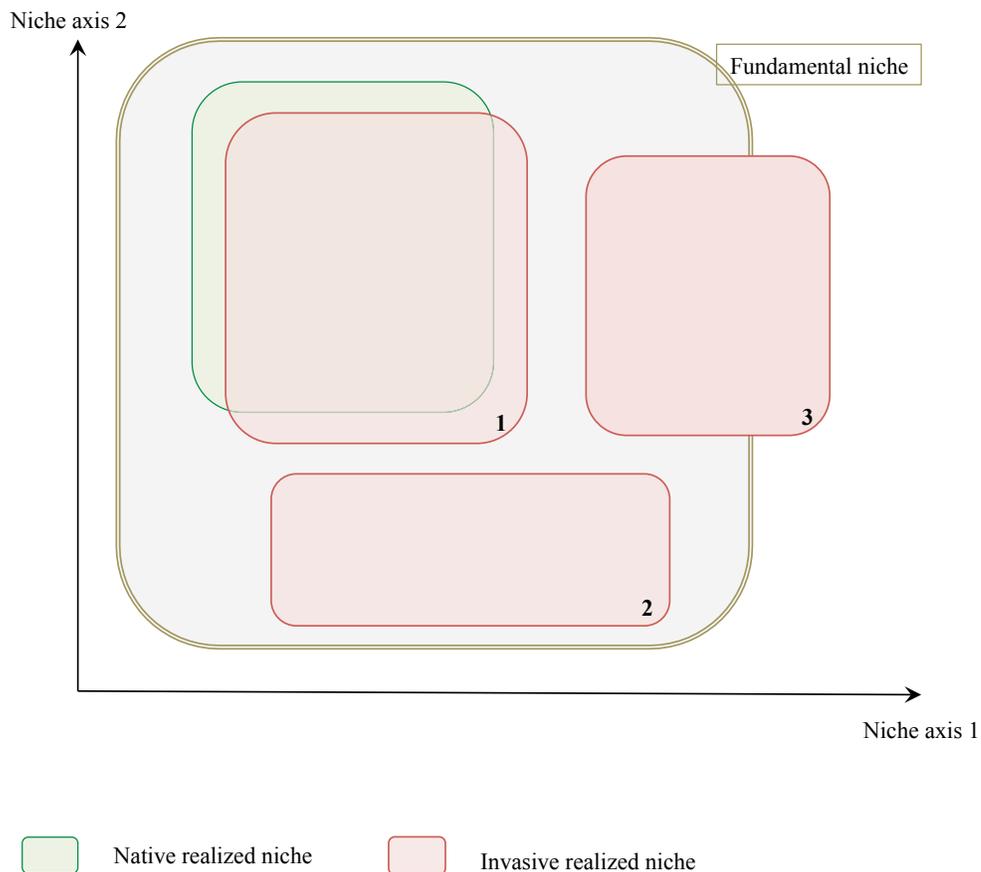


Ensemble modelling methods (Araujo & New 2007) are based on the use of different model algorithms (e.g. regressions, classification trees, machine learning). The rationale of ensemble forecasting is that different algorithms have different levels of accuracy under different circumstances and there is no single perfect algorithm (Elith *et al.* 2006). The model averaging method is an average of predicted probability maps of species presences from the different algorithms. In the committee averaging method, probability maps are not averaged but instead are transformed into binary maps, which are then averaged to obtain one single map of the final output. In other words, each model ‘votes’ in each site, forecasting a species’ presence or not. The main advantage of the committee averaging method is the use of ‘comparable outputs’ (binary presence–absences) instead of the raw algorithm outputs (continuous probabilities) that do not necessarily have the same meaning or the same range of variation. This method can easily incorporate the use of: (1) multiple selections of pseudo-absence data (minimizing the bias due to a specific set of selected pseudo-absences), (2) several repetitions of cross-validation procedures (calibration and evaluation procedures are repeatedly carried out on different sub-datasets), and (3) various model algorithms. For both methods, it is also possible to keep only the best performing models (*model selection*) for the final output by setting a selection threshold based on predictive accuracy metrics.

Performing experiments in controlled conditions is one possible way of investigating the physiological characteristics that limit species survival. Although useful for disentangling the factors responsible for the presence of the species in a given environment, experiments are time-consuming and cannot be applied to complex systems, such as natural landscape, where many environmental variables are intertwined. On the other hand, models that can statistically relate species' occurrences to environmental variables in order to identify suitable environmental conditions (also called Habitat Suitability Models, HSMs; Guisan & Thuiller 2005) can cope with the multiplicity of the environmental conditions and provide broad-scale screenings. Nonetheless, statistical models also have their limits as they rely on observed species occurrences which are implicitly influenced by historical contingencies, demographic dynamics, and biotic interactions. However, since HSMs have the advantage of being rapid and easily applicable to a large number of species (especially worldwide distributed invaders) this tool was chosen as the most suitable for this PhD.

Modelling invasive species niches is not straightforward. Although methodological improvements have been proposed to overcome single model algorithm deficiencies, such as ensemble forecasting (Thuiller 2004, Marmion *et al.* 2009; see Box 4), several technical issues remain. So far, researchers have either calibrated models using the species' native range to extrapolate the relationships into the adventive range (e.g. Beerling *et al.* 1995, Peterson *et al.* 2003, Ibanez *et al.* 2009) or simply calibrated the model in the adventive range with the assumption that the environments of the native and the adventive ranges must to be similar to enable a successful invasion to occur (e.g. Panetta & Mitchell 1991, Scott & Panetta 1993). These two different strategies for calibrating HSMs are prone to predict substantial amounts of false presences and false absences in the invaded range, at least for the three following cases (discussed in more detail below): (1) when there are environmental differences between the native and adventive regions, (2) when the invader distribution did not reach equilibrium in its adventive range, and eventually (3) when there is rapid evolution of the invader's niche in the adventive range.

## BOX 5. DIFFERENT REALIZED NICHES



This figure represent the *realized niche dilemma* in predicting invasion risk based on habitat suitability models. Indeed, when comparing native (in green) and invaded (red) regional niches, three main types of situation can occur:

- Case 1: The realized niche in the invaded range is similar to the one in the native range. This situation can occur when environmental conditions and biotic interactions are similar in both ranges.
- Case 2: The realized niche in the invaded range may be very different from the one in the native range, but still lie in the fundamental niche. This situation can occur when the two ranges show differences in environmental conditions and/or biotic interactions (e.g. enemy release).
- Case 3: The realized niche in the invaded range may be partially outside the fundamental niche of the species because of rapid genetic adaptation.

It can be noted that estimating a “global” niche by using all realised niches can allow buffering the regional effects of biotic interactions, and thus provide a niche estimate that is closer to the fundamental niche of the species (i.e. environmental niche only).

*(1) Environmental conditions differ across regions*

In cases where the model is calibrated using occurrence data from a region that is not the one of interest (e.g. calibration in native range and predictions on adventive range), biased predictions can occur if there are environmental differences between the two regions. For example, it is possible that different variables limit species presences in the two ranges. Thus, even though the model performs well in the calibration range this is not necessarily the case in the adventive region (e.g. when there is a new habitat type in the adventive range). Another example is that of environmental extrapolation. Even in a case where both the native and adventive range share the same environmental filters, it is still possible that the range of values for these variables and their interactions differ between the regions, thus leading to predictions that are not consistent with the observed data. To overcome such problems, a solution for worldwide invaders is to use all known occurrences to calibrate the model (from both the native and invaded ranges), in other words to estimate a *global niche* (Gallien *et al.* 2010, Broennimann & Guisan 2008). This is particularly interesting when not only abiotic but also biotic factors limit species distribution, because combining all *realised niches* from different regions of the world can buffer single region limitations and thus help us to get closer to the fundamental niche (see Box 5).

*(2) The invader is not at equilibrium in the adventive region*

A non-equilibrium situation occurs when a species is present on unfavourable sites and/or is not occupying all favourable sites (omitting natural source-sink dynamics). Such situations are common for invasive species in their adventive ranges, as the species can be frequently transported to unsuitable sites (e.g. harbours, road sides), and when the process of regional colonisation to favourable sites has not yet finished (due to dispersal limitations or/and time constraints). Therefore, when predictive models are calibrated directly using the adventive range, presences in unfavourable sites can bias the niche estimation towards the environmental conditions of the introduction sites, while absences



in favourable sites can cause underestimation of the niche in regions that have not yet been explored by the invader.

### *(3) Rapid niche evolution in the adventive region*

Rapid adaptive evolution has been shown to occur frequently during invasions (e.g. Lavergne & Molofsky 2007, Duglosh & Parker 2008, Urban *et al.* 2008). Even if there have not been any descriptions of rapid environmental niche evolution on a genetic basis so far, several studies have suggested that it is possible (e.g. Broennimann *et al.* 2007, Gallagher *et al.* 2010, but see Petitpierre *et al.* 2012). In such cases, calibrating the model with occurrences from native range only would fail to identify favourable environmental conditions in the adventive range (see Box 5, case 3). To overcome this problem, one could calibrate the model with all known occurrences and make sure that the model also includes occurrences from the invaded range where rapid adaptation is ongoing (Gallien *et al.* 2010).

## **4. THE ROLE OF THE BIOTIC FILTER IN PLANT INVASION ECOLOGY**

Biotic interactions are the next filter to biological invasions and influence their establishment phase in natural communities. They are not easy to capture and model, especially at a regional scale where a large number of species could potentially interact. Currently, there is a lack of basic empirical and theoretical knowledge of how biotic interactions shape invasive and native species' geographic distributions (see Lavergne *et al.* 2010 for a recent review). Historically, most studies characterizing biotic interactions focused on both intra-guild interactions such as competition and facilitation (e.g. plant-plant interaction; Chesson 2000, Callaway *et al.* 2002, Tilman 2004), and on species network dynamics (e.g. engineering effects of keystone species, Jones *et al.* 1994). More than twenty hypotheses linking biotic interactions to plant invasion



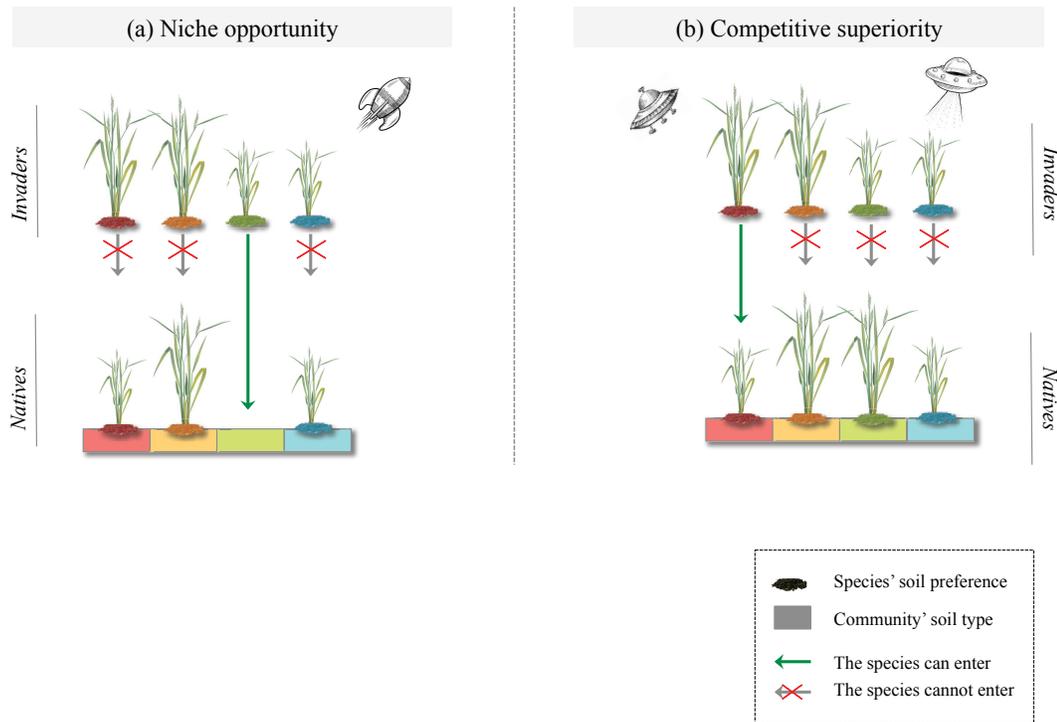
have now been listed (Mitchell *et al.* 2006). Because it is essential to understand the processes occurring within one trophic level well before it is possible to integrate multiple levels, the focus here has been on plant species interactions and specifically on competition.

In particular, I have studied competitive interactions between invaders and native communities in the context of two long ongoing questions in invasion ecology: What makes a successful invader (i.e. *species invasiveness*; Rejmanek 2005, Pysek & Richardson 2007)? And, what makes a community resistant to invasion (i.e. *community invasibility*; Davis *et al.* 2000, Levine *et al.* 2004, Tilman 2004, Richardson *et al.* 2005)? It is now increasingly accepted that the link between the characteristics of the introduced species and the ones of the recipient communities is of prime importance (Lodge 1993, Blumenthal 2006, MacDougall *et al.* 2009). However, it remains unclear as to whether competitive interactions, in particular, play a prominent role in driving the establishment success of an invader.

#### 4.1. TWO COMPETITIVE INTERACTION STRATEGIES

It is possible for an invader to take one of two major strategies in order to be successful in establishing itself when faced with competitive interactions (review in Thuiller *et al.* 2010). On the one hand, the invader can occupy an empty niche (e.g. an unused resource at the community level) regardless whether it is a good or bad competitor (MacArthur 1970, Hierro *et al.* 2005). This strategy, also called *niche opportunity*, is possible where there are niche differences between the invader and the native species. The niche opportunity strategy can be linked to the hypothesis made by Darwin (1859, based on de Candolle's observations), in which he suggested that immigrant species were more likely to become naturalized when they belong to genera that have no native species in the region. Darwin's hypothesis relied on the assumption that closely related species share more similar functional traits and thus use more similar resources than distantly related ones. On the other hand, the invader can succeed in establishing itself by being a better competitor than at least one other

## BOX 6. NICHE OPPORTUNITIES vs. FITNESS DIFFERENCES



This figure illustrates the two possible strategies for an invader to enter a plant community.

(a) In the first case, the invader can only enter when a favourable niche is empty (i.e. niche opportunity).

(b) In the second case, the invader can only enter when it is a better competitor than the native species that occupies the same niche as the invader (i.e. competitive superiority).

native species. In this case the invader can replace the inferior native via *competitive exclusion*: In other words the invader has a positive fitness difference in comparison to the native (e.g. differences in abilities to draw on common limiting resources or in predator susceptibility, MacDougall *et al.* 2009; see Box 6).

Up to now, even though these questions have been studied intensively, no final consensus has been reached on the main drivers of invasion success (see examples in Levine & D'Antonio 1999, Lonsdale 1999, Shea & Chesson 2002, Mitchell *et al.* 2006, Proches *et al.* 2008). Here I argue that there are three major methodological issues that can help to explain the inconsistency in the results of these studies.

## 4.2. METHODOLOGICAL ISSUES

### *(1) Performance differences in data, metrics and statistical tests*

To identify the different strategies that lead to invasion success, a wide variety of data (e.g. taxonomic, functional, phylogenetic), metrics (e.g. diversity indices, alpha-niche metrics) and tests (e.g. null models, regressions, t-test) have been applied. However, it remains unclear as to whether these data, metrics and tests are equally efficient at identifying the two strategies. Indeed, when using taxonomic, functional or phylogenetic data, different implicit assumptions are involved. For example, taxonomic richness relies on the assumption that the richer a community is, the more the niches within it are filled (see Shea & Chesson 2002). Using phylogenetic distances between species relies on the assumption that closely related species share similar niches, and thus assumes that species niche differences show a phylogenetic signal along the phylogeny (Thuiller *et al.* 2010, Mouquet *et al.* 2012). Additionally, to date there is no study for testing the performance of the different indices and statistical tests in



revealing the different invasion strategies (note that a study has been done for native assembly indices, Münkemüller *et al.* 2012).

*(2) The two competition strategies should be tested simultaneously*

Usually, only one of the two invasion strategies detailed above (i.e. niche opportunity vs. competitive exclusion) is explicitly tested. This is an important issue as they can act at different stages of invasion, and they can also act differently within different groups of species. For example, it is possible that niche opportunity and environmental filtering could prevail at early stages of invasion when the invader occupies disturbed habitats (i.e. transport and colonization stages), and that the importance of competitive ability increases when the invader is establishing itself in natural communities (i.e. establishment and spread stages, Theoharides & Dukes 2007). As a second example, it seems probable that all species of a community do not necessarily compete for the same resources (e.g. light interception or soil nutrients) and thus different types of interactions can prevail across different types of organisms. In this way, for light interception, *niche opportunity* is probably of prime importance for the coexistence of different growth forms (e.g. herbaceous vs. woody species), while *competitive ability* could be expected to a key driver in the outcome of assembly between woody species (e.g. Kunstler *et al.* 2012, but see also Fargione *et al.* 2003).

*(3) Interactions between environmental and competition filters*

Because environmental and biotic filters are often thought to be independent, the role of the niche in determining community membership is underestimated (HilleRisLambers *et al.* 2012, but see Graham *et al.* 2009 for an example with native species assemblages). In fact, the position of the community within the environmental niche of the invasive species is often ignored. Nevertheless, the importance of competition may vary according to the strength of the environmental stress (Körner 1999, Callaway *et al.* 2002,



see also Welden & Slauson 1986, Kikvidze *et al.* 2011). It is suspected that the importance of the competition filter is lower at the niche edge where physiological constraints limit species distribution (e.g. in cold alpine environments, Callaway & Walker 1997, De Roy *et al.* 2013), while it can be higher at another niche edge when the distribution of the focal species is limited by negative biotic interaction with another species (e.g. presence of a predator, Wisz *et al.* 2013). Moreover, because ecological patterns are inherently scale-dependent, the resolution at which communities are sampled can have a major impact on the detection of the competition process (Huston 1999, Willis & Whittaker 2002, Qian & Kissling 2010). Theoretically, one can anticipate more competition at finer resolutions where species co-exist (e.g. community scale) than at coarser resolutions where species are solely co-occurring. Some studies have indeed shown that both patterns of competition and environmental filtering between native species can appear within the same system, but at different spatial resolutions (Cavender-Bares *et al.* 2006, Swenson *et al.* 2006, Carboni *et al.* 2012).

## **5. THE ROLE OF THE EVOLUTIONARY FILTER IN PLANT INVASION ECOLOGY**

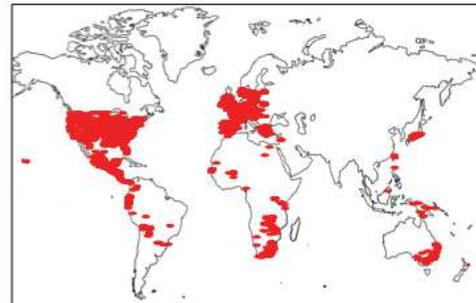
### **5.1. WHY STUDY NICHE EVOLUTION?**

The individuals belonging to any species are not genetically identical and this means that their ecological niche can evolve in space and time. Although niche evolution applies to both abiotic and biotic niches, here, the focus will essentially be on the evolution of the climatic niche, because this niche is the easiest to handle in the context of plant invasions. In fact, the existence of climatic niche shifts has been demonstrated in a number of invasive organisms, by comparing their observed distribution in the new range with the one predicted from the native range (Broennimann *et al.* 2007, Gallagher *et al.* 2010, but see PetitPierre *et al.* 2012). Such shifts, however, only reveal changes

## BOX 7. AMBROSIA ARTEMISIIFOLIA L.



*A. artemisiifolia*'s seed



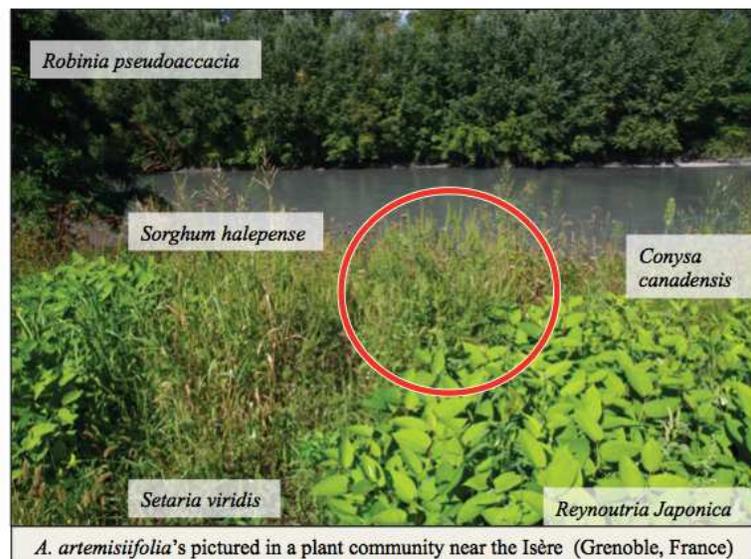
*A. artemisiifolia*'s distribution from GBIF database



*A. artemisiifolia*'s flower



*A. artemisiifolia*'s in a crop



*A. artemisiifolia*'s pictured in a plant community near the Isère (Grenoble, France)

*Ambrosia artemisiifolia* L. (Asteraceae family) is an annual weed native from North America, and has several characteristics that make it a dangerous invader:

- Large seed banks: seed dormancy up to 35 years (Brandes & Nitzsches 2007).
- Long seed dispersal: anthropochorous (usually transported with soil), and hydrochorous.
- High seed production: ~2000 seeds per individual.
- High pollen production (anemogamous), which is very allergenic for human (Ziska *et al.* 2011).

This species is usually common in disturbed habitats, croplands, along roadsides, railroads, and riversides (such as the upper figure, where *A. artemisiifolia* [red circle] is present in a highly invaded community, the positions of other invaders is highlighted).

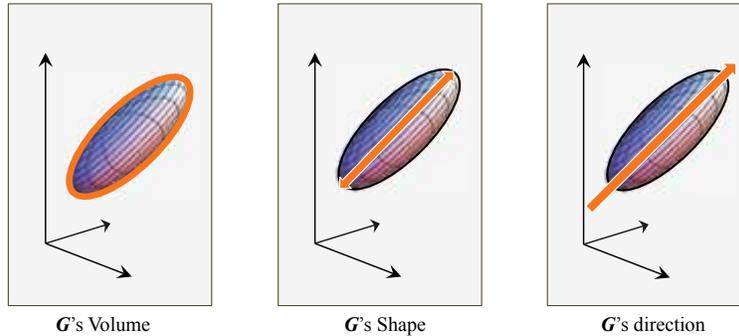
in a species' realized climatic niche, and may not be driven by adaptive changes in their climatic affinities (Gallien *et al.* 2012). At best, observed shifts between native and invasive climate niches may only reflect sorting of populations or lineages that are pre-adapted to particular climatic conditions (Treier *et al.* 2009, Lachmuth *et al.* 2010), without the necessary shift in the genetically based climatic tolerance of introduced populations. An important step towards forecasting the future spread of invasive species is thus to predict the fundamental limits of their climatic ranges. It is, however, not clear whether climatic niches of invasive species can be considered as fixed species features or whether they can experience rapid evolution, as shown for other species characteristics (e.g. Maron *et al.* 2004, Duglosch & Parker 2008, Urban *et al.* 2008). If niche evolution appears to happen frequently during invasions, it will have a large impact on invaders' potential distribution and thus probably increase their impacts on the resident native species. This is particularly alarming given that most invaders come from horticultural and ornamental production, exactly selected for their hardiness and rapid early growth in a large range of environmental conditions (Duglosch & Parker 2008, Richardson & Pysek 2012). Therefore the primary questions that need to be answered are: Is there evidence for a genetic based niche evolution of the climatic niche of invaders? Will it continue in the future?

## **5.2. WHAT DRIVES ADAPTIVE EVOLUTION DURING INVASION?**

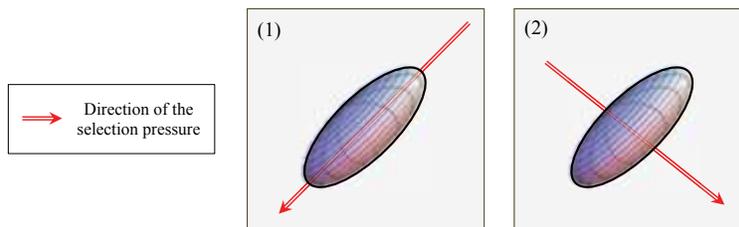
There are several ways for the proper study of evolution in invasion ecology that can be carried out: (1) comparing native vs. adventive populations' genetically-based niche differences (e.g. Kolbe *et al.* 2004, Lavergne & Molovsky 2007, Rey *et al.* 2012), (2) reciprocal-transplanting of individuals in different conditions of the niche (e.g. Sexton *et al.* 2011), and (3) studying the genetic patterns of invasive populations over the spatial and environmental space of the adventive region. Though the three approaches should ultimately be applied concurrently to gain a deeper understanding of the adaptive evolution during the colonisation process, this is often not feasible for economical or

## BOX 8. $\mathbf{G}$ MATRICES & SELECTION SKEWERS METHOD

The  $\mathbf{G}$  matrices: matrices of additive genetic variance-covariance



The selection skewer method



To estimate whether genetic correlations limit the ability of a population to respond to selection, one can study the “lines of least resistance” in the population’s additive genetic variance covariance matrix  $\mathbf{G}$  (Schluter 1996).  $\mathbf{G}$  describes the level of genetic variation for a set of selected traits (e.g. traits measured in a common garden), and the extent to which these traits are genetically correlated to one another.  $\mathbf{G}$  matrices are built at a population level, and enable the comparison of genetic trait correlations and trade-offs across populations.

$\mathbf{G}$  matrices can be represented as ellipses (in an orthogonal trait space), and three measures are usually used to characterise them.  $\mathbf{G}$ 's volume, the total genetic variance, is measured as the sum of all individual trait variances (i.e. the trace of  $\mathbf{G}$ ).  $\mathbf{G}$ 's shape, the strength of correlations, is taken as the proportion of the total genetic variance due to  $P_{max}$ , the first eigenvector. Finally,  $\mathbf{G}$ 's direction, the main direction of the genetic variation (line of least resistance, Cheverud 1996, Kirkpatrick 2009), is given by the direction of  $P_{max}$ .

The potential of trait evolution for several populations can then be estimated using the populations'  $\mathbf{G}$  and the selection skewers method (Calsbeek & Goodnight 2009). This method compares the response of the different  $\mathbf{G}$  matrices to a selection vector (with a given type and intensity of selection) to test whether they will result in alternative evolutionary responses. In the above figure, in case (1) the evolutionary potential will be greater than in case (2), as the direction of selection is oriented on the line of least resistance of  $\mathbf{G}$  in case (1), but not in case (2).

organisational reasons. In this thesis, I used the third approach to understand, in detail, what happened during the colonisation of the French Alps by an alien plant (the common ragweed, *Ambrosia artemisiifolia* L., see Box 7) across a wide range of niche conditions.

Niche evolution in an invasive species can only happen when there is *genetic innovation*, that is, the creation of new genotypes, and *functional evolvability*, which is the genetically based variation in phenotypic traits with adaptive significance. This type of critical process, which enables niche and range evolution to occur, often take place at niche margins, i.e. in marginal ecological conditions (Kawecki *et al.* 2008, Sexton *et al.* 2009). In the case of invasive species, repeated introductions and significant propagule pressure, as well as the ability to overcome population bottlenecks, have been shown to enhance rapid adaptive evolution (Kawecki *et al.* 2008). Post-introduction admixture actually has the potential for increasing genetic variation in selective traits and/or increasing individual fitness through hybrid advantage (Facon *et al.* 2006). However, it is yet to be established whether this occurs at the forefront of niche limits because most studies have used latitudinal or elevation gradients as surrogates of climatic gradients without *a priori* testing. Increased propagule pressure and gene flow may, nonetheless, also provoke gene swamping in climatically marginal populations and prevent them from developing local adaptations (Bridle & Vines 2007, Kawecki 2008). High levels of gene flow may also increase the probability of allele surfing (Lachmucht *et al.* 2010, Excoffier *et al.* 2009) and create the possibility of spurious allele-environment relationships occurring, which would be wrongly attributed to local adaptations. Therefore, it is necessary to contrast adaptive and non-adaptive genetic variation to tease apart the effects of adaptive processes from random changes due to the colonisation process. Finally, the selection leading to adaptation can only have an effect if there is both enough genetic variation for the traits involved in this adaptation to occur, and if their genetic correlations make it possible depending on the direction of the selection pressure (Cheverud 1996, Martin *et al.* 2008, Kirkpatrick 2009, Box 8).

---

**Genetic admixture:**  
Breeding of individuals from two previously separated populations

---

---

**Gene swamping:**  
Loss of genetic variance at a locus under selection because gene flow is too high

---

## BOX 9. THE FRENCH ALPS



Presentation of the studied region: French Alps (here circled in red). This region encompasses a wide diversity of climates, ranging from the Mount Blanc Massif, where precipitation is high, to the Mediterranean Alps that are much drier and warmer.

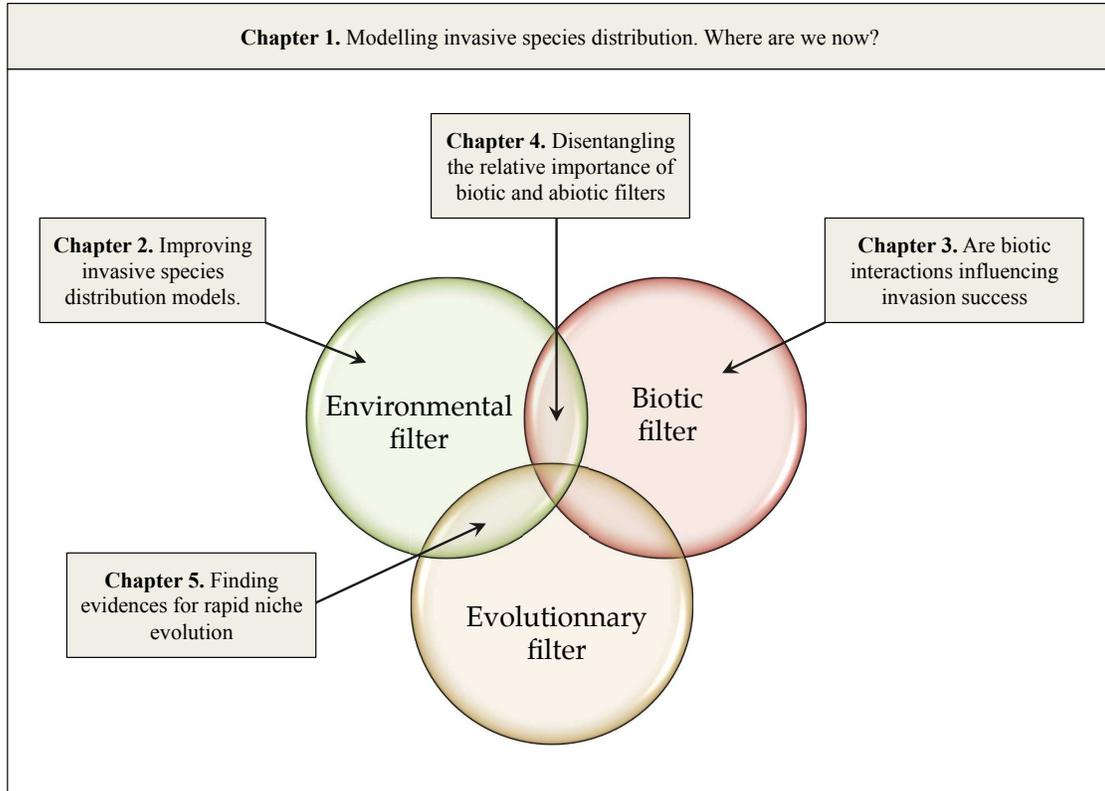
## **6. ORGANISATION OF MY WORK AND THE STUDY SYSTEM**

### **6.1. STUDY SYSTEM: THE FRENCH ALPS**

I chose the French Alps as a study system for some of my work because this region encloses a large number of interesting characteristics. First, the French Alps encompasses wide bioclimatic gradients and is at the ecotone between several biogeographic regions (Mediterranean, continental and Alpine), and harbours elevation range from 0 to 4800 m above sea level (Box 9). The recent alpine orogeny (late Mesozoic) also provides very different rock types. This high diversity of climatic conditions enables an extremely high diversity of species to persist, comprising more than half of French flora (Boulangeat *et al.* 2012b). Second, when the aforementioned climatic heterogeneity is combined with the high level of landscape fragmentation in the region (due to anthropogenic disturbances, such as deforestation) it makes it possible for a newly introduced species to reach very different types of communities within very short distances.

Moreover, since the dispersal of potential invaders is strongly linked to the level of anthropogenic disturbances, which occur more frequently in valleys than at mountaintops, dispersal roads are highly concentrated in the valleys, potentially carrying seeds over very long distances. Finally, the alpine region, as an insular system, harbours an elevated number of rare and endemic species, which are thought to be heavily threatened by climate change, and particularly climatic warming in the region (Thuiller *et al.* 2005, Engler *et al.* 2011, Dullinger *et al.* 2012). As a consequence, the recent increase of invasive species populations in the Alpine region (Walther *et al.* 2009) may constitute an additional threat for these species, particularly if they show rapid adaptation to the alpine climate.

## BOX 10. GENERAL ORGANISATION OF MY PHD

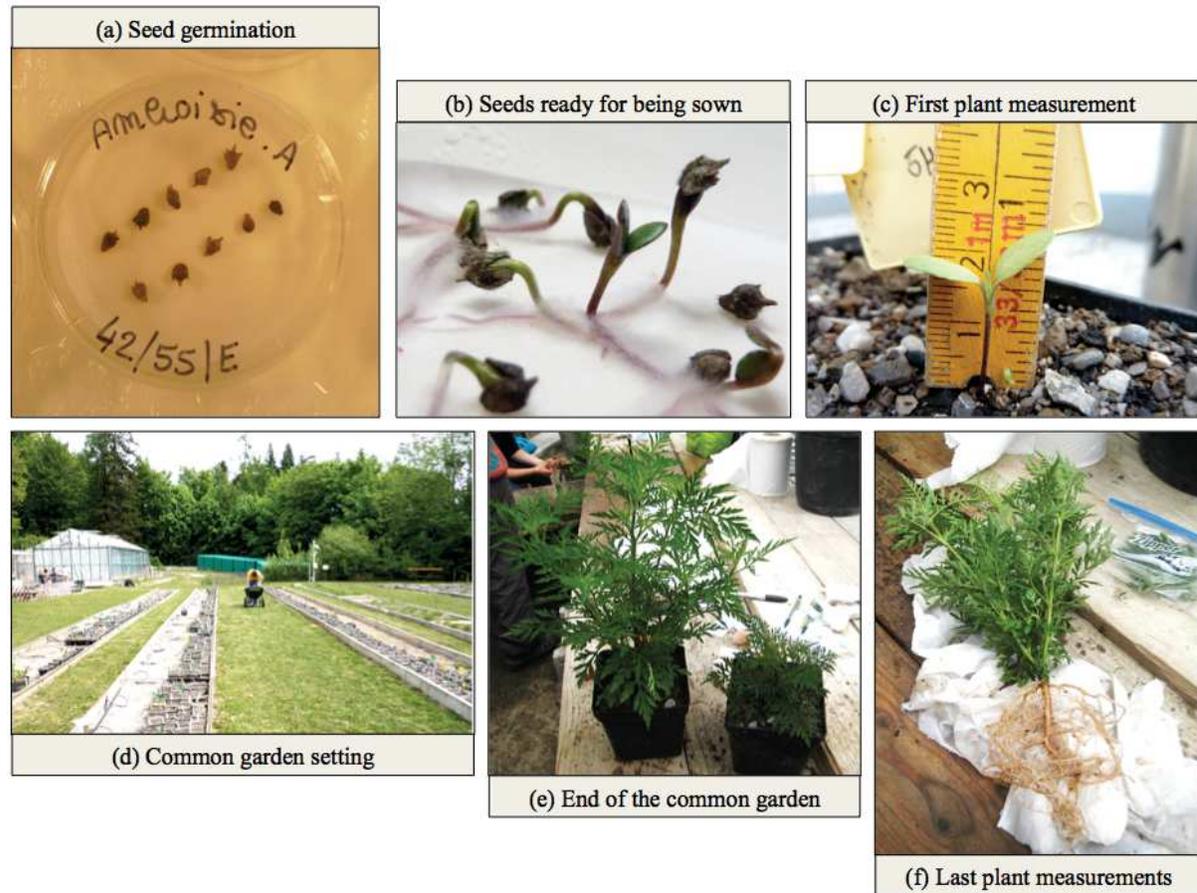


Schematic representation of the PhD chapters according to the three filters of invasion studied.

## 6.2. ORGANISATION OF THE WORK

The aim of my PhD was to reach a better understanding of three filters limiting biological invasions once introduced: the environmental filter, the biotic interactions, and the rapid adaptive evolution filter, as well as their interactions at regional scales (see Box 10). Firstly, I reviewed the current methodological and conceptual limitations of our capacity to predict invasion spread in a given region (Chapter 1, Gallien *et al.* 2010). These limitations comprise the usual trade-off between knowledge requirement (to set the model parameters) and data requirement (to calibrate phenomenological models). Based on this initial study, I developed a modelling framework for improving environmental niche estimations at regional spatial scale, by combining global and regional information, and applied it to the case study of colonisation of the French Alps by exogenous plants (Chapter 2, Gallien *et al.* 2012). The next step was to focus on different ways of integrating biotic interactions to the biogeographic study of invasion. I thus first focused on understanding the methods commonly used to identify processes of competition in ecological communities, as well as their limitations. To do so, I worked in collaboration with others to review the methodological limitation of particular approaches (Chapter 3.1, Thuiller *et al.* 2010), and to test our expectations of spatial scale limitations on invaded plant communities in Italian coastal dunes (Chapter 3.2 Carboni *et al.* 2012). To further test under which conditions the different metrics that are usually applied can truly reveal the processes driving competitive interactions in the invasion context, I implemented a simulation model of community assembly (Chapter 3.3 Gallien *et al.* in prep, this model had already been applied in the native assembly context, Appendix 1 Chalmandrier *et al.* 2013). As a natural follow up, I incorporated the indices previously used to estimate competitive interactions into environmental niche models in order to (i) identify which type of biotic interactions were the strongest drivers of plant invasions in alpine plant communities, (ii) whether their importance varied among invaders and (iii) across the environmental niche of these invaders (Chapter 4, Gallien *et al.* in review a). As a last step, I investigated the potential rapid adaptation of the

## BOX 11. COMMON GARDEN



This figure shows some steps of the common garden experiment on *A. artemisiifolia* at the National Botanical Conservatory of the Alps (Conservatoire Botanique National Alpin, CBNA) in Gap (French Alps). For this experiment about 3800 seeds from 27 populations were grown under the same conditions (Picture (a)). Once the seeds germinated (Picture (b)) they were placed into pots (Picture (c)) and settled outside according to a predefined random design (Picture (d)). Two months and half later (i.e. beginning of the flowering time), plants were removed from the garden (Picture (e)) and four traits were measured (Picture (f)): height, above- and below-ground biomass, and leaf dry matter content.

climatic niche of an invasive weed *Ambrosia artemisiifolia* L. suspected to rapidly adapt to the cold conditions of the French Alps. I combined information on its environmental niche, genetic structure and measured functional traits in common garden experiments (Box 11) to identify whether this species really would show rapid climatic niche evolution and investigate what would be the consequences of such evolution for further spread of this species (Chapter 5, Gallien *et al.* in review *b*).



# *Chapter 1*



# MODELLING INVASIVE SPECIES DISTRIBUTION

-

WHERE ARE WE NOW?

*MODELLISATION DES ESPECES INVASIVES*

-

*OÙ EN SOMMES NOUS AUJOURD'HUI?*



## PRESENTATION DU CHAPITRE

*Ces dernières années, la modélisation des espèces invasives et de leur distribution potentielle a soulevé un formidable intérêt. Cependant, la recherche dans ce domaine s'est développée dans deux directions opposées: la première, vers la mise au point de méthodes de criblage utilisant des modèles phénoménologiques, la deuxième, portant sur la prédiction de la dynamique des espèces invasives utilisant des modèles mécanistes. Dans ce chapitre, nous présentons le développement de modèles hybrides, comme une approche permettant de créer une passerelle et d'intégrer les avantages fournis par les deux domaines de recherche.*

*Dans un premier temps, nous rappelons brièvement les caractéristiques et les limitations des deux approches initiales. Ensuite, nous proposons une revue de la littérature sur les récents développements des modèles hybrides, discutons leurs problèmes actuels et proposons des pistes d'amélioration future.*

*Généralement, les modèles hybrides sont capables de combiner les avantages des approches phénoménologiques et mécanistes. Les défis majeurs lors de leurs développements restent les choix liés au niveau de détail approprié ainsi qu'à l'efficacité et au lien utilisé pour connecter leurs différents sous-modèles. Etant donné ces défis, nous discutons des liens entre les paramètres des modèles phénoménologiques et ceux des modèles mécanistes, des concepts sous-jacents de niche fondamentale et de niche réalisée, ainsi que des problèmes liés aux boucles de rétroaction entre les dynamiques de population et les facteurs environnementaux.*

*A partir du moment où ces défis seront résolus les modèles hybrides deviendront des outils de premier choix pour apporter des prédictions fiables de la distribution potentielle des espèces invasives, de leur dynamique et de leurs conséquences.*





## Predicting potential distributions of invasive species: where to go from here?

Laure Gallien\*, Tamara Münkemüller, Cécile H. Albert, Isabelle Boulangeat and Wilfried Thuiller

Laboratoire d'Ecologie Alpine, CNRS UMR  
5553, Université Joseph Fourier, BP 53, 38041  
Grenoble Cedex 9, France

### ABSTRACT

**Aim** There has been considerable recent interest in modelling the potential distributions of invasive species. However, research has developed in two opposite directions: the first, focusing on screening, utilizes phenomenological models; the second, focusing on predictions of invasion dynamics, utilizes mechanistic models. Here, we present hybrid modelling as an approach to bridge the gap and to integrate the advantages of both research directions.

**Location** Global.

**Methods** First, we briefly summarize the characteristics and limitations of both approaches (screening vs. understanding). Then, we review the recent developments of hybrid models, discuss their current problems and offer suggestions to improve them.

**Results** Generally, hybrid models are able to combine the advantages of currently used phenomenological and mechanistic approaches. Main challenges in building hybrid models are the choices of the appropriate degree of detail and efficiency and the decision on how to connect the different sub-models. Given these challenges, we discuss the links between the phenomenological and the mechanistic model parameters, the underlying concepts of fundamental and realized niches and the problem of feedback loops between population dynamics and environmental factors.

**Main conclusions** Once the above challenges have been addressed and the necessary framework has been developed, hybrid models will provide outstanding tools for overcoming past limitations and will provide the means to make reliable and robust predictions of the potential distribution of invasive species, their population dynamics and the potential outcomes of the overall invasion process.

### Keywords

Biological invasions, habitat suitability model, hybrid model, invasion dynamics, mechanistic model, species distribution model.

\*Correspondence: Laure Gallien, Laboratoire  
d'Ecologie Alpine, CNRS UMR 5553,  
Université Joseph Fourier, BP 53, 38041  
Grenoble Cedex 9, France.  
E-mail: laure.gallien@gmail.com  
All authors contributed equally to the paper.

### INTRODUCTION

Biological invasions, resulting in biotic exchange and subsequent homogenization, are a major component of global change (Vitousek *et al.*, 1997). The anthropogenic displacement of species when followed by permanent establishment, rapid colonization and uncontrolled spread, i.e. biological invasion (Pyšek *et al.*, 2004), modifies native diversity, ecosystem functioning and associated goods and services (Vitousek *et al.*, 1997). Predicting and understanding invasion processes is therefore essential for management actions and policies. The search for common patterns among different invasion events

has produced a large body of literature focussing on the intrinsic properties of invaders (reviewed in Rejmánek *et al.*, 2005; Pyšek & Richardson, 2007), the propensity of natural communities to be invaded (Rejmánek *et al.*, 2005) and the relationship between invaders' distributions and environmental factors (Thuiller *et al.*, 2006; Wilson *et al.*, 2007). Although insights into this work has improved our understanding of invasions and has fostered the development of improved approaches for screening, our ability to reliably predict invasion processes is still very limited. A number of limitations result from the fact that studies traditionally either focused on 'brute-force' broad-scale screening and multi-species

predictions (Peterson *et al.*, 2008) or on incorporating local-scale processes to analyse species-specific dynamic outcomes (Higgins *et al.*, 1996). However, increased computer power now allows combining the advantages of these two approaches, offering a promising avenue towards better models for predicting which species could invade and what could be the course and outcome of invasions.

Broad-scale screening approaches aim to predict which species have the ecological niche to potentially maintain viable populations in a given area (Peterson & Vieglais, 2001). They rely on phenomenological habitat suitability models (HSMs) that describe and extrapolate patterns and relationships (Daehler *et al.*, 2004; Kolar, 2004). HSMs are based on the ecological characteristics of known occurrences in the native distribution of a species and aim to identify the suitable local areas in a potentially available new range (Peterson & Vieglais, 2001). Screening approaches do not directly account for underlying processes but assume that the influence of local processes can be captured indirectly by analysing patterns at larger spatial scales (Ficetola *et al.*, 2007; Beaumont *et al.*, 2009a; Roura-Pascual *et al.*, 2009b). However, this underlying assumption might be violated when extrapolating into new regions and under global change (e.g. climate or land use changes), resulting in potentially erroneous predictions (Davis *et al.*, 1998; Dormann, 2007). Nonetheless, screening approaches, promoted by the increasing availability of environmental and distributional data, have been successfully applied to describe and extrapolate presence/absence patterns for large numbers of potentially invading species and over large areas (Daehler & Carino, 2000; Roura-Pascual *et al.*, 2004; Thuiller *et al.*, 2005; Ficetola *et al.*, 2007). Although HSM have been criticized (see next part, Mack, 1996; Hulme, 2003), their efficiency in predicting invasions is of primary importance for preventive invasion management. The reason is that the attempt to eradicate invasive species after their establishment causes colossal costs and is often unsuccessful (Perrings *et al.*, 2005; Pimentel *et al.*, 2005).

Alternatively, approaches that aim to predict the spread and dynamic outcomes of invasions usually incorporate demographic processes and/or landscape structure. They are mostly applied to address questions focusing on demographic dynamics of invasive species after their establishment: How is the species likely to spread (Higgins *et al.*, 1996)? How is the species going to influence the native community? Mechanistic simulation models are the tools of choice for such purposes as they are able to explicitly incorporate local-scale processes and dynamics (Table 1). As these models directly simulate the mechanistic link between the environment, biotic interactions and the invaders' demographic responses, they are supposed to be less prone to produce erroneous predictions for new regions and under global change (Morin & Lechowicz, 2008). However, building mechanistic models is highly data demanding and involves more complex model structures for which better expert knowledge and process-based understanding is required.

Today, there is a growing awareness that the advantages of both phenomenological models (most notably their efficiency

on broad spatial scales and for many species) and mechanistic models (most notably their ability to model new situations) are necessary to improve our ability to predict accurately the outcome of invasion (Morin & Lechowicz, 2008; Thuiller *et al.*, 2008; Brook *et al.*, 2009; Franklin, 2010). This is because of the fact that invaders often encounter completely new settings in the adventive range. These new settings are not captured by the broad-scale relationship perceived in the native range. For example, invaders may encounter new types of landscapes, new barriers, new competitors or enemies. To account for such differences between the native and the invaded range, we need to model the processes that are sensible to these differences. The idea of modelling species distributions on the basis of large-scale relationships while at the same time considering the most important processes has recently led to the development of so-called hybrid models (Morin & Lechowicz, 2008; Thuiller *et al.*, 2008; Brook *et al.*, 2009). We believe that these hybrid models can solve some of the most important problems occurring when projecting species distributions in space and time and aim at advancing their use in invasion ecology.

In the following, we introduce the different models used either for *screening and broad-scale predictions* or *predictions of dynamics outcomes* and discuss their respective purposes and limitations. We then briefly review how these contrasting approaches have been combined to hybrid models to overcome the conceptual and statistical shortcomings underlying the single approaches. However, hybrid models have only recently been developed and can be improved in several ways. Therefore, we finally develop a set of rules of thumb to facilitate and improve the use of hybrid models for predicting invasion events and suggest solutions to overcome some of their current limitations.

## APPROACHES TO PREDICT INVASIONS

### Screening and broad-scale predictions

Screening studies are based on phenomenological habitat suitability models (HSM), which statistically relate species occurrences to environmental variables (Guisan & Thuiller, 2005; Franklin, 2010). Although species distributions are co-determined by various physical factors (e.g. temperature or soil pH), biotic interactions (e.g. predation or pollination) and disturbances, climate is often seen as the main driver at large spatial scales (Woodward & Williams, 1987; Willis & Whittaker, 2002). Thus, at first, HSMs were often solely based on climatic data (Franklin, 1995; Guisan & Zimmermann, 2000; Heikkinen *et al.*, 2006) but they have been refined afterwards with data representing other aspects of the environment, such as land use, soil or productivity (Pearson *et al.*, 2004; Bradley & Mustard, 2006; Ficetola *et al.*, 2007). Generally, researchers have either calibrated models using the species' native range to extrapolate the found patterns into the adventive range (e.g. Beerling *et al.*, 1995; Welk *et al.*, 2002; Peterson *et al.*, 2003; Richardson & Thuiller, 2007; Ibanez *et al.*, 2009) or simply calibrated the model in the adventive range to predict the

**Table 1** Broad classification of different modelling techniques mentioned in the article and their associated key references. This table is a toolbox for hybrid model builders

Type of model	Description	Key reference	Example of use in invasion ecology
Curve fitting model (CFM)	A CFM is a formula-based description of a process or a pattern, typically analytically solvable. It is often used as a sub-model in a more complex model. Examples are CFMs describing dispersal kernels (e.g. fat-tailed negative exponential models) or population dynamics (e.g. logistic model)	(May, 1976)	Parameterization of dispersal ability of an invasive species (Skarpaas & Shea, 2007)
Matrix population model (MPM)	A MPM describes the growth process of individuals or cohorts via life-stages and transition probabilities (e.g. using Leslie matrices) and is analytically solvable. Examples of applications are population viability analyses. There is no information on space	(Caswell, 2001)	Evaluation of the local dynamic of an invasive species (Sebert-Cuvillier <i>et al.</i> , 2007)
Metapopulation model (MM)	A MM describes the demographic dynamics of a population living on suitable habitat patches within a hostile matrix of unsuitable habitat. The main focus is on extinction and colonization of local populations. The simpler metapopulations MMs are analytically solvable (e.g. incident models). More complex MMs can be spatially explicit and can describe dispersal, reproduction and competition explicitly	(Hanski & Gaggiotti, 2004)	Evaluation of the risk of introduction of a non-native species (Deines <i>et al.</i> , 2005)
Cellular automaton (CA)	CAs are stochastic spatially explicit models that may be used to describe spread and spatial interactions. Each cell on a grid evolves through discrete time steps according to a set of rules based on the states of neighbouring cells. It is typically used to explore colonization processes and patterns	(Bolliger <i>et al.</i> , 2003)	It could be used to evaluate the influence of initial spatial structure in the spread of an invasive species (Ferrari & Lookingbill, 2009)
Landscape model (LM)	LMs are spatially explicit models aiming at projecting a landscape (structure, function, composition) over time. They can include spatial interactions, community dynamics or/and ecosystem processes. LMs are typically used to simulate different management or global change scenarios. Two broad classes of examples are gap/landscape models (e.g. LANDIS, ForCLIM) and dynamic vegetation models (e.g. IBIS, LPJ)	(Scheller & Mladenoff, 2007)	There are few examples of landscape models in invasion ecology. It could be used to evaluate the colonization dynamic of a species (Albert <i>et al.</i> , 2008)
Individual-based model (IBM)	IBMs are models that focus on units (e.g. individuals, populations...) and their interactions. It describes processes at small scales that directly influence the units. IBMs are typically used to investigate patterns emerging at larger scales and to make predictions	(Grimm & Railsback, 2005)	It could either describe qualitatively the invasion process (Travis <i>et al.</i> , 2007) or quantify results of invasion process (Nehrbass & Winkler, 2007)
Mechanistic niche model (MNM)	MNMs are based on niche theory and describe the link between a species and its environment from the relationship between species' characteristics (behaviour, morphology, physiology...) and environmental factors. They are mainly used to predict patterns of species distribution over space and/or time	(Kearney & Porter, 2009)	Predictions of the cane toad's distribution under future climatic scenario (Kearney <i>et al.</i> , 2008)
Habitat suitability model (HSM)	HSMs are statistical models that are based on niche theory and fit the link between a species and its environment from occurrence or abundance data and environmental data. They are mainly used to predict patterns of species distribution over space and/or time	(Guisan & Thuiller, 2005)	Large-scale predictions of the risk of invasion by an alien species (Thuiller <i>et al.</i> , 2005)

potential extent of species' distribution (Zalba *et al.*, 2000; Roura-Pascual *et al.*, 2004; Parker-Allie *et al.*, 2009). By using such an environmental-based approach, scientific efforts have

focused on defining potential invasive species through environmental matching (Peterson *et al.*, 2008). The use of this approach is related to one of the main hypothesis in invasion

ecology stating that the environment of native vs. adventives' ranges has to be similar to allow for a successful invasion (Panetta & Mitchell, 1991; Scott & Panetta, 1993).

Habitat suitability models have a limited accuracy in providing predictions of future invasions as they do not explicitly incorporate demographic processes driving species distribution and invasion rates (e.g. fecundity and dispersal ability). However, they are particularly efficient to assess the invasive potential of large numbers of species before their introduction (Peterson & Vieglais, 2001) and are often reasonable alternatives when other modelling tools are missing or are excessively time or money consuming.

Besides specific limitations for the application to invasions, HSMs also have some further well known and described limitations that we will not detail here (Guisan & Thuiller, 2005; Bahn & McGill, 2007; Dormann, 2007). In the context of biological invasions, HSMs are prone to predict substantial false presences and false absences because of the non-equilibrium nature of the invader's distribution. False presences can be predicted when environmental variables non-introduced in the models (such as soil type, disturbance regime or interspecific interactions) are limiting the naturalization of a species in the invaded range. False absences occur if a species' potential distribution has not been realized in its native range because of non-equilibrium dynamics, e.g. because of historical constraints attributable to human influences or because of physical barriers that prevent full range occupancy (Curnutt, 2000).

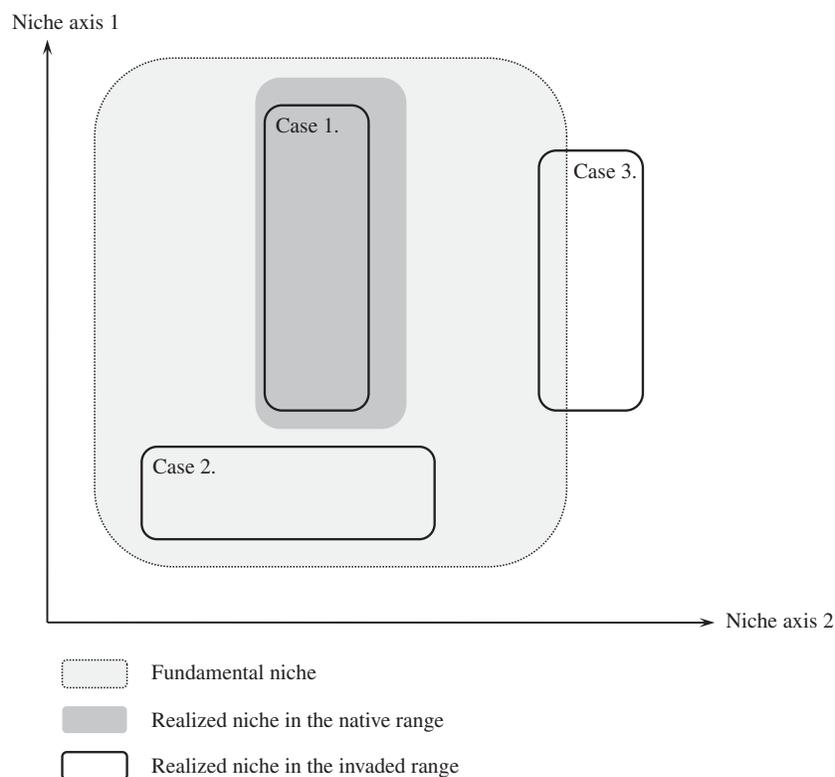
In the native range, a given species occurs at the intersection of suitable (climate, resource), available (biotic interactions, habitat disturbance) and reachable (dispersal) habitats (Sobron, 2007). In the absence of source-sink dynamics, this intersection, commonly called the realized niche of the species (Hutchinson, 1957), is theoretically smaller than the species' fundamental niche (Pulliam, 2000). Comparing the realized niche within the native vs. the invaded ranges can lead to three non-exclusive theoretical cases (Fig. 1). First, in the invaded range, the species could use a similar or smaller realized niche than in the native range. This case is expected when the environment and the outcomes of biotic interactions in an adventive area are comparable to the native area (case 1 in Fig. 1). Only in this case, the assumptions of HSMs are fully met and we can expect reliable predictions from a HSM exclusively calibrated with data from species' native ranges (Thuiller *et al.*, 2005). Second, the introduced species may occupy a realized niche very different from the one in the native area, for instance because of new predator community, multiple sites of introduction, niche differentiation (e.g. in ploidy level, Treier *et al.*, 2009) or different environmental conditions (case 2 in Fig. 1). In this case, a model exclusively calibrated with data from species' native ranges will fail by predicting erroneous potential ranges (Broennimann *et al.*, 2007; Fitzpatrick *et al.*, 2007). This problem can be partly addressed for world-wide invaders by using all known occurrences (both from the native and invasive ranges) to calibrate the model (e.g. Kearney *et al.*, 2008; Beaumont *et al.*, 2009b). Third, the species could undergo rapid genetic

adaptation. Genetic adaptation violates the underpinning assumption of slow niche evolution when predicting species distribution with HSMs (Holt, 1992) and is probably most difficult to account for (case 3 in Fig. 1). In the last years, several studies challenged the assumption of slow niche evolution by demonstrating that some invasive species have rapidly evolved during the course of invasion because of genome size reduction (Lavergne & Molofsky, 2007), genetic bottleneck, converging selection, mutations (Phillips *et al.*, 2008b) or hybridization (Hall *et al.*, 2006). In this case, the realized niche may extend outside of the species initial fundamental niche. The only way to address this issue in a HSM framework is to calibrate the model with all known occurrences and make sure that the model also includes occurrences from the particular invaded range where rapid adaptation is ongoing. Calibrating habitat suitability models on all known occurrence could also lead to some particular problems which depend on the overall goal of the analysis. Indeed, models calibrated on all known occurrences are likely to over-predict the distribution in the invaded range of a species currently invading in a particular area. The researcher will have to decide whether this is a problem or not. In the goal of predicting the potential distribution of the species for prevention, it is clearly welcome to know where the species could further invade and one would be more tolerant with regard to false presence predictions. In the case of understanding and possibly eradicating the species, a model producing a better match with the current distribution in the invaded range is probably more acceptable.

### Processes and predictions of dynamics outcomes

Although phenomenological HSMs have been the tool of choice for screening purposes, they have not improved our understanding of the dynamics underlying invasions and their outcomes. A better understanding of invasions requires a better understanding of the demographic processes that drive invasion spatio-temporal dynamics and of the characteristics of the invaders, the recipient communities and the environmental variables that influence these processes. Two key demographic processes: dispersal (Hastings *et al.*, 2005) and growth (Jongejans *et al.*, 2008), are known to influence invasions differently at different invasion stages (Dawson *et al.*, 2009). More specifically, Allee-effects (Taylor & Hastings, 2005), interspecific interactions (Davis *et al.*, 1998; Richardson *et al.*, 2000), phenotypic plasticity (Wilson *et al.*, 2009), genetic adaptations (e.g. hybridization, Hall *et al.*, 2006) or increasing dispersal abilities (Travis *et al.*, 2007; Phillips *et al.*, 2008a) and disturbances (Edward *et al.*, 2009) are examples for processes that have been identified as being important during certain invasions. Improving our understanding of the causal role of demographic processes in invasions can be achieved either on the basis of experiments, for example field, greenhouse (e.g. Leishman & Thomson, 2005) and microcosm experiments (e.g. Davis *et al.*, 1998), or through mechanistic models (With, 2002; Nehrbaas *et al.*, 2006).

**Figure 1** The realized niche dilemma in predicting invasion risk based on habitat suitability models. Several possibilities for the realized niche in the invaded range compared to the realized niche in the native range. *Case 1*: The realized niche in the invaded range is similar to the one in the native range. It can occur if the outcome of biotic interactions is similar in both ranges. *Case 2*: The realized niche in the invaded range may be very different from the one in the native range. It can occur because of different biotic interaction like enemy release, different access to sites because of introduction, different environmental conditions, or different niches because of populations having different ploidy levels. *Case 3*: The realized niche in the invaded range may be partially outside the fundamental niche because of rapid genetic adaptation.



Mechanistic models of invasions simplify the natural system and reduce it to its basic processes to improve the understanding of the underlying invasion mechanisms (Wissel, 1989). Deciding on how much reality should be simplified (i.e. choosing the best level of ecological details) is one of the hardest questions. The answer to this depends on the research question and may offer an array of different solutions ranging from theoretical to applied models (Bolker, 2008). On the one hand, mechanistic models can be *theoretical models* developed to explore a concept without reference to a particular species or place. The results of such theoretical mechanistic models show qualitative hints and trends and can be generalized within the framework of 'a priori' assumptions (Bolker, 2008). Theoretical models also contributed to growing consensus, such as the importance of long-distance dispersal events for range expansions, though rare and difficult to predict (Hastings *et al.*, 2005). Moreover, dispersal kernels (i.e. the probability function of dispersal distances) might not remain static during invasion events. The process of invasion itself may induce strong selection pressure on species' dispersal abilities, resulting in increased dispersal at the expanding front (either through mutations or because of higher fitness and resulting agglomeration of strong dispersers, Travis & Dytham, 2002; Phillips *et al.*, 2008a). On the other hand, mechanistic models can be *applied models* developed with the aim of providing quantitative and detailed predictions on specific cases and then striving to incorporate more ecological details. For example, Nehrbass *et al.* (2006) parameterized and compared a deterministic matrix model and an individual-based model to analyse why a harmful invasive species, the Giant Hogweed

*Heracleum mantegazzianum*, has shown long-term range expansion but short-term population decline. They identified temporal variability in demographic factors as the main driver of such dynamic and concluded from their model comparison that taking into account invader's demography can lead to strong practical implications for control measures (Nehrbass & Winkler, 2007). The development of applied mechanistic models is constrained by available expert knowledge used to formulate model rules and functions and by the data needed to parameterize the model. Spatially detailed information on key environmental factors such as pH or soil water humidity is often lacking and can obviously preclude model building.

Both theoretical and applied mechanistic models utilize a broad range of different modelling techniques. Among them are mechanistic niche models (MNM), matrix population models (MPM), metapopulation models (MM), individual-based models (IBM) and landscape models (LM, Table 1). However, a comprehensive description and general classification of different modelling techniques goes beyond the scope of this article and has been presented elsewhere (e.g. Grimm & Railsback, 2005; Jorgensen & McLachlan, 2008; Kearney & Porter, 2009).

## HYBRID MODELS – THEIR PRESENT AND THEIR FUTURE

### What has been carried out so far?

Recent years have seen the emergence of hybrid models (e.g. Morin & Lechowicz, 2008; Thuiller *et al.*, 2008; Franklin, 2010)

that aim to overcome former statistical and conceptual limitations by integrating both (1) the predictive accuracy of phenomenological models at large spatial scales and (2) the ability to capture dynamics of mechanistic models.

A number of studies have successfully hybridized different model types to predict the spread of invasive species or endangered species extinction threats (Jeltsch *et al.*, 2008; Thuiller *et al.*, 2008). The simplest model combination, which is so far also the most commonly used one for predictive biogeography, is the association of a HSM with a spatially explicit applied mechanistic model such as spread, metapopulation or landscape models (Albert *et al.*, 2008; Keith *et al.*, 2008; Anderson *et al.*, 2009; Brook *et al.*, 2009; Dullinger *et al.*, 2009; Jacobs & MacIsaac, 2009; Roura-Pascual *et al.*, 2009a; Smolik *et al.*, 2009; Table 1). Model combinations can also adopt different forms, but in most cases, the spatially explicit model parameters (e.g. mortality/survival, carrying capacity, dispersal rate) are constrained by the outputs of a habitat suitability model (e.g. probability of presence or presence/absence). The biological reasoning is that such constraints on the parameters by the HSM mimic the change of species' characteristics and performances throughout the environment (Thuiller *et al.*, 2010). These types of hybrid models rely on the assumption that large-scale environmental gradients (commonly climate) determine which species could persist in a given environment (i.e. habitat filtering, Diamond, 1975), while population dynamic processes take place at smaller spatial scales (Weither & Keddy, 1995; Lortie *et al.*, 2004).

For example, Roura-Pascual *et al.* (2009a) successfully reconstructed the invasion spread of the Argentina ant in Catalonia by constraining the metapopulation dynamics governing the cell-state transition by a topo-climatic-based habitat suitability (see also, Smolik *et al.*, 2009 on *Ambrosia artemisiifolia*). Only extinction and colonization rates were restricted (i.e. linearly weighted) by habitat suitability.

The form of such hybrid models can still increase in ecological details and therefore complexity. To model the population dynamics of an endangered bird species, Wintle *et al.* (2005) proposed a three-step hybridization where vegetation dynamics were modelled by a spatially explicit landscape model (step 1) (LANDIS, Mladenoff & He, 1999). This landscape model in turn fed the bird habitat suitability model (step 2) which constrained the metapopulation dynamics of the bird (step 3) (RAMAS GIS – Metapop, Akçakaya *et al.*, 2003).

Hybridization of models to predict the spread and dynamic of invasive species is not restricted to habitat suitability and metapopulation models. There are several examples of models developed for a given target species. For instance SPAnDX, a detailed climate-driven process-based population cohort model, combining the approaches of forest growth models and community dynamics models, has been specifically developed to model the population dynamics of *Acacia nitolica* (Kriticos *et al.*, 2003). Such complex models focused on a single species are obviously not easily applicable to many species but they can be highly robust and accurate.

## Rules of thumb for the hybrid-building process

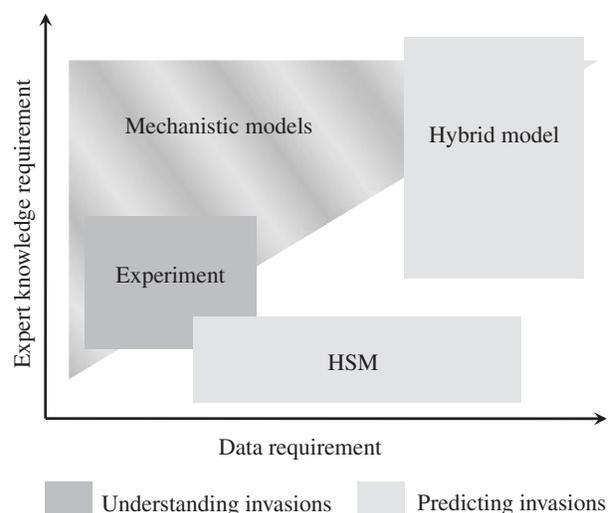
Typically, hybrid models combine phenomenological habitat suitability models (from moderate to high data requirement and low to moderate expert knowledge Fig. 2), with reasonably complex mechanistic models (low data requirement and moderate to strong expert knowledge) and are complex and data-demanding models (shift towards the upper right corner in Fig. 2). Then, one of the major challenges is to select the most appropriate sub-models regarding at the same time: the research question, the required expert knowledge and data availability. But how much complexity is still reasonable? The theoretical answer is clear: the minimum overall error is obtained at moderate levels of complexity. Consequently, increasing complexity does not automatically increase model performance (Wissel, 1989). To help the decision-making about sub-models selection, we propose here a guideline based on four key questions (see example in Box 1).

### BOX 1

The use of a guideline based on the four rules of thumb can be exemplarily shown based on the study of Williams *et al.* (2008). They developed a hybrid model to predict the potential spread of the orange hawkweed (*Hieracium aurantiacum*) from the Bogong High Plains to alpine areas of Australia. The goal of their study was to facilitate early detection of new populations before high abundance threatens native biodiversity, that is to say a moderate need of mechanistic understanding (Question 1). The whole target area was an alpine region, its grain size was 20 × 20 m, and the near future was the temporal scale (Question 2). From this information, one could expect them to combine a HSM with processes such as interspecific interactions, disturbance, demography or evolutionary adaptation.

In fact, with their few data but high expert knowledge and information from the literature, they created a HSM on the conditions of high likelihood of the hawkweed establishment (including native vegetation community type, wetness and disturbance). However, they did not possess enough empirical records to create a mechanistic model, so they decided to model the likelihood of seed dispersal from known populations according to the wind direction, solely based on literature information and expert knowledge. Integrating their knowledge (Question 3) and the data available (Question 4) on the processes involved in the spread of the target species, they could partly account for climate (wetness), indirect species interactions (native vegetation community type), disturbance (e.g. distance to roads) and demography (dispersal). Finally, this modelling strategy has been particularly interesting for the detection of newly established populations via wind dispersion, impossible to reach through the use of simple HSMs.

**Figure 2** Requirements and objectives of different kinds of models. To understand invasions, experiments or conceptual models can be used to simulate virtual worlds based on known processes. To predict invasions without much knowledge, phenomenological models like habitat suitability models (HSMs) are really useful. Mechanistic models may be more accurate in predicting invasions but need lot of knowledge to implement processes. Hybrids models may be a compromise to improve predictions without detailing all processes.



The first two questions are of equal importance: (Question 1) How much understanding (vs. screening) do we need to fulfil the study goal? When the aim of a study is a screening procedure among a large number of invasive species, then only the demographic processes that are essential for all species can be included. For screening, we often may want to include sub-models for dispersal and/or local extinction processes to HSMs. When a study aims at predicting future distributions of a single invasive species, then more detailed expert knowledge about the species' ecology can be used to incorporate a larger number of important processes and sub-models. (Question 2) Which processes of invasion are relevant for the studied system at both spatial and temporal scales? For example, if we would build a hybrid model for a species' increasing dispersal abilities at the leading edge (e.g. Phillips *et al.*, 2008b), we may consider including not only dispersal but also the evolution of the species' dispersal abilities along with the pre-defined habitat suitability. Decisions on the importance of processes can be aided by recalling that usually HSMs implicitly already incorporate all demographic processes of the species. It is only necessary to explicitly include in a hybrid model those processes that are prone to change species' relationships with the habitat and the environment during the invasion process.

These first two questions on the choice of prediction detail level and on the selection of the potential processes of interest allow for the delineation of the 'maximal hybrid model'. This maximal model ideally contains all processes that are important for the study purpose, regardless of the information required and available for the implementation and parameter estimation procedure. Subsequently, the two last questions deal with the feasibility of the hybrid model: (Question 3) For which of the chosen processes do we have sufficient expert knowledge to implement rules and equations? (Question 4) For which of the chosen processes do we have enough available data to parameterize the model? The ultimately selected processes should simultaneously meet the expert knowledge and data requirements. The hybrid model structure chosen through such a hierarchical design contains

the most relevant process combination, avoiding the development of too complex models that could decrease prediction reliability.

### Hybrid model limitations and suggested improvements

Hybrid models do not aim to predict perfectly but to overcome specific limitations of traditional models. As discussed before, most of the existing hybridizations concern HSM and meta-population or landscape models. However, there are different challenges that hamper a more extensive use of hybridization approaches. These challenges concern the form, the strength and the direction of the link between the demographic parameters and the HSM, as well as circularity problem.

#### *Form and strength of the relationship between HSM and demographic parameters*

Two essential questions need to be addressed before hybridization: (1) what parameters of the mechanistic model should be constrained by the habitat suitability measure? (2) What link should be established between these model parameters and habitat suitability? The question (1) is rarely addressed explicitly, and the most important parameters of the mechanistic model are generally constrained based on expert knowledge (e.g. carrying capacity, dispersal rate, growth rate). For the question (2), the link between habitat suitability and for example carrying capacity (Keith *et al.*, 2008; Anderson *et al.*, 2009) or survival/fecundity (Wintle *et al.*, 2005; Albert *et al.*, 2008; Dullinger *et al.*, 2009) is generally assumed to be linear or logistic (but see Kearney *et al.*, 2008). This assumption is not fully supported by experiments or observational analyses. Thuiller *et al.* (2010) showed that the link between habitat suitability and plant performance can be rather idiosyncratic, not always consistent between and within species and not always following the expected direction (e.g. negative relationships instead of positive), corroborating the

few other studies having investigated these relationships (Wright *et al.*, 2006; Elmendorf & Moore, 2008). At the moment, using linear links between selected parameters and a given HSM is the most simple approach and given the limited data the only available alternative. Solving these problems would require reproducing experimental and observational analyses in different environments with different species. Additionally, several articles have shown that HSM outputs can strongly vary depending on the used statistical models (Albert & Thuiller, 2008). This raises the question of selecting one given HSM or a combination of the most reliable ones (e.g. ensemble forecasting, Araújo & New, 2007; Marmion *et al.*, 2009; Roura-Pascual *et al.*, 2009b).

Alternatively, to address the limitations mentioned above, hybrid models could use only presence/absence predictions instead of using habitat suitability with a continuous scale between 0 and 1. This way, the HSM only gives the areas where the species could occur and then the applied mechanistic model simulates the demography based on competition, dispersal, extinction and disturbance (Albert *et al.*, 2008). This would avoid dealing with potentially erroneous assumptions on the type and form of relationships between habitat suitability and model parameters. However, using presence/absence predictions requires the transformation of the continuous habitat suitability information into binary presence/absence using a particular threshold. Selecting for an optimal threshold has been reviewed extensively in the past (e.g. Liu *et al.*, 2005; Hirzel *et al.*, 2006) and need to be carefully thought in the invasion context.

#### *One-way or two-way interactions*

An additional shortcoming of hybrid models is that they are mostly based on a one-way interaction between a model that is supposed to give patch quality or habitat suitability and another model that is supposed to simulate population and community dynamics. However, in the case of invasive species, this one-way interaction could be of limited relevance if the invader is known to modify the environment and the availability of resources. Examples range from nitrogen-fixing plant species that modify ecosystem functioning (Vitousek *et al.*, 1997) and resource use to animal invaders that could influence dispersal dynamics of vegetation. Future developments of hybrid models for modelling invasions should focus on implementing two-way interactions between sub-models to allow for feedbacks.

#### *Circularity*

Using habitat suitability models to constrain the invader's population dynamics raises another problem linked to the circularity of the modelling process. There is an ongoing debate on the exact meaning of the output from an HSMs: Do they represent species habitat vs. species niche (Kearney, 2006), the realized vs. the fundamental niche or the realized vs. potential distribution (Soberon, 2007)? Ideally, the HSM

should predict the fundamental niche of the invader and not the realized niche to be used to influence the population dynamics (e.g. demography) of the target species in mechanistic models. This is not necessary true if the mechanistic model only concerns dispersal for instance. However, HSMs implicitly and indirectly accounts for biotic interactions, disturbance effects, land use legacy and dispersal limitations. This might be problematic because using habitat suitability model outcomes calibrated on observed distributions will lead the hybrid model to account for biotic interactions twice. This is likely to result in under-predictions of the potential distribution of the invader. A way to deal with this problem could be the use of very liberal models (to avoid false absences) that do not overfit (down-weight false presences) to depict only the broad range limits of the invader and let the applied mechanistic model simulate the population dynamics in the potential range.

## CONCLUSIONS

While tremendous progress has been made on many aspects related to the building and evaluation of phenomenological HSMs and theoretical and applied mechanistic models in the context of biological invasions, future efforts should focus on combining the advantages of these various approaches. Phenomenological habitat suitability models have been mostly applied to predict the potential distribution of many species in adventive ranges, ignoring population dynamics and resistance of the native communities, while mechanistic models have been used to understand invasion dynamics once the invader was introduced, mainly ignoring the influence of environmental conditions.

Recent years have seen the emergence of a new generation of models that capitalize on the strength and advantage of both approaches and concepts to make more reliable and useful predictions. These hybrid models typically use phenomenological models to constrain demographic parameters of meta-population or landscape models. Important aspects of hybrid models requiring deeper examination include: (1) the form and strength of the link between habitat suitability and demographic parameters; (2) the potential circularity involved in the use of habitat suitability models – that indirectly already account for biotic interactions and limited dispersal – to constrain demographic parameters (3) the one-way interaction between HSMs and mechanistic sub-models which may not always be robust, especially for invasions which may influence the environment in return. Additionally, we argue that the conception of a hybrid model should not utilize a general a priori design but should follow simple strategic steps based on the following criteria: (1) the ultimate goal is to predict, understand or both; (2) relevant processes for the studied system; (3) selection of processes with enough expert knowledge; (4) selection of processes with enough available data.

Once these challenges are addressed and the framework is rigorously built-up, hybrid models provide outstanding tools

to overcome past limitations and to make reliable and robust predictions of the potential distribution of an invader but also its population dynamics and the outcomes of the overall invasion process.

## ACKNOWLEDGEMENTS

We thank David Richardson for the invitation to write this article and for his insightful comments and editing of the manuscript. We also thank Marcel Rejmánek and four anonymous reviewers who helped to significantly improve the article. This work was funded by ANR SCION (ANR-08-PEXT-03) and DIVERSITALP (ANR-07-BDIV-014) projects. We also received support from European Commission's FP6 ECOCHANGE (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe, No 066866 GOCE) project.

## REFERENCES

- Akçakaya, H.R., Mladenoff, D.L. & He, H.S. (2003) *Ramas landscape: integrating metapopulation viability with LANDIS forest dynamics model. User manual for version 1.0. Applied biomathematics*. Setauket, New York.
- Albert, C. & Thuiller, W. (2008) Favourability functions against probability of presence: advantages and misuses. *Ecography*, **31**, 417–422.
- Albert, C.H., Thuiller, W., Lavorel, S., Davies, I.D. & Garbolino, E. (2008) Land-use change and subalpine tree dynamics: colonization of *Larix decidua* in French subalpine grasslands. *Journal of Applied Ecology*, **45**, 659–669.
- Anderson, B.J., Akçakaya, H.R., Araújo, M.B., Fordham, D.A., Martinez-Meyer, E., Thuiller, W. & Brook, B.W. (2009) Dynamics of range margins for metapopulations under climate change. *Proceedings of the Royal Society B, Biological Sciences*, **276**, 1415–1420.
- Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42–47.
- Bahn, V. & McGill, B.J. (2007) Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography*, **16**, 733–742.
- Beaumont, L.J., Gallagher, R.V., Downey, P.O., Thuiller, W., Leishman, M.R. & Hughes, L. (2009a) Modeling the impact of *Hieracium* spp. on protected areas in Australia under future climates. *Ecography*, **32**, 757–764.
- Beaumont, L.J., Gallagher, R.V., Thuiller, W., Downey, P.O., Leishman, M.R. & Hughes, L. (2009b) Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions*, **15**, 409–420.
- Berling, D.J., Huntley, B. & Bailey, J.P. (1995) Climate and the distribution of *Fallopia japonica*: Use of an introduced species to test the predictive capacity of response surfaces. *Journal of Vegetation Science*, **6**, 269–282.
- Bolker, B.M. (2008) *Ecological models and data in R*. Princeton University Press, Princeton.
- Bolliger, J., Sprott, J.C. & Mladenoff, D.J. (2003) Self-organization and complexity in historical landscape patterns. *Oikos*, **100**, 541–553.
- Bradley, B.A. & Mustard, J.F. (2006) Characterizing the landscape dynamics of an invasive plant and risk of invasion using remote sensing. *Ecological Applications*, **16**, 1132–1147.
- Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701–709.
- Brook, B.W., Akçakaya, H.R., Keith, D.A., Mace, G., Pearson, R.G. & Araújo, M.B. (2009) Integrating bioclimate with population models to improve forecasts of species extinctions under climate change. *Biology Letters*, **5**, 723–725.
- Caswell, H. (2001) *Matrix population models: construction, analysis, and interpretation*. Sinauer, Sunderland, MA, USA.
- Curnutt, J.L. (2000) Host-area specific climatic-matching: similarity breeds exotics. *Biological Conservation*, **94**, 341–351.
- Daehler, C.C. & Carino, D.A. (2000) Predicting invasive plants: prospects for a general screening system based on current regional models. *Biological Invasions*, **2**, 93–102.
- Daehler, C.C., Denslow, J.S., Ansari, S. & Kuo, H.C. (2004) A risk-assessment system for screening out invasive pest plants from Hawaii and other Pacific islands. *Conservation Biology*, **18**, 1–9.
- Davis, A.J., Jenkinson, L.S., Lawton, J.H., Shorrocks, B. & Wood, S. (1998) Making mistakes when predicting shifts in species range in response to global warming. *Nature*, **391**, 783–786.
- Dawson, W., Burslem, D. & Hulme, P.E. (2009) Factors explaining alien plant invasion success in a tropical ecosystem differ at each stage of invasion. *Journal of Ecology*, **97**, 657–665.
- Deines, A.M., Chen, V.C. & Landis, W.G. (2005) Modeling the risks of non-indigenous species introductions using a patch-dynamics approach incorporating contaminant effects as a disturbance. *Risk Analysis*, **25**, 1637–1651.
- Diamond, J.M. (1975) Assembly of species communities. *Ecology and evolution of communities* (ed. by M.L. Cody and J.M. Diamond), pp. 342–444. Harvard University Press, Cambridge.
- Dormann, C.F. (2007) Promising the future? Global change projections of species distributions *Basic and Applied Ecology*, **8**, 387–397.
- Dullinger, S., Kleinbauer, I., Peterseil, J., Smolik, M. & Essl, F. (2009) Niche based distribution modeling of an invasive alien plant: effects of population status, propagule pressure and invasion history. *Biological Invasions*, **11**, 2401–2414.
- Edward, E., Munishi, P.K.T. & Hulme, P.E. (2009) Relative role of disturbances and propagule pressure on the invasion of humid tropical forest by *Cordia alliodora* (Boraginaceae) in Tanzania. *Biotropica*, **41**, 171–178.
- Elmendorf, S.C. & Moore, K.A. (2008) Use of community-composition data to predict the fecundity and abundance of species. *Conservation Biology*, **22**, 1523–1532.

- Ferrari, J.R. & Lookingbill, T.R. (2009) Initial conditions and their effect on invasion velocity across heterogeneous landscapes. *Biological Invasions*, **11**, 1247–1258.
- Ficetola, G.F., Thuiller, W. & Miaud, C. (2007) Prediction and validation of the potential global distribution of a problematic alien invasive species – the American bullfrog. *Diversity and Distributions*, **13**, 476–485.
- Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, **16**, 24–33.
- Franklin, J. (1995) Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, **19**, 474–499.
- Franklin, J. (2010) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, **16**, DOI: 10.1111/j.1472-4642.2010.00641.x.
- Grimm, V. & Railsback, S.F. (2005) *Individual-based modeling and ecology*. Princeton University Press, Princeton.
- Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Hall, R.J., Hastings, A. & Ayres, D.R. (2006) Explaining the explosion: modeling hybrid invasions. *Proceedings of the Royal Society Biological Sciences*, **273**, 1385–1389.
- Hanski, I. & Gaggiotti, O.E. (2004) *Ecology, genetics, and evolution of metapopulations*. Academic Press, San Diego.
- Hastings, A., Cuddington, K., Davies, K.F., Dugaw, C.J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, U., Melbourne, B.A., Moore, K., Taylor, C. & Thomson, D. (2005) The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, **8**, 91–101.
- Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W. & Sykes, M.T. (2006) Methods and uncertainties in bioclimatic envelope modeling under climate change. *Progress in Physical Geography*, **30**, 751–777.
- Higgins, S.I., Richardson, D.M. & Cowling, R.M. (1996) Modeling invasive plant spread: the role of plant-environment interactions and model structure. *Ecology*, **77**, 2043–2054.
- Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, **199**, 142–152.
- Holt, R.D. (1992) Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. *Evolutionary Ecology*, **6**, 433–447.
- Hulme, P.E. (2003) Biological invasions: winning the science battles but losing the conservation war? *Oryx*, **37**, 178–193.
- Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbour Symposia on Quantitative Biology*, **22**, 415–427.
- Ibanez, I., Silander, J.A., Wilson, A.M., Lafleur, N., Tanaka, N. & Tsuyama, I. (2009) Multivariate forecasts of potential distributions of invasive plant species. *Ecological Applications*, **19**, 359–375.
- Jacobs, M.J. & MacIsaac, H.J. (2009) Modeling spread of the invasive macrophyte *Cabomba caroliniana*. *Freshwater Biology*, **54**, 296–305.
- Jeltsch, F., Moloney, K.A., Schurr, F., Köchy, M. & Schwager, M. (2008) The state of plant population modelling in light of environmental change. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 171–189.
- Jongejans, E., Shea, K., Skarpaas, O., Kelly, D., Sheppard, A.W. & Woodburn, T.L. (2008) Dispersal and demography contributions to population spread of *Carduus nutans* in its native and invaded ranges. *Journal of Ecology*, **96**, 687–697.
- Jorgensen, M.A. & McLachlan, G.J. (2008) Wallace's approach to unsupervised learning: the snob program. *The Computer Journal*, **51**, 571–578.
- Kearney, M. (2006) Habitat, environment and niche: what are we modeling? *Oikos*, **115**, 186–191.
- Kearney, M. & Porter, W.P. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334–350.
- Kearney, M., Phillips, B.L., Tracy, C.R., Christian, K.A., Betts, G. & Porter, W.P. (2008) Modelling species distributions without using species distributions: the cane toad in Australia under current and future climates. *Ecography*, **31**, 423–434.
- Keith, D.A., Akçakaya, H.R., Thuiller, W., Midgley, G.F., Pearson, R.G., Phillips, S.J., Regan, H.M., Araújo, M.B. & Rebelo, A.G. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560–563.
- Kolar, C. (2004) Risk assessment and screening for potentially invasive fishes. *New Zealand Journal of Marine and Freshwater Research*, **38**, 391–397.
- Kriticos, D.J., Brown, J.R., Maywald, G.F., Radford, I.D., Nicholas, D.M., Sutherst, R.W. & Adkins, S.W. (2003) SPAnDX: a process-based population dynamics model to explore management and climate impacts on an invasive alien plant, *Acacia nilotica*. *Ecological Modelling*, **163**, 187–208.
- Lavergne, S. & Molofsky, J. (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences USA*, **104**, 3883–3888.
- Leishman, M.R. & Thomson, V.P. (2005) Experimental evidence for the effects of additional water, nutrients and physical disturbance on invasive plants in low fertility Hawkesbury sandstone soils, Sydney, Australia. *Journal of Ecology*, **93**, 38–49.
- Liu, C.R., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385–393.

- Lortie, C.J., Brooker, R.W., Callaway, R.M., Choler, P., Kikvidze, Z., Michalet, R. & Pugnaire, F.I. (2004) Rethinking plant community theory. *Oikos*, **107**, 433–438.
- Mack, R.N. (1996) Predicting the identity and fate of plant invaders: emergent and emerging approaches. *Biological Conservation*, **78**, 107–121.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modeling. *Diversity and Distributions*, **15**, 59–69.
- May, R.M. (1976) Simple mathematical models with very complicated dynamics. *Nature*, **261**, 459–467.
- Mladenoff, D.J. & He, H.S. (1999) Design, behaviour and applications of LANDIS, an object-oriented model of forest landscape disturbance and succession. *Spatial modelling of forest landscape change* (ed. by D.J. Mladenoff and W.L. Baker), pp. 125–162. Cambridge University Press, Cambridge.
- Morin, X. & Lechowicz, M.J. (2008) Contemporary perspectives on the niche that can improve models of species range shifts under climate change. *Biology Letters*, **4**, 573–576.
- Nehrbass, N. & Winkler, E. (2007) Is the giant hogweed still a threat? An individual-based model approach for local invasion dynamics of *Heracleum mantegazzianum*. *Ecological Modelling*, **201**, 377–384.
- Nehrbass, N., Winkler, E., Pergl, J., Perglova, I. & Pyšek, P. (2006) Empirical and virtual investigation of the population dynamics of an alien plant under the constraints of local carrying capacity: *Heracleum mantegazzianum* in the Czech Republic. *Perspectives in Plant Ecology Evolution and Systematics*, **7**, 253–262.
- Panetta, F.D. & Mitchell, N.D. (1991) Homoclimate analysis and the prediction of weediness. *Weed Research*, **31**, 273–284.
- Parker-Allie, F., Musil, C.F. & Thuiller, W. (2009) Effects of climate warming on the distributions of invasive European annual grasses: a southern African perspective. *Climatic Change*, **94**, 87–103.
- Pearson, R.G., Dawson, T.P. & Liu, C. (2004) Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, **27**, 285–298.
- Perrings, C., Dehnen-Schmutz, K., Touza, J. & Williamson, M. (2005) How to manage biological invasions under globalization. *Trends in Ecology and Evolution*, **20**, 212–215.
- Peterson, A.T. & Vieglais, D.A. (2001) Predicting species invasions using ecological niche modelling: new approaches from bioinformatics attack a pressing problem. *BioScience*, **51**, 363–371.
- Peterson, A.T., Papes, M. & Kluza, D.A. (2003) Predicting the potential invasive distributions of four alien plant species in north America. *Weed Science*, **51**, 863–868.
- Peterson, A.T., Stewart, A., Mohamed, K.I. & Araújo, M.B. (2008) Shifting global invasive potential of European plants with climate change. *PLoS ONE*, **3**, 1–7.
- Phillips, B.L., Brown, G.P., Travis, J.M. J. & Shine, R. (2008a) Reid's paradox revisited: the evolution of dispersal kernels during range expansion. *The American Naturalist*, **172**, 34–48.
- Phillips, B.L., Chipperfield, J.D. & Kearney, M.R. (2008b) The toad ahead: challenges of modelling the range and spread of an invasive species. *Wildlife Research*, **35**, 222–234.
- Pimentel, D., Zuniga, R. & Morrison, D. (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, **52**, 273–288.
- Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349–361.
- Pyšek, P. & Richardson, D.M. (2007) Traits associated with invasiveness in alien plants: where do we stand? *Biological Invasions* (ed. by W. Nentwig), pp. 97–122. Springer-Verlag, Berlin Heidelberg.
- Pyšek, P., Richardson, D.M., Rejmánek, M., Webster, G.L., Williamson, M. & Kirschner, J. (2004) Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. *Taxon*, **53**, 131–143.
- Rejmánek, M., Richardson, D.M., Higgins, S.I., Pitcairn, M.J. & Grotkopp, E. (2005) Ecology of invasive plants: state of the art. *Invasive alien species: A new synthesis* (ed. by H.A. Mooney, R.N. Mack, J.A. McNeely, L. Neville, P.J. Schei and J. Waage), pp. 104–161. Island Press, Washington, D.C.
- Richardson, D.M. & Thuiller, W. (2007) Home away from home – objective mapping of high-risk source areas for plant introductions. *Diversity and Distributions*, **13**, 299–312.
- Richardson, D.M., Allsopp, N., D'antonio, C.M., Milton, S.J. & Rejmanek, M. (2000) Plant invasions: the role of mutualisms. *Biological Reviews*, **75**, 65–93.
- Roura-Pascual, N., Suarez, A.V., Gómez, C., Pons, P., Touyama, Y., Wild, A.L. & Peterson, A.T. (2004) Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proceedings of the Royal Society London*, **271**, 2527–2534.
- Roura-Pascual, N., Bas, J.P., Thuiller, W., Hui, C., Krug, R.M. & Brotons, L. (2009a) From introduction to equilibrium: reconstructing the invasive pathways of the Argentine ant in a mediterranean region. *Global Change Biology*, **15**, 2101–2115.
- Roura-Pascual, N., Brotons, L., Peterson, A.T. & Thuiller, W. (2009b) Consensual predictions of potential distributional areas for invasive species: a case study of Argentine ants in the Iberian peninsula. *Biological Invasions*, **11**, 1017–1031.
- Scheller, R.M. & Mladenoff, D.J. (2007) Forest landscape simulation models: tools and strategies for understanding large-scale forested ecosystems. *Landscape Ecology*, **22**, 491–505.
- Scott, J.K. & Panetta, F.D. (1993) Predicting the Australian weed status of southern African plants. *Journal of Biogeography*, **20**, 87–93.
- Sebert-Cuvillier, E., Paccaut, F., Chabrierie, O., Endels, P., Goubet, O. & Decocq, G. (2007) Local population dynamics of an invasive tree species with a complex life-history cycle: a stochastic matrix model. *Ecological Modelling*, **201**, 127–143.
- Skarpaas, O. & Shea, K. (2007) Dispersal patterns, dispersal mechanisms, and invasion wave speeds for invasive thistles. *American Naturalist*, **170**, 421–430.

- Smolik, M., Dullinger, S., Essl, F., Kleinbauer, I., Leitner, M., Peterseil, J., Stadler, L.-M. & Vogl, G. (2009) Integrating species distribution models and interacting particle systems to predict the spread of an invasive alien plant. *Journal of Biogeography*, **37**, 411–422.
- Soberon, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115–1123.
- Taylor, C.M. & Hastings, A. (2005) Allee effects in biological invasions. *Ecology Letters*, **8**, 895–908.
- Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modeling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234–2250.
- Thuiller, W., Richardson, D.M., Rouget, M., Procheş, Ş. & Wilson, J.R. U. (2006) Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology*, **87**, 1755–1769.
- Thuiller, W., Albert, C., Araújo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G.F., Paterson, J., Schurr, F.M., Sykes, M.T. & Zimmermann, N.E. (2008) Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology Evolution and Systematics*, **9**, 137–152.
- Thuiller, W., Albert, C., Dubuis, A., Randin, C. & Guisan, A. (2010) Variation in habitat suitability does not always relate to variation in species' plant functional traits. *Biology Letters*, **6**, 120–123.
- Travis, J.M. J. & Dytham, C. (2002) Dispersal evolution during invasions. *Evolutionary Ecology Research*, **4**, 1119–1129.
- Travis, J.M. J., Münkemüller, T., Burton, O.J., Best, A., Dytham, C. & Johst, K. (2007) Deleterious mutations can surf to high densities on the wave front of an expanding population. *Molecular Biology and Evolution*, **24**, 2334–2343.
- Treier, U.A., Broennimann, O., Normand, S., Guisan, A., Schaffner, U., Steinger, T. & Mueller-Schaerer, H. (2009) Shift in cytotype frequency and niche space in the invasive plant *Centaurea maculosa*. *Ecology*, **90**, 1366–1377.
- Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Rejmánek, M. & Westbrooks, R. (1997) Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, **21**, 1–16.
- Weither, E. & Keddy, P.A. (1995) Assembly rules, null models, and trait dispersion: new questions from old patterns. *Oikos*, **74**, 159–165.
- Welk, E., Schubert, K. & Hoffman, M.H. (2002) Present and potential distribution of invasive garlic mustard (*Alliaria petiolata*) in north America. *Diversity and Distributions*, **8**, 219–233.
- Williams, N.S. G., Hahs, A.K. & Morgan, J.W. (2008) A dispersal-constrained habitat suitability model for predicting invasion of alpine vegetation. *Ecological applications*, **18**, 347–359.
- Willis, K.J. & Whittaker, R.J. (2002) Species diversity-scale matters. *Science*, **295**, 1245–1248.
- Wilson, J.R. U., Richardson, D.M., Rouget, M., Procheş, Ş., Amis, M.A., Henderson, L. & Thuiller, W. (2007) Residence time and potential range: crucial considerations in modelling plant invasions. *Diversity and Distributions*, **13**, 11–22.
- Wilson, E.E., Mullen, L.M. & Holway, D.A. (2009) Life history plasticity magnifies the ecological effects of a social wasp invasion. *Proceedings of the National Academy of Sciences USA*, **106**, 12809–12813.
- Wintle, B.A., Bekessy, S.A., Venier, L.A., Pearce, J.L. & Chisholm, R.A. (2005) Utility of dynamic-landscape metapopulation models for sustainable forest management. *Conservation Biology*, **19**, 1930–1943.
- Wissel, C. (1989) *Theoretische ökologie*. Springer, New York, Berlin.
- With, K.A. (2002) The landscape ecology of invasive spread. *Conservation Biology*, **16**, 1192–1203.
- Woodward, F.I. & Williams, B.G. (1987) Climate and plant distribution at global and local scales. *Vegetatio*, **69**, 189–197.
- Wright, J.W., Davies, K.F., Lau, J.A., McCall, A.C. & McKay, J.K. (2006) Experimental verification of ecological niche modeling in a heterogeneous environment. *Ecology*, **87**, 2433–2439.
- Zalba, S.M., Sonaglioni, M.I., Compagnoni, C.A. & Belenguer, C.J. (2000) Using a habitat model to assess the risk of invasion by an exotic plant. *Biological Conservation*, **93**, 203–208.

## BIOSKETCH

The **EMABIO**'s team (Evolution, Modelling and Analysis of BIOdiversity) led by Wilfried Thuiller focuses at revealing the evolutionary mechanisms that have generated extant biodiversity and its spatial patterns, especially climatic niche evolution and its effect on species diversification. The EMABIO's team also aims to address the mechanisms governing the assembly of biotic communities, such as the genetic effects of keystone species on plant and microbial communities, but also the effects of dispersal limitation, environmental filtering and resource competition. Finally, the EMABIO's team proposes new modelling tools and generate biodiversity assessments at various spatial scales (continental, national and regional).

Author contributions: All authors conceived the ideas and participated in the writing, L.G. led the writing.

Editor: Marcel Rejmanek

## *Chapter 2*



INVASIVE SPECIES DISTRIBUTION MODELS

-

TAKING ADVANTAGE OF THEIR LIMITATIONS

*MODELES DE DISTRIBUTION DES INVASIVES*

-

*COMMENT BENEFICIER DE LEURS*

*IMPERFECTIONS*



## PRESENTATION DU CHAPITRE

*Deux hypothèses clés des modèles de distribution sont nécessairement violées quand on les applique aux cas des espèces invasives. En effet, les espèces invasives ne sont pas à l'équilibre avec leur environnement, et la quantification et transférabilité de leur niche dans le temps et l'espace sont limitées. Dans ce chapitre nous testons si la combinaison de données à l'échelle mondiale et à l'échelle régionale dans un nouveau cadre conceptuel permet de palier à ces limitations. Au delà de simplement améliorer la modélisation des niches régionales pour des espèces invasives, ce nouveau cadre permet de tirer profit de la violation de l'hypothèse d'équilibre, et ainsi d'estimer également le stage d'invasion, le niveau de d'occupation et le risque de propagation dans le future proche pour 27 espèces de plantes invasives dans les alpes françaises.*

*Pour chaque espèce invasive nous construisons trois jeux de modèles de distribution (SDMs): un modèle global et deux modèles régionaux (un modèle conventionnel et un utilisant les résultats du modèle global pour pondérer ses pseudo-absences régionales). Les performances des modèles sont comparées grâce aux indices AUC, TSS, sensibilité et spécificité. Ensuite, les prédictions faites pour chaque site où l'espèce a été observée sont extraites et comparées entre le modèle mondial et les modèles régionaux. Cette comparaison permet d'identifier si les espèces invasives sont observées à l'intérieur de leurs niches globales et régionales.*

*Finalement, cette étude permet de valider la mise en place d'un nouveau cadre conceptuel et méthodologique permettant d'améliorer la modélisation des espèces invasives à l'échelle régionale. Dans cette nouvelle méthode, les prédictions d'un modèle global sont utilisées afin de pondérer les pseudo-absences d'un modèle régional, et améliorent significativement les performances du modèle régional. De plus, les comparaisons des résultats du modèle global avec ceux du modèle régional permettent de révéler différents patrons de niches et différents niveaux d'occupation d'aires potentielles des espèces invasives. Ces différences permettent de conclure sur les stades d'invasion des espèces et de leurs risques de propagation dans un futur proche (tous deux en accord avec les opinions d'experts). Ce cadre peut facilement être appliqué pour un grand nombre d'espèce et ainsi être utile au contrôle des espèces invasives et à la planification de leur éradication.*



MACROECOLOGICAL  
METHODS



# Invasive species distribution models – how violating the equilibrium assumption can create new insights

Laure Gallien<sup>1\*</sup>, Rolland Douzet<sup>2</sup>, Steve Pratte<sup>1</sup>, Niklaus E. Zimmermann<sup>3</sup> and Wilfried Thuiller<sup>1</sup>

<sup>1</sup>Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France, <sup>2</sup>Station Alpine Joseph Fourier, CNRS UMS 2925, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France, <sup>3</sup>Landscape Dynamics, Swiss Federal Research Institute for Forest, Snow and Landscape Research, 8903 Birmensdorf, Switzerland

## ABSTRACT

**Aim** Two core assumptions of species distribution models (SDMs) do not hold when modelling invasive species. Invasives are not in equilibrium with their environment and niche quantification and transferability in space and time are limited. Here, we test whether combining global- and regional-scale data in a novel framework can overcome these limitations. Beyond simply improving regional niche modelling of non-native species, the framework also makes use of the violation of regional equilibrium assumptions, and aims at estimating the stage of invasion, range filling and risk of spread in the near future for 27 invasive species in the French Alps.

**Innovation** For each invader we built three sets of SDMs using a committee averaging method: one global model and two regional models (a conventional model and one using the global model output to weight pseudo-absences). Model performances were compared using the area under the receiver operating characteristic curve, the true skill statistic, sensitivity and specificity scores. Then, we extracted the predictions for observed presences and compared them to global and regional models. This comparison made it possible to identify whether invasive species were observed within or outside of their regional and global niches.

**Main conclusions** This study provides a novel methodological framework for improving the regional modelling of invasive species, where the use of a global model output to weight pseudo-absences in a regional model significantly improved the predictive performance of regional SDMs. Additionally, the comparison of the global and regional model outputs revealed distinct patterns of niche estimates and range filling among the species. These differences allowed us to draw conclusions about the stage of invasion and the risk of spread in the near future, which both correspond to experts' expectations. This framework can be easily applied to a large number of species and is therefore useful for control of biological invasions and eradication planning.

## Keywords

**Adaptation, biological invasions, colonization, ecological niche, equilibrium, invasion stage, non-native plant species, spatial scale.**

\*Correspondence: Laure Gallien, Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France.  
E-mail: laure.gallien@gmail.com

## INTRODUCTION

Invasive plant species pose significant challenges with regard to managing and maintaining indigenous biodiversity in natural ecosystems (Olden *et al.*, 2004). Given that once introduced species become established they are often extremely difficult to eradicate (Rejmánek *et al.*, 2005), preventing their introduction

is by far the most cost-effective form of management. To this end, a range of modelling tools have been developed in order to understand the drivers of species invasions and project the potential distribution of naturalized and invasive species (*sensu* Richardson *et al.*, 2000; to simplify, both naturalized and invasive species will be referred to as *invasives* hereafter) in space or time (Peterson, 2003; reviewed in Gallien *et al.*, 2010). Among

these tools, species distribution models (SDMs; Guisan & Thuiller, 2005), i.e. phenomenological models that statistically relate observed species occurrences to environmental variables, have been used prolifically (see the review in Gallien *et al.*, 2010). They rely on the ecological niche concept and use observed occurrences and thereby model the realized niches of the focal species in the region studied (Pulliam, 2000; Soberón, 2007).

In the context of invasion ecology it is important to specify that there are (at least) three possible views of an invasive species' niche (Gallien *et al.*, 2010). Firstly, the *global niche* corresponds to the broad abiotic and, to a lesser extent, the biotic conditions, under which the species persists. It is built from all data collected across a species' range (i.e. the sum of all its realized niches) and is the most complete estimate of the entire ecological niche without laboratory experimentation (Vetaas, 2002). Secondly, at the scale of the study region, the *regional niche at equilibrium* is limited by both the small-scale abiotic conditions and biotic interactions (e.g. competition, predation, pathogens) in this region (but we often don't have access to it). Thirdly, the *realized regional niche* differs from the regional equilibrium niche when the invader is not in quasi-equilibrium with the regional environment, and is thus limited by abiotic conditions, biotic interactions, invasion history and dispersal constraints (Wilson *et al.*, 2007).

Following this differentiation between invasive species niches, SDMs have been used to predict the potential distributions of invasives in adventive regions (using the realized regional niche; e.g. Rouget *et al.*, 2004), at the global scale (using the global niche; e.g. Beaumont *et al.*, 2009) and even under environmental change scenarios (e.g. Roura-Pascual *et al.*, 2004). However, the suitability of SDMs for modelling invasive species can be questioned on the grounds that two of their critical assumptions are usually seriously flawed. First, SDMs assume that the species' ecological niche is stable in space and time. In other words, the invasive species in its adventive region occupies similar environmental conditions as in the native range. Second, to ensure reliability, SDMs assume that the species of interest is at quasi-equilibrium with the environment in which it occurs. In other words, the invasive species has already reached all suitable places and is absent from all unsuitable sites (Guisan & Thuiller, 2005).

Concerning niche stability in space, realized regional niches may differ significantly between the native and invaded ranges (e.g. Broennimann *et al.*, 2007; Gallagher *et al.*, 2010). Compared to a species' global niche, the realized regional niche corresponding to a new adventive region can either: (1) occupy only a reduced part of the global niche (e.g. due to local biotic limitations such as competitors, predators or pathogens) similar to the realized niche in the native range; (2) occupy only a reduced part of the global niche, different from the realized niche in the native range; or (3) partly occupy areas outside of the global niche thanks to rapid genetic adaptations in the adventive range (Fig. 1 in Gallien *et al.*, 2010). In order to address this issue and to project the potential distribution of an invasive species in an adventive range, it has been recommended that all data available throughout the world (for both native and invasive ranges) are

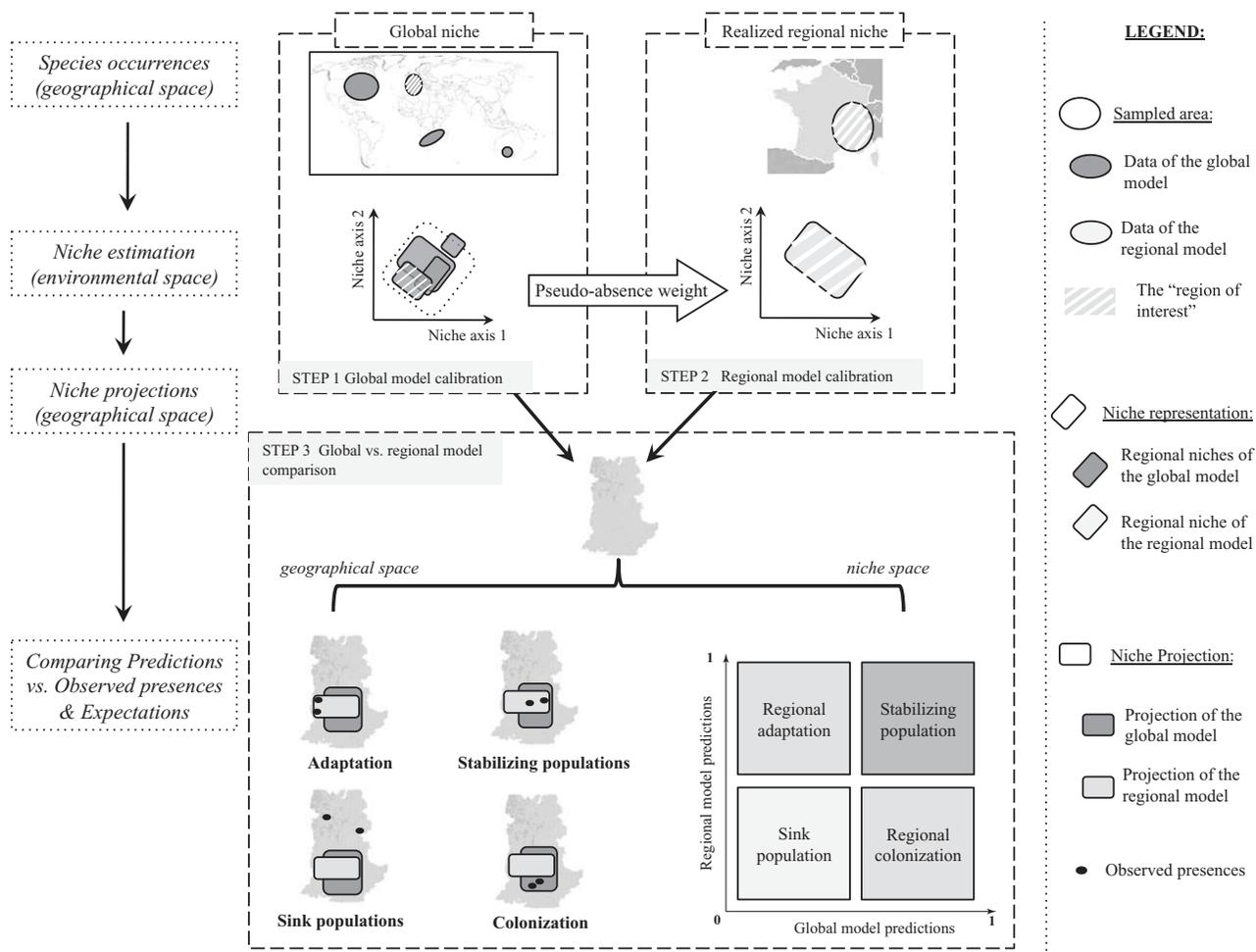
used in order to estimate its ecological niche or at least the full-range biotic and abiotic niche requirements of the species (i.e. *the global niche*; Beaumont *et al.*, 2009; Ibáñez *et al.*, 2009). However, a global description of the niche does not account for the specificities of local adventive ranges (local environment, local biotic interactions and specific human uses). This explains why, when predicting the potential distribution of the species of interest in a specific invaded region, some researchers prefer to use occurrence data from the invaded range only (e.g. Dullinger *et al.*, 2009). Therefore, using a combination of both global and realized regional niches has the potential to produce improved estimates of the potential distribution of a given invader in a study region (e.g. Roura-Pascual *et al.*, 2009).

In theory, the equilibrium between the invader and the environment varies according to the stage of invasion: *introduction*, *colonization* or *establishment* (Theoharides & Dukes, 2007). These invasion stages are themselves strongly influenced by five elements: (1) the introduction history (e.g. propagule pressure, position of founder populations, time of residence; Wilson *et al.*, 2007); (2) the spatial distribution of suitable habitats (Alofs & Fowler, 2010); (3) the invader's characteristics (e.g. dispersal capabilities; Aikio *et al.*, 2010); (4) the invader's potential for rapid adaptation (Travis *et al.*, 2009); and (5) interactions between the invader and the native communities (Davies *et al.*, 2010). These factors make it difficult to infer the stage of regional invasion by simply using distribution data, without prior expert knowledge. Consequently, the observed distribution of the species does not always inform its potential distribution in the region.

Here, we take advantage of the difference between the *realized regional niche* and the *global niche* and propose a novel framework capable of producing more reliable predictions of the distribution of an invasive species in its adventive range of interest, on the one hand, and an improved estimation of its invasion stage and risk, on the other. This framework consists of two development phases (Fig. 1).

Firstly, we develop a hierarchical approach to improve regional SDM performance while simultaneously accounting for both global and regional information (Fig. 1, steps 1–2). When observed absences or pseudo-absence data are used to build a SDM, it is generally assumed that they represent 'true' absences (i.e. sites where the species cannot survive). In the case of invasive species, it is likely that some represent 'false' absences because the species is not at equilibrium (i.e. sites where the species could survive but is currently absent due to dispersal limitations; Le Maitre *et al.*, 2008). This problem can be partly overcome by exploiting the estimation of the invader's global niche to attribute a weight to each (pseudo-)absence, i.e. if the pseudo-absence obtains a low probability of suitability in the global model then it will have a higher weighting in the regional model because it is more likely to be a 'true' absence, and vice versa. This is a first step towards closing the gap between the *realized regional niche* and the *regional niche at equilibrium*.

Secondly, we compare the invader's global (where the species could spread) and realized regional (where the species is already



**Figure 1** Flowchart summarizing the three steps proposed in our analysis. Step 1 represents the global model calibration and step 2 represents the regional model calibration (integrating the output of the global model to weight the regional pseudo-absences); both steps relate to the methodological innovation for modelling invasive species. Then, step 3 presents the theoretical framework developed to inform about the stages of invasion of an invasive species' population. If a species is observed (1) within its regional niche only, then it is potentially a population that develops adaptation to novel environmental conditions. If however a species is observed (2) within both its global and regional niches, then it is likely to be a stabilizing population. If a species is observed (3) only within its global niche, then it may represent a population that participates to the colonization process of various environmental conditions in the region. If finally a species is observed (4) outside both niches, then it probably represents a 'sink' population.

observed) niches with the observed presences to take advantage of both the disequilibrium and the global niche estimation. Theoretically, this comparison allows us to infer both the stage of invasion for each population in the ecological niche space and the degree of regional range filling of the invading species in geographical space (Fig. 1, step 3). In the niche space of an adventive region (assuming the best set of explanatory variables and no data bias), a species is at quasi-equilibrium when at the same time its observed presences are located within both the global and realized regional niches and they fully fill the regional niche range. However, if the regional niche range is not filled then populations are approaching but still away from stabilization. When the species observations instead cover the global niche but not the realized regional niche (i.e. the regional model cannot predict some of the observed presences), then regional

quasi-equilibrium is not yet reached (e.g. colonization from different sources in various environments). Alternatively, if some populations within the realized regional niche are outside the global niche, this indicates that these populations may have adapted to new (abiotic and/or biotic) environments (e.g. rapid local adaptations; Lavergne & Molofsky, 2007). Finally, if a species occurs regionally outside of both the global and the regional niches, then it is probable that the observed presences belong to sink populations (e.g. those introduced into unsuitable areas that are unlikely to provide opportunities for stable population development). Extending these comparisons from niche space to geographical space allows us to infer the degree of range filling for both a species' *global niche* and *realized regional niche*. Combining the information on species' population stages during invasion and range filling has the potential to provide

interesting insights into the invader's future dynamics and potential threat.

Here, we use 27 invasive plant species in the French Alps to test and illustrate the overall development of the approach. The results are then consolidated with the expert knowledge of national botanists. Finally, we make suggestions regarding the further use of the framework to generate testable hypotheses of interest in invasion biology.

## MATERIALS AND METHODS

### The study system

Since 1980, 142 non-native plant species have been identified in the French Alps (source: National Botanical Conservatory of the Alps and Mediterranean, CBN). We used this list of species to create a global and a regional occurrence database. At the global scale, we extracted species occurrences from the Global Biodiversity Information Facility (GBIF; <http://data.gbif.org>) at a minimum resolution of 2.5' (c. 4.5 km). We post-processed the data to remove all records from botanical gardens or those with unrealistic coordinates. At a regional scale, we used the occurrence records from the CBN containing 30 years of botanical surveys at a minimum resolution of 100 m (Boulangeat *et al.*, 2011). We only selected species with more than 200 records in the French Alps to avoid biased estimations of species environmental preferences and to remove casual species. This left us with a dataset of 27 species.

### Modelling framework

At both global and regional scales and for each species we built a set of SDMs – with presence records and randomly generated pseudo-absences – and applied a committee averaging method (Box 1) to extract a single output.

### The global model

The committee averaging method (Box 1) was used to construct an estimation of the species' global niche for each of 27 invasives, using the both the GBIF and CBN presence data with two datasets of 20,000 random pseudo-absences each and the WorldClim climatic database (Hijmans *et al.*, 2005; <http://www.worldclim.org/>). The high number of pseudo-absences artificially reduces the prevalence in the models, which in return influences the probabilities of occurrence of the models (reducing the overall probability values). To avoid this problem here, we did not use the raw probabilities from the models but instead we transformed the probabilities into binary presence/absence data via a threshold (see below). In order to restrict the choice of pseudo-absences to realistically reachable locations, we created a buffer zone of 20 km around any of the presence records used and we randomly allocated absences inside these buffer zones. In this way we avoid areas where invasive species have not been inventoried. This strategy follows the one advocated by Phillips *et al.* (2009) who recommended following the same sampling design for selecting pseudo-absences as for selecting presences. From the 19 available bioclimatic variables we selected the five which had the lowest pair-wise correlations for our dataset (Spearman rank-correlation < 0.6), i.e. (1) maximal temperature in the warmest month, (2) annual temperature range, (3) mean temperature in the coldest quarter, (4) precipitation in the wettest month, and (5) precipitation in the driest month. To model species distributions, we used five algorithms available in the BIOMOD library (version 1.1–6.3; Thuiller *et al.*, 2009) in R (R Development Core Team, 2010): (1) a regression method [generalized additive model (GAM) with four degrees of smoothing and a stepwise variable selection based on the Akaike information criterion (AIC); Hastie & Tibshirani, 1990]; (2) a classification method [classification tree analysis (CTA), with a 50-fold cross-validation; Breiman *et al.*, 1984]; (3) a mix

#### Box 1

Committee averaging method. The committee averaging method is an ensemble forecasting method (Araújo & New, 2007) based on the use of different model algorithms (e.g. regressions, classification trees, machine learning). The rationale of ensemble forecasting is that different algorithms have different levels of accuracy under different circumstances and there is no single perfect algorithm (Elith *et al.*, 2006). In the committee averaging method, predicted probability maps of species presences from the different algorithms are not averaged, but instead are transformed into binary maps (using for each model the threshold that maximizes both sensitivity and specificity) which are then averaged to obtain one single map of the final output. In other words, each model 'votes' for each site whether it forecasts a species' presence or not. It is therefore not a probability of occurrence that is measured but rather a percentage of agreement on species presence between the various

algorithms. The main advantage of the committee averaging method is the use of 'comparable outputs' (binary presence-absences) instead of the raw algorithm outputs (continuous probabilities) that do not necessarily have the same meaning or the same range of variation. This method can easily incorporate the use of: (1) various model algorithms; (2) multiple selections of pseudo-absence data (minimizing the bias due to a specific set of selected pseudo-absences); and (3) several repetitions of cross-validation procedures (calibration and evaluation procedures are repeatedly carried out on different subdatasets). It is also possible to keep only the best performing models (i.e. reliable models only) for the final output by setting a selection threshold based on predictive accuracy metrics. In the end, the number of potential 'voting maps' is: number of algorithms × number of pseudo-absence datasets × number of cross-validation runs.

between regression and classification methods [multivariate adaptive regression splines (MARS); Friedman, 1991]; (4) a boosting algorithm [boosted regression trees (BRT) with the optimal number of trees selected by cross-validation; Ridgeway, 1999]; and (5) a machine learning method [artificial neural networks (ANN) with the best amount of weight decay and the number of units in the hidden layer selected using five-fold cross-validation; Ripley, 1996].

We implemented a split-sample cross-validation procedure to avoid circular reasoning for evaluating the models (i.e. different data portions are used to construct and to evaluate the model; Araújo *et al.*, 2005). Specifically, models were calibrated for each species on 70% of the initial data and then evaluated on the remaining 30% with the true skill statistic (TSS; Allouche *et al.*, 2006) and the area under the receiver operating characteristic curve (AUC; Swets, 1988). The AUC scores vary from 0 for a model whose predictions are systematically wrong, over 0.5 for a random fit, to 1 for a model achieving perfect agreement with the observed data. The TSS has a range of  $-1$  to  $+1$ , with  $-1$  and  $+1$  representing systematically wrong predictions and systematically right predictions, respectively, and 0 representing a random fit.

For each algorithm and for each species, two pseudo-absence datasets were randomly selected, and four cross-validations performed. Forty different models were therefore calculated in total. Of these models, only those obtaining both a TSS and AUC score above 0.6 and 0.8, respectively, were used to build the committee averaging map. There are no specific guidelines for both TSS and AUC scores as they depend on the extent and (obviously) on the goal of the study. We chose 0.6 and 0.8, respectively, based on visual inspections of the output and on the proposed thresholds used in the literature (e.g. Araújo *et al.*, 2005; Engler *et al.*, 2011).

### The regional models

At the scale of the French Alps we used the committee averaging method to model the regional distribution of each invasive species using two different approaches: a conventional approach which used only the data from the French Alps to construct the models, and our proposed approach which additionally integrates global niche information (from the global model built in the section above). The single difference between the two approaches concerns the weights attributed to the pseudo-absences ('true' absence data are not available), after they had been randomly generated.

For both procedures, models were built for the French Alps region at a 100-m spatial resolution using the CBN occurrence records and a set of five environmental variables (four pedoclimatic and one land cover) known to be important for species establishment and spread. Four climatic variables originated from the meteorological model Aurelhy (Bénichou & Le Breton, 1987), based on interpolated measurements at a resolution of 100 m  $\times$  100 m, summarizing climatic information over the last 30 years (here 1971–2000). These variables were: (1) mean annual solar radiation; (2) maximum temperature of the summer as an index of extreme temperatures (this variable was

highly correlated to minimum temperature in winter); (3) standard deviation of annual precipitation as an index of seasonality; and (4) soil water-holding capacity. Land-cover information was extracted from the CORINE Land Cover Map for Europe (as suggested in Polce *et al.*, 2011). Six statistical algorithms (used to attribute different weights to the data) were selected: GLM (generalized linear model), GAM, CTA, GBM, RF (random forest) and ANN. As for the global-scale models, we also ran the regional models with two random selections of pseudo-absence data, followed by four cross-validation repetitions (70–30% as for the global model), giving a total of 48 models per species for each procedure. In order to remove inaccurate models, only those which obtained TSS and AUC scores of over 0.6 and 0.8, respectively, were used to build the committee averaging map.

For the conventional approach, models were constructed using observed species presences and two sets of random pseudo-absence data (10,000 absences per dataset). Traditionally, pseudo-absence data have the same weight as presence data. In other words, the algorithms attribute equal confidence to the pseudo-absence data as to the observed presence data (i.e. pseudo-absences are considered as 'true' absences).

In our proposed approach, models were built using the same observed presences and pseudo-absence data as for the conventional approach, but without assuming that all pseudo-absences represent 'true' absences (i.e. we assume that some pseudo-absences probably reflect environmental conditions where the species cannot survive, while others reflect locations where the species has not yet arrived due to dispersal limitations). We used the global model projections applied to the region to weight each pseudo-absence. Where the global model showed a high level of agreement with the pseudo-absence (i.e. a low habitat suitability) we attributed a high weight to the pseudo-absence (i.e. it probably represents a 'true' absence), and vice versa. The weight was given by means of an inverse logistic transformation (equation 1) to obtain stronger discrimination between the predictions of absences and presences:

$$\text{Weight}(x) = \frac{1}{1 + \left( \frac{\text{projG}(x)}{\text{projG}(x) - 1} \right)^2} \quad (1)$$

where  $\text{Weight}(x)$  is the weight attributed to the pseudo-absence  $x$ , which depends on  $\text{projG}(x)$  the global model prediction at the location of  $x$  [if  $\text{projG}(x) = 1$  then  $\text{Weight}(x) = 0$ ].

### Interpreting outputs

The global and regional prediction accuracy was estimated using the two aforementioned indices (AUC and TSS). At the regional scale, we also compared the regional model predictions with the global model predictions both for the observed presences and for the pseudo-absences separately. The comparison looked at sensitivity (proportion of presences correctly predicted) and specificity (proportion of absences correctly predicted), respectively. Note that a threshold value had to be selected in order to convert continuous model predictions into a discrete prediction

of presences or absences. A threshold of 0.5 is not usually optimal when model outputs are true probabilities (Santika, 2011), but given that the outputs were the agreements between models, we selected this value in order to make sure that the majority of model agreements were decided (i.e. at least half of the models agree), because we wanted to create neither conservative nor liberal predictions.

The estimation of a species' stage of invasion was inferred according to the theoretical framework displayed in Fig. 1 (step 3). It is based on a comparison of the predictions made by the models at the global and at the regional scale for each observed presence in the French Alps. Whether the observed presence is predicted as a presence or an absence in one or both models can theoretically relate to the stage of invasion for each invader's population. In addition, for each species we estimated its range filling in the French Alps (regional scale). We compared three pairs of data in the geographical space: (1) the observed presences against the regional model predictions (Svenning & Skov, 2004); (2) the observed presences against the global model predictions; and (3) the regional against the global model predictions. We used these comparisons as proxies to assess: (1) how well species currently fill their projected regional niche; (2) how well species currently fill their projected global niche; and (3) how (dis-)similar are the regional and the global niche projections. In order to have comparable inputs (occurrence, regional and global projection resolutions), we used a grid at 2.5' resolution (i.e. the global model resolution) where the pixel occupancy of the regional model at 100-m resolution and the observations were scaled up: if at least one of the pixels at 100-m resolution was occupied then the aggregated pixel was considered to be occupied.

## RESULTS

### Model performances

Global models showed good performances at the global scale (AUC > 0.8 and TSS > 0.6 for all species), but low to moderate performances at the regional scale (AUC between 0.15 and 0.8; TSS between 0 and 0.5; Fig. 2a). In other words, an invader's observed presences at regional scale did not fill the ranges predicted by the global models (absences within and/or presences outside of the projected niches). Global-scale information alone was thus not sufficient to predict regional-scale distribution of invaders.

The comparison of the performance of the two regional modelling approaches (weighted versus unweighted) revealed that in 78% of the cases, weighting the pseudo-absences significantly improved discrimination between areas where the species was observed as being present and where it was not recorded (pseudo-absences) (Fig. 2b). This difference was essentially due to the fact that: (1) presences were generally equally predicted by the unweighted model when they occurred outside of the global niche, but (2) absences were better predicted by the weighted model (Fig. 2c).

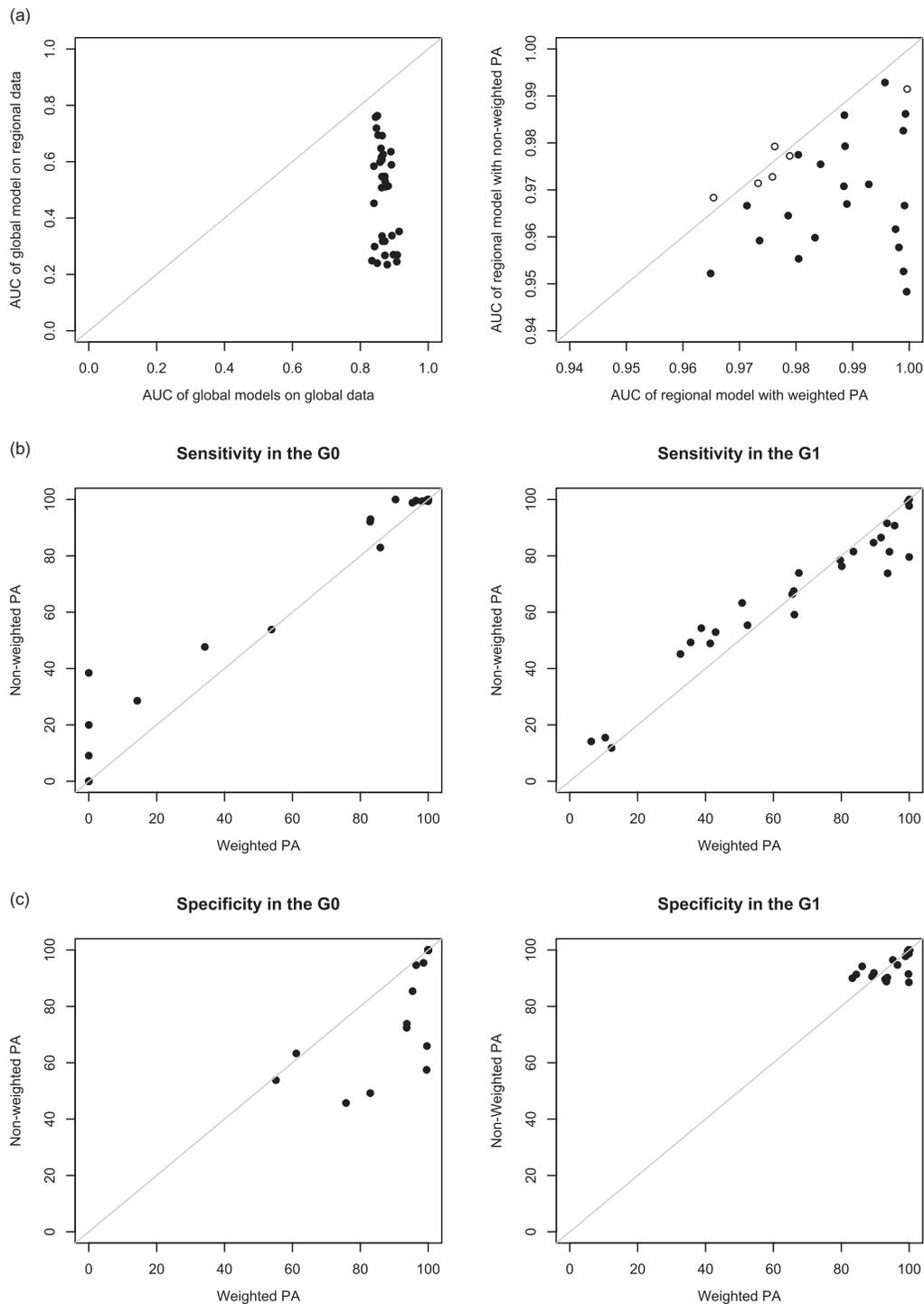
### Stage of invasion

The comparison of the performance of the global and regional models with the aim of inferring the stage of invasion was only carried out using the regional model with weighted pseudo-absences given that its performance was better (or equivalent) to the unweighted model in all comparisons (Fig. 2b).

Using the proposed theoretical framework (Fig. 1) we compared the invader's (global and realized regional) niches with observed presences and inferred the current stage of invasion for our studied species from this comparison. In general, four situations were distinguished (Fig. 1): (1) species with stabilizing populations in the region (e.g. *Panicum capillare*), for which the observed presences were included in both observed regional and global niches; (2) species that are probably undergoing rapid local adaptation (e.g. *Ailanthus altissima*, *Artemisia annua*), for which many observed presences were within the regional niche but outside of the global niche; (3) species that are engaged in the colonization process (e.g. *Solidago gigantea*, *Buddleja davidii*), for which many observed presences were outside of the regional niche but within the global niche; and (4) species with populations at different stages of invasion (e.g. *Sorghum halepense*), for which observed presences were both within and outside of both the regional and global niches. None of the species was found to only have sink populations in the region, probably because we selected species with at least 200 occurrences (see Appendix S1 in Supporting Information for the results on the 27 species).

### Range filling and risk of invasion

The level of range filling in the French Alps complemented the information on the stage of invasion for all species (Fig. 3, Appendix S1). Four situations were distinguished: (1) species at quasi-equilibrium with stabilizing populations widely distributed over the entire regional niche projection (e.g. *P. capillare*), representing a low risk of invasion in the near future; (2) species with stabilizing populations but not filling the entire regional niche projection, representing species that could be classified at risk of invasion in the future because not all suitable environments have yet been occupied; (3) species in disequilibrium due to local adaptations outside of their global niche; (4) species in disequilibrium due to a colonization process. In the specific case of species with populations undergoing local adaptations, if the regional niche projection was filled by presences (e.g. *Ailanthus altissima*) these species could be considered as having a higher risk of spread than if it was not filled (e.g. *Artemisia annua*). The assumption behind it is that a filled niche projection provides more opportunities for gene flow between populations that could increase the spread of the adaptations. In the case of species undergoing a colonization process, if the regional niche projection was filled by presences (e.g. *B. davidii*), then these species could be considered as having a lower risk of spread than if the regional niche projection was not filled (e.g. *Solidago gigantea*) because the suitable but still unoccupied sites were rare.

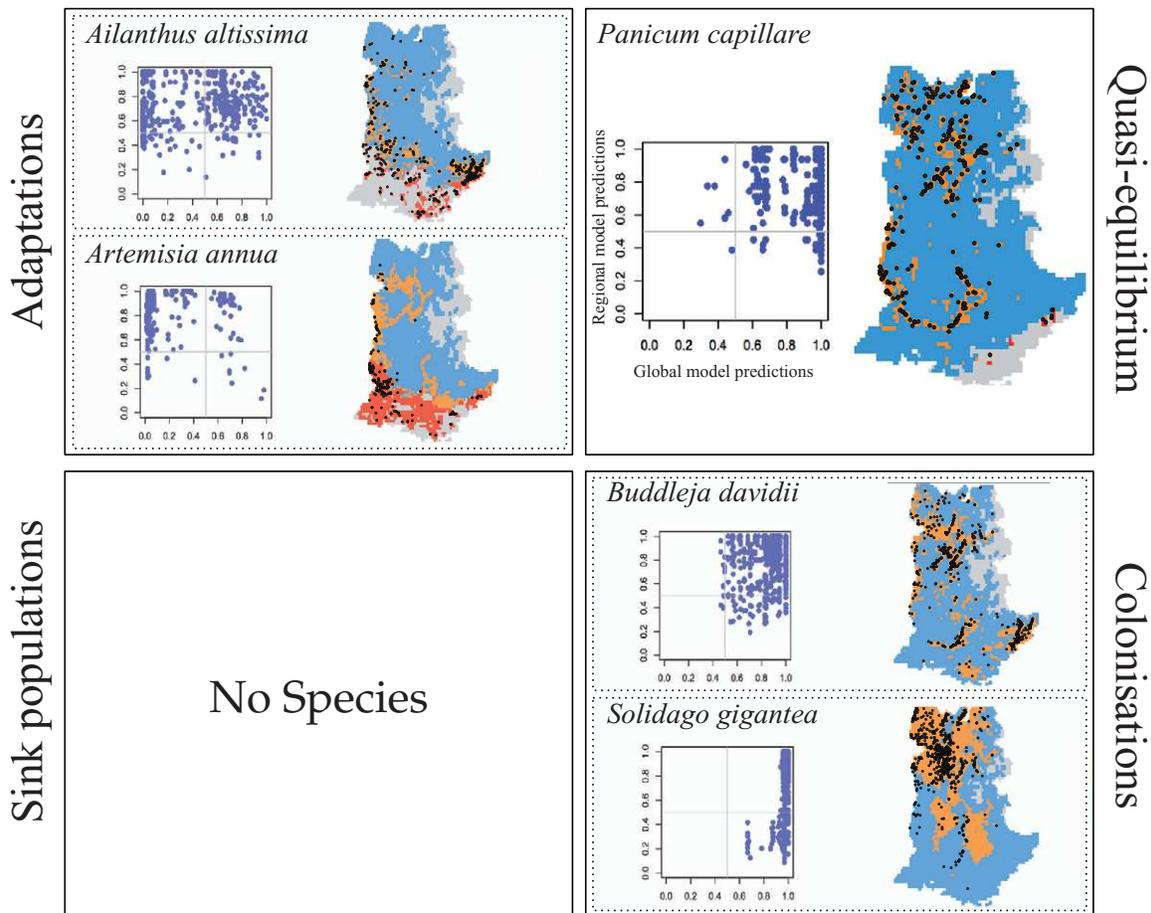


**Figure 2** Comparison of model performance. (a) Global model performance (area under the receiver-operating characteristic curve, AUC) for global versus regional data. (b) Regional model performance (AUC) for weighted versus unweighted pseudo-absences (PA), where filled circles represent significant differences and open circles non-significant differences. (c) Sensitivity and specificity of the two regional models within (G1) and outside (G0) of each species' global niche. The diagonal lines show the values of identical performance between the models compared.

## DISCUSSION

The conceptual and statistical improvement proposed herein relies on the recognition and use of the multiple regional niches

the species exhibited in their global distribution (Gallien *et al.*, 2010). It makes it possible to obtain the most complete estimate of a species' ecological niche from observational data. The use of the global niche allows us to simultaneously obtain a better



**Figure 3** Invasion stages for five invasive species according to our novel methodological framework (Fig. 1). For each species' observed occurrence we plotted the values of the global and regional model predictions, and mapped their geographical representation. Global niches are shown in blue, overlapping regional and global niches in orange, regional niches in red, areas outside of both niches in grey, and observations of occurrences are represented with black dots.

estimation of a species' climatic limitations, and to realistically remove potentially false absences in the regional pseudo-absence datasets. Indeed, because both biotic and abiotic conditions differ between regions and because species climatic tolerances may change (e.g. local adaptation, genetic drift or phenotypic plasticity) the observed niches estimated in different regions (both native and invasive) may vary substantially. Using only one estimate of realized regional niches (e.g. the native range) may misrepresent the species' environmental preferences and result in incomplete predictions (Broennimann *et al.*, 2007; Fitzpatrick *et al.*, 2007).

Weighting the pseudo-absences at a regional scale using information from the global model increased the predictive accuracy of the regional models by: (1) decreasing the influence of (regional) false absences, (2) invariably accepting true (regional) presences, and (3) letting the regional climate, soil and land use refine the regional niche estimation. Other approaches have been proposed to constrain pseudo-absence selection, for example choosing only those found outside the species' climatic tolerance (e.g. Le Maitre *et al.*, 2008), or selecting pseudo-absence data with the same sampling bias as the one of the observed presence

datasets (e.g. Phillips *et al.*, 2009). Although these approaches have the advantage of considering the species' global niche, or diminishing the sampling bias, the former does not ensure the model to match observed regional distribution more closely and the latter does not resolve the problem of false absences.

Our proposed framework is more than just a model improvement as it has the potential to advance our understanding of species invasion and associated risk. Based on theoretical expectations, under the assumption that novel climatic conditions do not largely exceed those present in the global distribution, and when using adequate data, the framework provides insights into characteristics of invader populations in a region, as to whether they are: (1) at quasi-equilibrium, (2) potentially adapting to new local conditions, (3) still in the colonization process, or (4) represent sink populations (Fig. 3). These four different cases can then be formally tested using observations or common-garden experiments.

For instance, *Panicum capillare* L. (an annual tropical grass, Poaceae family) is judged currently to be at low risk of expansion in the French Alps. Most of the species' populations are 'stabilizing' and fill the regional range (Fig. 3). This corroborates

the history of the species, introduced 250 years ago together with corn seeds. It does not spread a long way from the cornfields but has large seed banks and long dormancy making the populations relatively stable.

Amongst the set of investigated invasive species, some are well predicted by the regional model but also seem to occur outside of their global niches, such as *Ailanthus altissima* (Mill.) Swingle and *Artemisia annua* L. (Simaroubaceae and Asteraceae family, respectively; Fig. 3). One basic explanation for this pattern could be methodological. The estimated global niche might not be sufficient for describing a more complex niche (e.g. it may be missing important variables). Another non-exclusive explanation is that individuals within these populations have managed to modify their environmental preferences. This could be due to increased genetic variation from multiple introductions generating genetic novelties through recombination (e.g. Lavergne & Molofsky, 2007), genome characteristics for rapid adaptation (e.g. neopolyploidy; Ramsey & Schemske, 2002) or hybridization with adapted native congeneric species (cf. Dietz & Edwards, 2006). These populations are of particular concern as their distributions are likely to continue to expand in the near future. Interestingly, *Ailanthus altissima* fills a larger part of its regional range than *Artemisia annua*, suggesting that populations of *Ailanthus altissima* that rapidly adapted outside of the global niche have more chance of subsisting and spreading along the colonization front (e.g. via mutation surfing; Travis *et al.*, 2010). This fits with the characteristics of *A. altissima*, a tree introduced 300 years ago for ornamental reasons, which has a high potential of tilling for efficient short-distance dispersal. In contrast, *Artemisia annua* is known to have sporadic populations that are less likely to adapt locally.

Similarly, the risk of future spread can be relatively different for two colonising species such as *S. gigantea* Aiton (Asteraceae) and *B. davidii* Franchet (Buddlejaceae). *Solidago gigantea* has a wider projected regional niche but it has filled its regional niche less than does *B. davidii*. This broad geographical scatter of the former probably results from multiple independent introduction events.

Finally, individual populations of the same species may be at various stages of invasion, like for example *Sorghum halepense*, a subtropical graminoid species. This species harbours a combination of functional trait advantages (perennial, C4 metabolism and vegetative reproduction) and is able to spread and colonize away from cultivated fields, and eventually develop adaptations to new environmental conditions (with the possibility of hybridisation with *Sorghum bicolor* that is cultivated in the region, Morrell *et al.*, 2005).

There are obviously possible pitfalls associated with the proposed framework that need to be carefully discussed. For instance, the capacity of the framework to determine the degree of range filling will ultimately depend on the heterogeneity of the regional environment. In highly heterogeneous environments, a model calibrated at the coarse (global) resolution could fail to fully capture the environmental variability that is observable at the fine (regional) resolution. In our case study, we

up-scaled our projections from the regional model to the same resolution as the global model. This ascertains that the number of occupied pixels becomes comparable, and it gives a first approximation of a species' range filling. However, in a perfect case, the global and regional models should be calibrated at the same resolution with exactly the same variables (climate, land cover, soil information). Having the species distribution and the climate data available at a 100-m spatial resolution at the global scale is currently too demanding. In addition, sampling bias at both global and regional scales could also influence the species niche estimations and thus affects the overall relevance of the analysis. In the case of a strong sampling bias, we would advise the collection of additional data in order to lessen the bias before applying the proposed methodological framework.

In conclusion, we discuss how the proposed framework could also be used to generate testable hypotheses that link the concept of ecological niche to invasion ecology.

1. What makes a good invader? Identifying which functional traits characterize invasive species has a long history in ecology, and some key traits have been consistently reported as favouring invasion (e.g. clonality, high seed production, hybridization potential; Pyšek & Richardson, 2007). Our framework could make it possible to identify, for a large number of species, whether invaders suspected of having evolved their niches (e.g. Beaumont *et al.*, 2009) have certain specific characteristics (e.g. mean functional traits, large intra-specific functional trait variability, phylogenetic position in a rapidly evolving clade, generalist versus specialist species).

2. Could we retrieve the invasion history and dynamics? Being able to re-create the invasion history of a particular species and understand its dynamics is crucial. This is especially true for pest species undergoing evolutionary adaptation in their adventive range. Our framework is capable of identifying the populations that are likely to undergo rapid adaptation, which could further be sampled for trait measurements and genetic analyses (e.g. Albert *et al.*, 2010). In the case of a known invasion history, our framework can help identify which factors contribute to colonization and adaptation or understand whether stabilizing populations are acting as source of gene flow toward non-equilibrium populations.

3. The interactions between the invader and the recipient native community. The reasons why some native communities are more resistant to invasion than others have been under investigation since Darwin's time (Darwin 1859), but this area still requires research (Thuiller *et al.*, 2010). Our proposed framework can be used here to test new hypotheses about the interactions between the native and invasive species, such as whether the types of interaction differ according to the population's stage of invasion.

## ACKNOWLEDGEMENTS

This work was funded by ANR SCION (ANR-08-PEXT-03) project. Thanks to T. Münkemüller, C. H. Albert, K. Schiffers, F. Boucher, Núria Roura-Pascual and two anonymous referees for comments that greatly improved earlier versions of the manu-

script. We would also like to thank the Alpine and Mediterranean National Botanical Conservatories for providing the observed data at regional scale and the list of invasive plant species in the region.

## REFERENCES

- Aikio, S., Duncan, R.P. & Hulme, P.E. (2010) Herbarium records identify the role of long-distance spread in the spatial distribution of alien plants in New Zealand. *Journal of Biogeography*, **37**, 1740–1751.
- Albert, C.H., Edwards, T., Jr, Graham, C.H., Yoccoz, N.G., Zimmermann, N.E. & Thuiller, W. (2010) Applied sampling in ecology and evolution – integrating questions and designs. *Ecography*, **33**, 1028–1037.
- Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.
- Alofs, K.M. & Fowler, N.L. (2010) Habitat fragmentation caused by woody plant encroachment inhibits the spread of an invasive grass. *Journal of Applied Ecology*, **47**, 338–347.
- Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.
- Araújo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005) Validation of species–climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.
- Beaumont, L.J., Gallagher, R.V., Thuiller, W., Downey, P.O., Leishman, M.R. & Hughes, L. (2009) Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions*, **15**, 409–420.
- Bénichou, P. & Le Breton, O. (1987) Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. *La Météorologie*, **7**, 23–34.
- Boulangeat, I., Lavergne, S., Van Es, J., Garraud, L. & Thuiller, W. (2011) Niche breadth, rarity and ecological characteristics within a regional flora spreading over large environmental gradients. *Journal of Biogeography*, **39**, 204–214.
- Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984) *Classification and regression trees*. Chapman & Hall, New York.
- Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701–709.
- Darwin, C.R. (1859) *The origin of species*. John Murray, London.
- Davies, K.F., Cavender-Bares, J. & Deacon, N. (2010) Native communities determine the identity of exotic invaders even at scales at which communities are unsaturated. *Diversity and Distributions*, **17**, 35–42.
- Dietz, H. & Edwards, P.J. (2006) Recognition that causal processes change during plant invasion helps explain conflicts in evidence. *Ecology Letters*, **87**, 1369–1367.
- Dullinger, S., Kleinbauer, I., Peterseil, J., Smolik, M. & Essl, F. (2009) Niche based distribution modelling of an invasive alien plant: effects of population status, propagule pressure and invasion history. *Biological Invasions*, **11**, 2401–2414.
- Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Engler, R., Randin, C.F., Thuiller, W. *et al.* (2011) 21st century climate change threatens mountain flora unequally across Europe. *Global Change Biology*, **17**, 2330–2341.
- Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, **16**, 24–33.
- Friedman, J. (1991) Multivariate adaptive regression splines. *Annals of Statistics*, **19**, 1–141.
- Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology*, **98**, 790–799.
- Gallien, L., Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, **16**, 331–342.
- Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
- Hastie, T.J. & Tibshirani, R. (1990) *Generalized additive models*. Chapman and Hall, London.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Ibáñez, I., Silander, J.A., Wilson, A.M., Lafleur, N., Tanaka, N. & Tsuyama, I. (2009) Multivariate forecasts of potential distributions of invasive plant species. *Ecological applications*, **19**, 359–375.
- Lavergne, S. & Molofsky, J. (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences USA*, **104**, 3883–3888.
- Le Maitre, D.C., Thuiller, W. & Schonegevel, L. (2008) Developing an approach to defining the potential distributions of invasive plant species: a case study of *Hakea* species in South Africa. *Global Ecology and Biogeography*, **17**, 569–584.
- Morrell, P.L., Williams-Coplin, T.D., Lattu, A.L., Bowers, J.E., Chandler, J.M. & Patterson, A.H. (2005) Crop-to-weed introgression has impacted allelic composition of Johnsongrass populations with and without recent exposure to cultivated sorghum. *Molecular Ecology*, **14**, 2143–2154.
- Olden, J.D., Poff, N.L., Douglas, M.R., Douglas, M.E. & Fausch, K.D. (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology and Evolution*, **19**, 18–24.
- Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology*, **78**, 419–433.

- Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J.R. & Ferrier, S. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181–197.
- Polce, C., Kunin, W.E., Biesmeijer, J.C., Dauber, J., Phillips, O.L. & The ALARM Field Site Network (2011) Alien and native plants show contrasting responses to climate and land use in Europe. *Global Ecology and Biogeography*, **20**, 367–379.
- Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349–361.
- Pyšek, P. & Richardson, D.M. (2007) Traits associated with invasiveness in alien plants: where do we stand? *Biological Invasions* (ed. by W. Nentwig), pp. 97–122. Springer-Verlag, Berlin.
- R Development Core Team (2010) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.
- Ramsey, J. & Schemske, D.W. (2002) Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics*, **33**, 589–639.
- Rejmánek, M., Richardson, D.M., Higgins, S.I., Pitcairn, M.J. & Grotkopp, E. (2005) Ecology of invasive plants: state of the art. *Invasive alien species: a new synthesis* (ed. by H.A. Mooney, R.N. Mack, J.A. Mcneely, L. Neville, P.J. Schei and J. Waage), pp. 104–161. Island Press, Washington, DC.
- Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, **6**, 93–107.
- Ridgeway, G. (1999) The state of boosting. *Computing Science and Statistics*, **31**, 172–181.
- Ripley, B.D. (1996) *Pattern recognition and neural networks*. Cambridge University Press, Cambridge.
- Rouget, M., Richardson, D.M., Nel, J.L., Le Maitre, D.C., Egoh, B. & Mgid, T. (2004) Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions*, **10**, 475–484.
- Roura-Pascual, N., Suarez, A.V., Gómez, C., Pons, P., Touyama, Y., Wild, A.L. & Peterson, A.T. (2004) Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 2527–2534.
- Roura-Pascual, N., Bas, J.P., Thuiller, W., Hui, C., Krug, R.M. & Brotons, L. (2009) From introduction to equilibrium: reconstructing the invasive pathways of the Argentine ant in a Mediterranean region. *Global Change Biology*, **15**, 2101–2115.
- Santika, T. (2011) Assessing the effect of prevalence on the predictive performance of species distribution models using simulated data. *Global Ecology and Biogeography*, **20**, 181–192.
- Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115–1123.
- Svenning, J.-C. & Skov, F. (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, **7**, 565–573.
- Swets, K.A. (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293.
- Theoharides, K.A. & Dukes, J.S. (2007) Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. *New Phytologist*, **176**, 256–273.
- Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.
- Thuiller, W., Gallien, L., Boulangéat, I., de Bello, F., Münkemüller, T., Roquet-Ruiz, C. & Lavergne, S. (2010) Resolving Darwin's naturalization conundrum: a quest for evidence. *Diversity and Distributions*, **16**, 461–475.
- Travis, J.M.J., Mustin, K., Benton, T.G. & Dytham, C. (2009) Accelerating invasion rates result from the evolution of density-dependent dispersal. *Journal of Theoretical Biology*, **259**, 151–158.
- Travis, J.M.J., Münkemüller, T. & Burton, O.J. (2010) Mutation surfing and the evolution of dispersal during range expansions. *Journal of Evolutionary Biology*, **23**, 2656–2667.
- Vetaas, O.R. (2002) Realized and potential climate niches: a comparison of four *Rhododendron* tree species. *Journal of Biogeography*, **29**, 545–554.
- Wilson, J.R.U., Richardson, D.M., Rouget, M., Procheş, Ş., Amis, M.A., Henderson, L. & Thuiller, W. (2007) Residence time and potential range: crucial considerations in plant invasion ecology. *Diversity and Distributions*, **13**, 11–22.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Stage and risk of invasion for 27 invasive species in the French Alps.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

## BIOSKETCHES

**Laure Gallien** is currently a PhD student at the Laboratory of Alpine Ecology (LECA) in France. She is interested in alpine plant invasion, where she especially focuses on evolutionary biogeography, community ecology and population genetics.

Editor: José Alexandre F. Diniz-Filho



## *Chapter 3*



ARE BIOTIC INTERACTIONS INFLUENCING  
INVASION SUCCESS?

*EST CE QUE LES INTERACTIONS BIOTIQUES  
INFLUENCENT LE SUCCES D'INVASION?*



## PRESENTATION DU CHAPITRE

*Dans ce chapitre nous avons travaillé sur l'importance des interactions biotiques dans le succès d'invasion à travers trois études successives. Dans un premier temps, nous avons fait une revue de la littérature sur l'hypothèse originalement développée par Darwin qui postule que l'invasion d'une espèce est plus facile lorsque celle-ci est une lointaine parente des espèces natives locales (et donc fonctionnellement différente), car cela lui permet d'utiliser des ressources non utilisées par les autres espèces (opportunité de niche). Cette hypothèse récemment testée sur divers jeux de données, est parfois validée et parfois réfutée. Parmi les raisons qui peuvent expliquer cet absence de consensus, nous nous sommes particulièrement intéressé à l'influence de la résolution des jeux de données utilisés et aux méthodes statistiques utilisées pour inférer les processus d'invasion à partir de leurs patrons. C'est pourquoi dans un deuxième temps nous avons testé l'effet de la résolution spatiale des données sur le processus détecté. Pour ce faire, nous avons étudié les communautés végétales des zones dunaires de la côte ouest italienne échantillonnées à trois résolutions spatiales différentes. Nos résultats confirment les attendus théoriques et montrent bien que les interactions biotiques ne peuvent être détectée qu'à une échelle local, et que lorsque l'on augmente le grain de l'étude seul le filtre environnemental est détecté. Dans un troisième temps, nous avons décidé de tester les différentes méthodes statistiques habituellement utilisées à travers une approche de simulation. Pour ce faire, nous avons développé un modèle qui permet d'assembler des communautés à l'aide de règles connues a priori. Une fois les communautés assemblées, les indices normalement utilisés pour des données réelles sont estimés et testés à l'aide de modèles nuls (méthode de randomisation). Dans cette étude (en préparation), les premiers résultats montrent que tous les indices et toutes les méthodes de randomisation ne sont pas également fiables pour inférer les différents mécanismes d'invasion.*



## *Chapter 3.1*





# Resolving Darwin's naturalization conundrum: a quest for evidence

Wilfried Thuiller<sup>\*†</sup>, Laure Gallien, Isabelle Boulangeat, Francesco de Bello<sup>‡</sup>, Tamara Münkemüller, Cristina Roquet and Sébastien Lavergne<sup>†</sup>

Laboratoire d'Ecologie Alpine, UMR-CNRS  
5553, Université Joseph Fourier, BP 53, 38041  
Grenoble Cedex 9, France

## ABSTRACT

**Aim** The study of biological invasions has long considered species invasiveness and community invasibility as separate questions. Only recently, there is an increasing recognition that integrating these two questions offers new insights into the mechanisms of biological invasions. This recognition has renewed the interest in two long-standing and seemingly contradictory hypotheses proposed by Darwin: phylogenetic relatedness of invaders to native communities is predicted to promote naturalization because of appropriate niche-adaptation but is at the same time predicted to hamper naturalization because of niche overlap with native species. The latter is known as Darwin's naturalization hypothesis.

**Location** Global.

**Methods and Results** We review the studies that have tested these hypotheses and summarize their largely inconsistent outcomes. We argue that most of the inconsistency arises from discrepancies in the applied conceptual frameworks and analytical approaches and not from different model organisms and different ecological contexts. First, observed patterns and results can be seriously flawed by different spatial and phylogenetic scales, which do not equally reveal community assembly mechanisms. Second, different studies have used different metrics, which may test for different specific hypotheses. Thus, we propose a set of metrics derived from the alpha niche concept to measure invaders relatedness to native communities. Finally, approximating species niche differentiation from phylogenetic relatedness is not exempt of assumptions, and invasive species naturalization may result from various ecological mechanisms of biotic resistance that are not necessarily revealed by species phylogeny alone.

**Main conclusions** The quest for resolving the conundrum of Darwin's naturalization hypothesis will only be successful if appropriate scales, metrics and analytical tests are thoroughly considered. We give several recommendations and suggest, whenever possible, to use trait-based measurements of species dissimilarity as the most promising avenue to unravel the mechanisms driving alien species invasions.

## Keywords

Alpha niche, biological invasions, community invasibility, community phylogenetics, ecological niche, environmental filtering, niche differentiation, null models, phylogenetic relatedness.

\*Correspondence: Dr Wilfried Thuiller, Laboratoire d'Ecologie Alpine, UMR-CNRS 5553, Université Joseph Fourier, BP 53, 38041 Grenoble Cedex 9, France.

E-mail: wilfried.thuiller@ujf-grenoble.fr

<sup>†</sup>Contributed equally to the paper.

<sup>‡</sup>Present address: Institute of Botany, Czech Academy of Sciences, CZ-379 82, Třeboň, Czech Republic.

## INTRODUCTION

The rise in human transportation has led to a substantial increase in species movements out of their native geographic ranges, ultimately resulting in biotic homogenization of

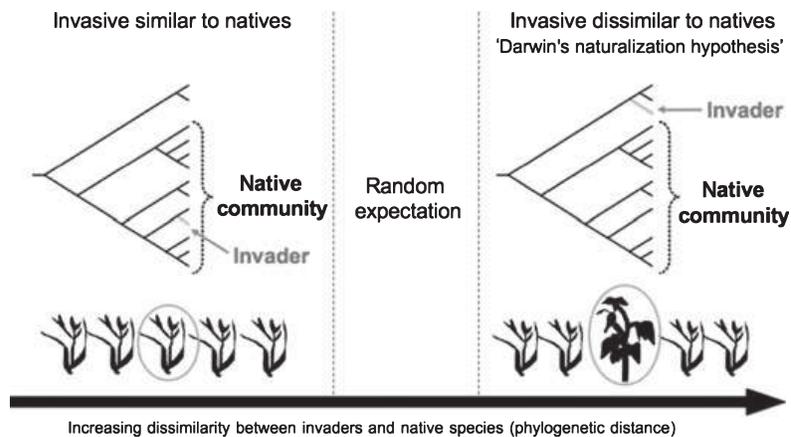
ecosystems world-wide and dramatic changes in ecosystem functioning (Mooney & Hobbs, 2000; Thuiller, 2007). Understanding and predicting the spread and impact of invasive species thus have become central research objectives in fundamental and applied ecology (Nentwig, 2007; Walther

*et al.*, 2009). In particular, invasion ecology has focussed on two questions: (1) which species traits make introduced species more likely to become invaders (Rejmánek, 1995; Thuiller *et al.*, 2006; Pyšek & Richardson, 2007)? and (2) why are some natural communities more prone to invasion than others (Davis *et al.*, 2000; Levine *et al.*, 2004; Tilman, 2004; Richardson *et al.*, 2005)?

Recently, there has been a renewed interest in long-standing hypotheses that merge the two questions by focusing on the phylogenetic relatedness between potential invaders and recipient communities (Fig. 1). Based on an original observation of De Candolle (1855), Darwin (Darwin, 1859), in *The origin of species*, hypothesized that immigrant species are more likely to naturalize when they belong to genera with no native species in the region. This hypothesis, termed ‘Darwin’s naturalization hypothesis’ (Rejmánek, 1996), states that introduced species that are phylogenetically unrelated to local communities should be more successful because they can exploit unfilled ecological niches in native communities (Fig. 1). It implies niche differentiation and niche gap-filling from invaders to be the main drivers of invasion success. However, Darwin also recognized that immigrant species from native genera might have a better chance to naturalize because they share similar pre-adaptations to local environmental conditions with allied species. Following this line of argument, an increase in the phylogenetic relatedness between an introduced species and its recipient community increases its probability of invasion (Fig. 1). This implies that related species have similar environmental requirements and/or benefit from mutualistic or facilitative interspecific interactions because of their shared evolutionary history (Bruno *et al.*, 2003; Wiens & Graham, 2005). These two seemingly contra-

dicting hypotheses, i.e. that introduced species are more likely to naturalize when they are phylogenetically similar versus dissimilar to the native community, have both been originally proposed by Darwin (1859) and are therefore encapsulated under the term ‘Darwin’s naturalization conundrum’ (Diez *et al.*, 2008). Both hypotheses make testable predictions: if species with non-overlapping niches in time or space are more likely to co-exist (Chase & Leibold, 2003), and if species niches have been conserved during evolutionary history, then successful invaders should exhibit a particular phylogenetic position relative to native communities.

A number of recent studies have tested these predictions with empirical data. They have in common that they have treated the two hypotheses as mutually exclusive (with the exception of Diez *et al.*, 2008 and Procheş *et al.*, 2008). However, few, if any, general patterns emerged (Table 1). Of course, the discrepancy between studies may partly be explained by different biological systems and environmental settings that may influence the relative importance of environmental filtering versus biotic interactions in driving community assembly. However, we argue that much of the inconsistency is ostensible and arises from discrepancies in the applied conceptual frameworks and analytical approaches. To our understanding, the main three points that have obscured a general understanding of community invasibility by the mean of species dissimilarity are a matter of spatial and phylogenetic scale, a matter of metric and null expectations and a matter of quantification of niche (dis)similarity. The application of a standard framework across different biological systems should ultimately allow us to assess whether Darwin’s naturalization hypotheses can explain current patterns of biological invasions.



**Figure 1** Conceptual diagram depicting classic hypotheses about species naturalizations and phylogenetic relatedness. Bold arrow at the bottom represents a gradient of species dissimilarity (phylogenetic distance) between invasive and native species. Left panel illustrates the hypothesis that invaders tend to be closely related to native than expected under random expectation. Right panel depicts the actual Darwin’s naturalization hypothesis that immigrant species that phylogenetic unrelated to the native species will be more likely to naturalize because they may harbour different traits (invader represented with a different morphology) and possibly exploit distinct niches than native species. In both panel, the invasive species is represented in grey (grey branch in the phylogenetic tree or grey circle in the drawing of hypothetical community).

**Table 1** Studies on phylogenetic patterns of invasions in a nutshell. Different studies build on different hypotheses (see Fig. 1) and therefore compare successful invaders with different potential species pools (cf. Box 1 in Procheş *et al.*, 2008), use different spatial and phylogenetic scales, consider different additional information and utilize different statistical tests. The conclusions of the different studies vary greatly. Some promote the hypothesis that invaders are phylogenetic more dissimilar to natives than would be expected by chance (cf. 'Darwin's naturalization hypothesis', indicated by + in the conclusions' column), others contradicting this hypothesis (–), finding no clear pattern (0) or draw more detailed conclusions depending on different tested scales or species pools. [The following pseudo-code is used in the table to describe regression models: response variable  $\sim$  explanatory variable<sub>1</sub> + explanatory variable<sub>2</sub>; GL(M)M: generalized linear (mixed) model].

Reference	Taxa	Spatial scale/spatial grain	Phylogenetic level	Statistic model (test)	Species pool	Additional information	Conclusion
Rejmánek (1996)	Plants (Gramineae, Compositae)	California/California	Genus	Number of naturalized species vs. species pool against number of species in European only vs. shared genera (contingency table, Chi-square test)	'Available' species from area of origin (Europe)		+
Daehler (2001)	Plants (Angiosperms)	Hawai/Hawai	Family (pooling multiple genera)	Probability that naturalized species belongs to native genera (expectation under binomial distribution)	1. Global species of families with naturalized species 2. All naturalized (early vs. later naturalized) 3. All accidentally naturalized (early vs. later naturalized)		1., 2. and 3. –*
Duncan & Williams (2002)	Plants (Angiosperms, Gymnosperms)	New Zealand/New Zealand	Genus	Naturalization rate (number of naturalized species as a proportion of pool) $\sim$ 'genus having at least one native species (fixed effect) + family (random effect) (GLMM)	Genera containing introduced species		–
Ricciardi & Atkinson (2004)†	Aquatic systems (fishes, invertebrate, algae and vascular plants)	Global/sites	Genus	Number of high-impact invaders vs. number of low-impact invaders against number of invaders in genera shared vs. unshared with natives (meta-analysis of region-specific contingency tables, Fisher Exact tests)	All invaders		+
Lambdon & Hulme (2006)	Plants	Islands of the Mediterranean Basin/regional, local and habitat	Genus, family, order, subclass	Naturalization status (0/1) $\sim$ presence of congeneric + species variables + island variables (GLM)	Common invaders	Species characteristics, island characteristics, habitat characteristics, introduction frequency	0‡

Table 1 (Continued)

Reference	Taxa	Spatial scale/spatial grain	Phylogenetic level	Statistic model (test)	Species pool	Additional information	Conclusion
Ricciardi & Mottiar (2006)	Fishes	Global/sites	Genus	Number of successful invaders vs. number of failed invaders against number of invaders in genera shared vs. unshared with natives (meta-analysis of region-specific contingency tables, Fisher Exact tests)	All introduced species		0
Strauss <i>et al.</i> (2006)†	Plants (Gramineae)	California/California	Phylogenetic supertree	Phylogenetic distance (mean distance to natives, distance to nearest native relative) ~ pest vs. non-pest invaders ( <i>t</i> -test)	All naturalized	Area of origin	+
Diez <i>et al.</i> (2008)	Plants	Auckland region/ Auckland region, habitat	Genus	1. Probability of naturalization ~ number of native congeners + abundance of native congeners 2. Exotic abundance ~ number of native congeners (region) + abundance of native congeners (region) 3. Exotic abundance ~ number of native congeners (habitat) + abundance of native congeners (habitat) (hierarchical Bayesian framework)	All introduced species	Habitat characteristics, stages of invasion (naturalization and spread), naturalization period	1. – (numb.) + (abund.) 2. 0 (numb.) – (abund.) 3. 0 (numb.) + (abund.)
Diez <i>et al.</i> (2009)	Plants	Australia and New Zealand/ Australia and New Zealand, Australia vs. New Zealand	Family, genus	Probability of naturalization ~ presence of native congeners + genus + climatic origin + family (random effect) (hierarchical Bayesian framework)	All introduced species	Climatic origin	–*

\*The general patterns supported the hypothesis of phylogenetic similarity of invaders and natives (i.e. Darwin's naturalization hypothesis) but there were few families with no pattern or a pattern contradicting the hypothesis.

†These studies test whether the impact (and not the potential to invade) of an invader is driven by phylogenetic dissimilarity.

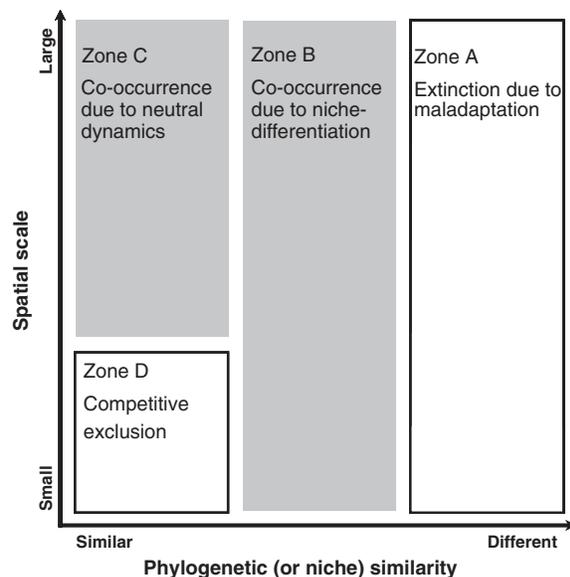
‡The authors concluded that identified patterns were marginal and prone to be generated through artefacts.

## A MATTER OF SCALE

In our opinion, the apparent contradiction between the outcomes of different studies (Table 1) can be partly solved by a detailed examination of the conceptual framework. In particular, we suggest to explicitly consider different spatial and phylogenetic scales (Procheş *et al.*, 2008) and to integrate theoretical understanding of neutral dynamics and niche-based dynamics (Chesson, 2000; Macdougall *et al.*, 2009) when formulating predictions. We use the term spatial scale to refer to the spatial resolution at which invaders naturalization and spread is assessed (e.g. community plot, region, continent), and the term phylogenetic scale to refer to the aggregation level of a phylogenetic reconstruction or classification, that is whether species, genera or families form the leaves of the phylogenetic tree.

### Spatial scale

Spatial scale and phylogenetic relatedness are the two key axes on which each working hypothesis makes unique testable predictions about ecological processes and resulting patterns. A point that has been little recognized, so far, is that these two axes should be considered together (Fig. 2). Under the



**Figure 2** Different predictions for the outcome of species introduction or immigration according to the spatial scale of study and the phylogenetic (or functional) similarity between introduced species and the native species assemblage. Spatial scale corresponds to the spatial scale at which the process of naturalization is quantified, from 'small' (e.g. community or plot level) to 'large' (e.g. regional or continental level). Phylogenetic similarity is the phylogenetic 'distance' of the introduced species to the native species assemblage (see Table 2 for different metrics of phylogenetic relatedness). Each prediction refers to the ecological processes that are hypothesized to be important for one given combination of spatial scale and phylogenetic relatedness.

theoretical assumption of niche conservatism and no possible evolutionary convergence, invaders that are phylogenetically very dissimilar to native communities have an almost null probability of invasion success, no matter of the spatial scale (zone A, Fig. 2). In this case, invaders are assumed to occupy niches so dissimilar to natives that these niches will be virtually absent in the recipient region (e.g. phylogenetically conserved biome affinities (Crisp *et al.*, 2009). In reality, niche conservatism may not be strict and trait or niche convergence may occur because of similar habitats in different regions and may lead to equally adapted but phylogenetically unrelated species. However, in average, phylogenetically very dissimilar invaders can be predicted to be maladapted and to have an increased chance to go extinct because of environmental filtering.

When invaders are phylogenetically moderately dissimilar to native communities, invasions have a much higher chance to be successful independently of the spatial scale (zone B, Fig. 2). In this case, invaders occupy slightly different niches than native species, and there is a good chance that the invader's niche exists in the recipient region. Phylogenetically distinct invaders are able to fill this niche and naturalize if either the niche is empty or the invader is able to outcompete native species occupying the niche (Procheş *et al.*, 2008). This niche-filling process of moderately dissimilar invaders produces a phylogenetic pattern consistent with Darwin's naturalization hypothesis (Ricciardi & Atkinson, 2004; Strauss *et al.*, 2006). When invaders are phylogenetically very similar to native communities, invasion success (comparable to processes of community assembly in general) should be highly dependent on spatial scale (Cavender-Bares *et al.*, 2006; Swenson *et al.*, 2006). At small spatial scales, invaders very similar to native species are more likely to coexist with native species because of competitive exclusion (Gause, 1934; Chesson, 2000) and/or clustering of common enemies such as herbivores and pathogens (Levine *et al.*, 2004; Mitchell *et al.*, 2006). At larger spatial scales, these invaders may be able to co-occur with phylogenetically related native species (at least over ecological time-spans) because of neutral processes and dispersal limitation (Chesson, 2000; Hubbell, 2001).

Thus, analyzing how phylogenetic relatedness between invaders and native species favours or hampers the process of invasion relates to the corpus of coexistence theory that has long explicitly considered the effects of spatial scale (Chesson, 2000). But this conceptual link with the field of biological invasions had been little emphasized so far (Macdougall *et al.*, 2009). Spatial scale may vary from small scales at which species frequently interact and potentially compete (coexistence) to large spatial scales at which species only rarely interact because of dispersal limitations (co-occurrence). As originally pointed out by Lambdon & Hulme (2006) and Procheş *et al.* (2008), the quest for evidence for Darwin's naturalization hypothesis at very large scales is probably meaningless if it supposes to test the outcome of processes of local species interactions and niche gap-filling (Duncan & Williams, 2002; Ricciardi & Mottiar, 2006; Diez *et al.*, 2008, 2009). We argue that, for the same reason, Darwin may have been wrong himself to suggest that

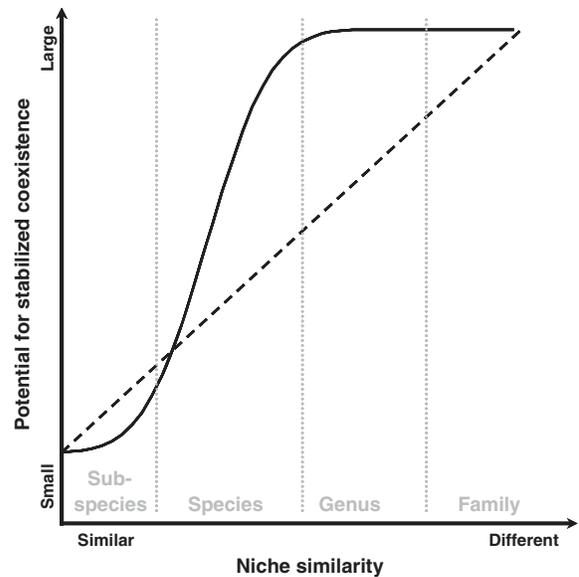
continental scale patterns of species naturalizations would be driven by biotic interactions between species. Patterns observable at large scales very unlikely reflect the outcome of species interactions but more likely reflect environmental filtering, regional heterogeneity and species dispersal abilities. Accordingly, studies performed at large spatial scales tend to confirm that introduced species naturalize when they have allied species in the native flora (Duncan & Williams, 2002; Diez *et al.*, 2009). Recent studies considering (even if only superficially) spatial scale showed that the scale alters the relationship between phylogenetic relatedness and invasion success (Lambdon & Hulme, 2006). In conclusion, when carefully considering the appropriate spatial scale, these two hypotheses are not conflicting but compatible in a common framework.

### Phylogenetic scale

The outcome of tests for phylogenetic patterns of invasions also depends on the phylogenetic scale (Procheş *et al.*, 2008). Species dissimilarities can be quantified from phylogenetic distances obtained from phylogenetic reconstructions or classifications, assuming that species relatedness captures (dis)similarity of ecological niches (discussed in the following text). Potential biases arising from the use of phylogenetic data to capture species ecological similarity are twofold.

First, as shown by studies of community assembly, phylogenetic scale can impact the observed signal of phylogenetic structure and therewith dissimilarity measures in natural communities (Hardy & Senterre, 2007). This can be illustrated by plotting – in an over-simplified example – a linear relationship between species niche similarity (represented across increasing aggregation levels) and the potential of species coexistence because of stabilizing mechanisms (Fig 3). If we measure differences between an invasive species and native species based on a small aggregation level (e.g. subspecies or species), we are in principal able to capture fine dissimilarities and therefore test for Darwin's naturalization hypothesis (Fig. 1). However, when measuring at higher aggregation levels (e.g. genera or family), we can only capture broad dissimilarities, which may preclude any accurate test of Darwin's naturalization hypothesis. For example, when niche differentiation is occurring only at lower aggregation levels (e.g. subspecies can invade, congeners not), an analysis conducted at higher levels (e.g. genus or family) will not detect any effect of phylogenetic relatedness although it is there. Instead, the effect of environmental filtering for broad adaptations would be prominent.

Different studies that have tested the effect of phylogenetic relatedness between introduced and native species on the outcome of invasions have used very different phylogenetic or taxonomic information (Table 1), ranging from very crude classification at the family level (e.g. Daehler, 2001) to more detailed classifications at the genus level (e.g. Diez *et al.*, 2008) and fully resolved phylogenies at the species level (e.g. Strauss *et al.*, 2006). As outlined earlier, this variation in the precision of phylogenetic information can strongly influence the



**Figure 3** Diagram of possible relationship between species similarities, the phylogenetic scale similarities are measured at and the potential for stabilizing coexistence dynamics. Sketched are a linear (dashed line) and a step-wise relationship (continuous line) between similarity and coexistence. According to this diagram and the underlying assumptions, some tests for invasion patterns necessitate small phylogenetic scales (subspecies and species) because strong species similarity can only be observed at these scales and a positive relationship between similarity and coexistence potential may only occur at these scales (given a step-wise function).

observed patterns of phylogenetic relatedness and invasion success and precludes any generalization, a problem widely underestimated.

Second, the majority of studies so far implicitly assume a linear relationship between species similarity and the potential for stabilized coexistence, i.e. increasing dissimilarity at all phylogenetic scales is related to constantly increasing stabilized coexistence. Coexistence theory does not necessarily support this assumption. There may not be an advantage of further differentiation when species already occupy different niches. However, if we relax the over-simplified assumption of a linear relationship between similarity and coexistence, potential limitations related to the choice of a large phylogenetic scale become evident. For example, when assuming a more step-like relationship, studies measuring dissimilarities at higher phylogenetic scales (e.g. genus or family) will tend to find no relationship between similarity and invasion success (Fig. 3).

To conclude, Darwin's naturalization hypothesis explicitly refers to direct species interactions as driving mechanisms for coexistence patterns. With increasing spatial and phylogenetic scales, these direct interactions get less and less important, and it therefore is not relevant to test Darwin's naturalization hypothesis at large scales. If the chosen phylogenetic scale is too broad, we expect phylogenetic similarity to be unrelated to invasion success, i.e. patterns that do not differ from random null-models (Table 1).

## A MATTER OF METRICS AND NULL MODELS

Darwin's naturalization conundrum is at least partly because of the diverse body of statistical metrics and models that have been utilized to test the specific predictions (Table 1), sometimes without clearly stating the underlying assumptions and limitations. We propose a set of metrics to quantify the similarity of introduced species to native communities that include both established and new approaches. We further propose a guideline for choosing the appropriate null model to test the 'significance' of the observed patterns and clarify the underlying theories of different commonly used null models.

### Phylogeny versus taxonomy

With the exception of very few studies (e.g. Strauss *et al.*, 2006), the relatedness between introduced and native species has only been quantified on the basis of taxonomic classification (Table 1). Many studies have hypothesized that invasion was related to the number of congeneric species occurring in the same region or habitat (Daehler, 2001; Duncan & Williams, 2002; Lambdon & Hulme, 2006), or alternatively related to the mean abundance of congeneric species (Diez *et al.*, 2008). In theory, the number and abundance of species congeneric to the invader (Table 2) are poor predictors of introduced species relatedness to the native flora, as they both assume that all congeneric species are equally related and that this relatedness does not vary between genera (coarse phylogenetic scale). Not only some genera had such a history of diversification that even congeneric species can be very distantly related and exhibit very dissimilar traits or niches (e.g. Hughes & Eastwood, 2006). Also, different genera largely differ in their evolutionary age so that species relatedness cannot be considered constant between different genera. Although clearly practical, it is not a generally acceptable assumption to consider that all species of a given genus are equally similar (in terms of niches), especially when this genus is more than 50 million years old, has a world-wide distribution and counts several hundreds of species. Furthermore, such an approach makes the results highly sensitive to taxonomy biases that may exist between different clades or biogeographic regions.

For the above-mentioned shortcomings, we advocate that taxonomy-based metrics of species relatedness should be avoided. With the increasing availability of gene sequence data and computational methods for the reconstruction of phylogenies with several thousand of taxa (e.g. Smith *et al.*, 2009), it is now a very reasonable objective to obtain a molecular phylogenetic tree for any biogeographic study performed at large spatial scale. Supertrees can also be combined from several published phylogenies in order to produce the most up-to-date phylogenetic hypotheses (Bininda-Emonds *et al.*, 2002). However, current available sequences or phylogenetic data may not always allow obtaining a fully resolved phylogenetic tree, which may limit analytical power, as discussed earlier ('Phylogenetic scale' section). Alternatively, some studies have focused on specific taxa for which fully

resolved phylogenies are available (e.g. Poaceae, Strauss *et al.*, 2006). Thus, independently of whether and how evolutionary relatedness captures niche similarity, we strongly recommend measuring invasive and native species relatedness from phylogenetic distances more than from taxonomic (mostly morphological) classifications.

### The $\alpha$ niche conceptual framework

Surprisingly, previous tests of Darwin's naturalization hypothesis have failed to relate it to the concept of  $\alpha$  niche. The  $\alpha$  niche of a species is a community-scale measurement that quantifies the resources exploited by this species in comparison with the resources exploited by co-existing species (Pickett & Bazzaz, 1978). In other words, the  $\alpha$  niche corresponds to a niche differentiation between a species and its community neighbours. It can thus be applied to functional and phylogenetic dissimilarities. For example, Ackerly & Cornwell (2007) define the  $\alpha$  trait niche as the deviation for a given species trait from the community average trait value. Thus, if a species exhibits functional traits radically different from the rest of the community where it occurs, it is considered to occupy a different trait niche than co-occurring species, probably reflecting a different resource use (Stubbs & Wilson, 2004).

This  $\alpha$  niche concept should be considered when testing whether naturalizations of introduced species are favoured or hampered by their dissimilarity to native communities, i.e. whether a species invasion success is driven by its potential  $\alpha$  niche in recipient communities. Following the niche gap-filling, a potential invader with a high  $\alpha$  niche has more chance to invade than an invader with a low  $\alpha$  niche value. Indeed, because of expected lower niche overlap with high  $\alpha$  niche, it is less likely to be under strong competitive pressure or to suffer from pathogens or herbivore attacks.

By analogy and in the absence of relevant functional traits, the  $\alpha$  phylogenetic niche of a species could be estimated as the mean phylogenetic distance of the species to the rest of the community. The  $\alpha$  phylogenetic niche can be estimated relative to the overall phylogenetic position in the community or just within a given family or guild depending on the objectives and hypotheses under investigations. The  $\alpha$  phylogenetic niche is the optimal representation of phylogenetic similarity between an invasive species to the rest of the community allowing to properly test Darwin's hypotheses.

In conclusion, a species  $\alpha$  niche can be computed from species dissimilarities matrices that are derived from a species phylogenetic tree or from a species classification based on functional traits. Testing Darwin's naturalization hypothesis implies testing how introduced species naturalizations are related to their  $\alpha$  phylogenetic niche or their  $\alpha$  trait niche.

### Different metrics to describe invader relatedness to native communities

To measure an invader's relatedness to native communities, we propose to use the  $\alpha$  niche concept together with a set of

**Table 2** Overview of different metrics, which can be used to quantify the phylogenetic relatedness (or niche/functional similarity) between an immigrant species and the native species assemblage (community, regional species pool). Each metric is described, along with its assumptions. The word distance refers to a quantitative measurement of species similarity that can be equally obtained from a phylogenetic reconstruction or a multivariate analysis of functional traits or niche dimensions.

Measure	Description	Assumptions
NCS Number of congeneric species	Number of native species belonging to the same genus than the invader	All species within the same genus are equally related (and functionally similar), and this relatedness does not vary between genera. Each native species contribute equally to the overall biotic resistance (or facilitation) of the community
MACS Mean abundance of congeneric species	Mean abundance of native species belonging to the same genus than the invader	All species within the same genus are equally related (and functionally similar), and this relatedness does not vary between genera. The contribution of each species to the overall biotic resistance (or facilitation) of the community depends on its relative abundance
MDNS Mean distance to the native species	Mean distance between each native species and the invader	The similarity of species niches or traits is well captured by their phylogenetic distance Each native species contribute equally to the overall biotic resistance (or facilitation) of the community
WMDNS Weighted mean distance to the native species	Mean distance between each native species and the invader, weighted by the abundances of native species	The similarity of species niches or traits is well captured by their phylogenetic distance The contribution of each species to the overall biotic resistance (or facilitation) of the community depends on its relative abundance
DMANS Distance to the most abundant native species	Distance between the invader and the most abundant native species	The similarity of species niches or traits is well captured by their phylogenetic distance Community's biotic resistance (or facilitation) is driven by the most abundant native species
DNNS Distance to the nearest native species	Distance between the invader and the closest native species	The similarity of species niches or traits is well captured by their phylogenetic distance, but relationship between coexistence potential and phylogenetic distance is stepwise (See Fig. 3) Community's biotic resistance (or facilitation) is driven by the native species that is the closest relative (or the most functionally similar) to the invader

distance-based metrics that can be computed from dissimilarity matrices (phylogeny or trait-based) and data of community structure (Table 2). If based on a phylogenetic tree, interspecific distances can be calculated using branch length information or by counting the number of nodes separating pairs of species when branch length information is missing. The following four metrics are adapted from or equal to metrics that are commonly used to depict the phylogenetic structure of natural communities (Kembel & Hubbell, 2006; Hardy & Senterre, 2007). First, the MDNS metric – Mean Distance of the introduced species relative to the Native Species – is the most straightforward one as it is directly derived from Ackerly & Cornwell's definition of  $\alpha$  niche (2007). Second, one may hypothesize that all native species will contribute to the overall biotic resistance (or facilitation) of the community according to their relative abundances, because they most likely interact with the invader for resource competition and are more likely to attract natural enemies (or conversely, pollinators). This motivates the use of the WMDNS metric – Weighted Mean Distance from the invader to all Native Species – in the community (which is analogous to the measure of functional and phylogenetic diversity including species relative abundances, De Bello *et al.*, 2009). Third, when a limited number of dominant species play a

central ecological role (so-called structuring or keystone species), one may use the DMANS metric – Distance to the single Most Abundant Native Species – in the community. Fourth, the biotic resistance (or facilitation) of a given community could be mainly driven by the native species that is most closely related to the invader because this close relative will most likely compete with similar resources, attract harmful enemies or conversely attract adapted pollinators or dispersers. In this case, an appropriate metric would just be the DNNS – Distance of the invader to its Nearest Native Species – in the native community (Table 2).

Each stage of species naturalization and invasion of native communities is certainly driven by a unique set of ecological mechanisms, and different mechanisms of biotic resistance (resource competition, apparent competition through natural enemies) may show different degrees of frequency dependence. Hence, we suggest that all four metrics should generally be calculated and compared when testing patterns of phylogenetic similarity and invasion success of introduced species.

### Choosing an appropriate null model

The proposed reference set of (phylogenetic or functional) similarity metrics can be used to relate similarity patterns – at

different spatial scales ranging from local sites to continents – to invasion measures such as probability of naturalization (Diez *et al.*, 2008), probability of becoming a pest (Strauss *et al.*, 2006) or regional spread measures (Lambdon & Hulme, 2006). However, on its own the reference set of similarity metrics only describes patterns and does not test to which extent observed patterns of phylogenetic dissimilarity are important for invasion success. To properly test this question, the observed patterns of dissimilarity must be compared with null models to confront observed patterns with random expectations. Again, comparative tests of observed patterns against null model patterns only make sense if applied at the appropriate scale and within a proper randomization scheme. We argue that for testing Darwin's naturalization hypothesis, patterns should be analyzed at the local community scale where individuals interact. Finally, the remaining challenge is to determine the appropriate null model.

In order to test whether species invasions are favoured by their phylogenetic relatedness to native species, the employed null model must break down the phylogenetic relationship between introduced and native species but not the phylogenetic relationships between native species of recipient communities. In short, the question tested by the null model has to be why is the invader and not another species of the available pool of potential invaders entering the recipient community? Here, we are interested with testing the mechanisms of a single event of the process of community assembly (i.e. the immigration of a non-native species), and the implementation of the randomization algorithm must not change the other evolutionary and ecological mechanisms that led to the current structure of native communities. Thus, the null model that seems the most appropriate to us keeps the structure of native communities unchanged and instead swaps the invader along the phylogenetic tree (or the functional dissimilarity matrix) containing all the species of a given species pool. This step is repeated  $N$  times to get a probability distribution of the metric (Table 2) under random conditions, and use the cumulative probability of the observed metric given this random distribution to test the observed value. Note that other null models might be applied, such as swapping each invader between all possible communities or only between communities where it could occur given its niche, or alternatively keeping the invader unchanged and swapping all co-occurring species. But these models seem less desirable as they would preclude the use of abundance-based metrics of  $\alpha$  niche (Table 2) or require more elaborate tests (i.e. Hardy & Senterre, 2007).

According to the hypotheses and assumptions, different null models can be generated, i.e. different algorithms can be used to swap the invader identity along the phylogenetic tree (Procheş *et al.*, 2008). Probably, the simplest algorithm is to draw these 'null' invaders from the entire regional species pool. When working on large environmental gradients, the species pool used for randomization in each given community can be constrained by previously inferred species niches along these gradients. Also, 'null' invaders could be drawn from a pool of

all species that have been introduced in the study region but never naturalized. This would test for analogous patterns (but at a lower spatial scale) than the ones tested in (Duncan & Williams, 2002; Diez *et al.*, 2008, 2009).

However, the general approach we outline here may have some pitfalls. First, community data generally contain *a posteriori* information of invasion. Nothing is known about the community structure before the invader's arrival, and especially it is unclear whether native species have been displaced by the invader. One way to get around this problem may be to compare the structure of invaded communities to the one of non-invaded communities occurring nearby or under similar environmental conditions. Second, a problem may arise when native communities are potentially invaded by more than one species. When testing for the patterns of one invasive species, the other invasive species occurring in the same communities could be considered 'native' for running the randomizations. However, this could lead to erroneous interpretations of observed patterns of species similarities, especially when closely related invaders have strong facilitative interactions with each other (Simberloff & Von Holle, 1999). This is a critical area of research, which needs further improvements given that multi-species invasions are not uncommon. Developing appropriate null-model for such complex situation where multi-species invasions occur in a given community is a pre-requisite for understanding invasion patterns in highly invaded regions.

## A MATTER OF IDENTIFYING NICHE SIMILARITY

### Species phylogenetic relatedness and niche similarity

An important assumption of both hypotheses in Darwin's naturalization conundrum, implicitly made by Darwin (1859) himself, is that species niches are conserved over time, so that closely related species should tend to have more similar niches than distantly related ones (Wiens & Graham, 2005). Although evidence for niche conservatism has been reported before (Peterson *et al.*, 1999; Prinzing *et al.*, 2001; Losos & Glor, 2003), the scope of niche conservatism may not be as broad as previously thought (Wiens & Graham, 2005; Losos, 2008). The observation that ecological niches have been very labile in the evolutionary history of some taxa, with some evolutionary convergences towards similar niches in distinct taxa, severely challenges the conventional assumption of niche conservatism (Pearman *et al.*, 2007). So far, no clear consensus has been reached about whether phylogenetic niche conservatism is an appropriate baseline assumption or not.

One potential limitation of Darwin's naturalization hypothesis lies in the assumption of most phylogeny-based approaches that phylogenetic relatedness equally predicts niche similarity for any pair of species, i.e. that the strength of phylogenetic niche conservatism is constant over the phylogenetic tree relating all study species. In fact, although this assumption has clear computational advantages, deviations from it may potentially distort the observed patterns.

Diez *et al.* (2009) nicely demonstrated that although there is an overall positive effect of congeneric species occurrence on the naturalization of introduced species, the effect estimate is highly variable between different families and even negative in a few families. This suggests that the strength of phylogenetic niche conservatism, hence the strength of biotic resistance mediated by native species related to the introduced species, is variable between different clades and that in certain clades, biotic resistance is driven by species that tend to be distantly related to the invader – probably because for evolutionary convergence. We thus recommend that, to be completely heuristic, further tests of Darwin's naturalization hypothesis should quantify the strength of phylogenetic niche conservatism (Blomberg *et al.*, 2003; Ackerly, 2009), to assess whether deviations from this assumption in certain clades will likely blur the observed patterns of phylogenetic relatedness and invasion success in introduced species.

The major restriction of using phylogenetic relatedness as a proxy of niche similarity is also the difficulty of relating it directly to the multidimensional nature of species niches. This has been viewed, originally, as a benefit given the premise that it is relatively easier to find pattern with a conglomerate measure like phylogeny than with individuals' traits, which can bring noisy information when intra-specific variability is too large (Albert *et al.*, 2010). However, the multidimensionality nature of species niches may lead to opposite processes resulting in random patterns from a phylogenetic point of view. For example, if pollination or other mutualistic interactions cause introduced species to naturalize where allied native species also occur (increased phylogenetic similarity of invaders and recipient communities), competition for resources or natural enemies may alternatively cause introduced species to naturalize in zones where they are phylogenetically novel (decreased phylogenetic similarity of invaders and recipient communities). This can cause the overall effect of phylogenetic relatedness on species naturalization to be very low or null, even though meaningful ecological mechanisms are currently at work. While we share the view that a phylogenetic approach to biological invasions is highly insightful, we foresee that, as large databases of species traits are being elaborated, a trait-based approach will be an interesting approach to unravel the ecological mechanisms that make introduced species naturalize and spread.

### Phylogeny and mechanisms of biotic resistance

Alien species naturalizations are potentially driven by a combination of ecological mechanisms, but it is not known which mechanisms of biotic resistance are likely to be modulated by phylogenetic relatedness between alien and native species. Invasive species interfere with a number of ecological processes within recipient ecosystems, including resource capture (Levine *et al.*, 2003, 2004), interactions with natural enemies such as herbivores or pathogens (Keane & Crawley, 2002; Hawkes, 2007) and mutualistic interactions such as with pollinators, seed dispersers or mycorrhizas

(Klironomos, 2002; Memmot & Waser, 2002; Milton *et al.*, 2007; Vilà *et al.*, 2009). Here, we review the relevant empirical work to assess the evidence for phylogenetic signals on how biotic resistance is mediated by resource capture, herbivory and pollination.

The breadth of environmental tolerance of closely related invasive and native species has been little investigated with experimental approaches (Brock & Galen, 2005; Braby & Somero, 2006; Geng *et al.*, 2006; Priddis *et al.*, 2009), so evidence for differences between invasive and native relatives in their response to large environmental gradients is limited. However, a larger number of studies support the hypothesis that ecophysiological traits can differ between closely related invasive and native plant species (Schierenbeck & Marshall, 1993; Mc-Dowell, 2002; Deng *et al.*, 2004; Willson *et al.*, 2008). Nevertheless, at least partial niche differentiation between congeneric species can be expected, suggesting that competition for resources between closely related plants may not be as intense as previously thought. To our knowledge, only one study tried to relate experimental measurements of competitive interactions to species phylogenetic relatedness (Cahill *et al.*, 2008). Based on a meta-analysis of competitive experiments involving 142 plant species, the authors show that Darwin's statement that competition should be stronger among closely related species is not supported and suggest that the outcome of interspecific competitive interactions should be better captured by species functional traits. Moreover, further experimental work suggest that competitive hierarchy between invasive and closely related native species can be reversed depending on resource availability (Burns, 2004; Garcia-Serrano *et al.*, 2007). Therefore, if the naturalization of introduced species is mainly driven by interspecific competitions, empirical evidence suggests that the importance of phylogenetic relatedness to native communities will be limited.

However, there is much more evidence that trophic or antagonistic interactions are influenced by species phylogenetic relationships, such as in prey–predator, prey–parasitoid, plant–parasitic fungi systems (Cattin *et al.*, 2004; Ives & Godfray, 2006; Vacher *et al.*, 2008; Rezende *et al.*, 2009). According to Darwin's naturalization hypothesis, biotic resistance of plant communities facing new invaders is predicted to be higher when native communities include species that are closely related to the introduced species. When natural enemies such as herbivores and pathogens mediate biotic resistance, it should be expected that introduced species that are phylogenetic novel to a native community should suffer less antagonistic interactions and ultimately be more likely to naturalize or increase in abundance. This prediction has actually received much support in the recent years. A number of studies have demonstrated that woody and herbaceous aliens tend to suffer less from herbivores attacks (mostly insects) when they were less related to the native flora (based on phylogenetic or taxonomic information), and this pattern seemed consistent between different bioclimatic contexts (Brändle *et al.*, 2008; Brown & Zuefle, 2009; Burton Hill & Kotanen, 2009; Dawson *et al.*, 2009; Pearse & Hipp, 2009), but

see Zuefle *et al.* (2008). Indeed, it seems that, for plant invasions, there is empirical support for a mechanism of Darwin's naturalization mediated by insect herbivores.

There is also some indication in the literature that networks of mutualistic interactions can be phylogenetically structured like for instance plant–pollinators, plant–frugivores and plant–mycorrhizal networks (Maherali & Klironomos, 2007; Rezende *et al.*, 2007). Concerning plant–pollinator interactions, there is evidence that closely related species may tend to have at least partly overlapping pollinator fauna and that pollinators can switch from native plants to congeneric aliens (Brown *et al.*, 2002; Vanparys *et al.*, 2008; Kandori *et al.*, 2009; Takakura *et al.*, 2009). However, these studies suggest that the negative effects of competition for pollinator service tend to be asymmetrical, that is alien species tend to have more negative effects on the reproductive success of their native congeners than native species on their aliens. Indeed, a recent meta-analysis showed that the negative effect of aliens on native species reproduction is consistently high and that this effect decreases with phylogenetic relatedness between alien and native species (Morales & Traveset, 2009). It generally seems that pollination mechanisms will tend to facilitate the naturalization of introduced species when species related to the invader occur within native communities, which goes against the prediction of Darwin's naturalization hypothesis. However, a proper test of this would be to examine whether pollinator limitation on reproductive success of alien species is increased when native communities are more related to the alien species. To our knowledge, such a test has not been published so far.

To conclude, naturalization of introduced species may result from varied ecological mechanisms of biotic resistance that are not necessarily affected by phylogeny. Although more empirical data are clearly needed especially on animal taxa, it appears that plant competitive hierarchies are poorly affected by species phylogenetic relationships. However, trophic interactions are strongly phylogenetically structured, and there are compelling evidences that biotic resistance through natural enemies may be affected by phylogenetic relatedness between invaders and native communities. Thus, there is a risk of observing no significant pattern of phylogenetic relatedness and invasion success of alien species, while varied mechanisms of biotic resistance might be truly acting, but in opposite directions.

### **Towards a trait-based measurement of species niche similarity**

As outlined earlier, phylogenetic relatedness will not always be a good predictor for the ecological mechanisms that favour biotic resistance of native communities to particular alien species. A very promising alternative would be to quantify species niche similarity through a set of functional traits (Violle *et al.*, 2007) depicting the strategy of species in terms of resource use and interspecific interactions. Such an approach ties with the original proposition of Elton (1958) that successful invaders should harbour original traits relative to

native communities. Although highly promising, this approach will also have its own problems, which may be potentially overcome. First, the same scale issues than the ones highlighted here will still apply with a trait-based approach (e.g. Fig 3). More importantly, the choice of a set of relevant traits, that adequately depicts different niche characteristics and mechanisms of biotic resistance and are measurable on large number of species ('soft traits') will certainly be critical (Violle *et al.*, 2007).

In a recent extensive review, Pyšek & Richardson (2007) analyzed multiple papers on plant invasions to determine whether there were any consistent sets of traits that could explain and predict a potential invader. However, these traits do not necessarily match with previously listed traits that would be useful to quantify species niches' overlaps (Violle *et al.*, 2007). Although the past literature had not reached a consensus so far, Pyšek & Richardson (2007) concluded that successful invaders possess some traits that unsuccessful invaders do not have. They also pleaded for explicitly filtering out the effects of residence time and other biases to reveal inherent trait-related determinants of invasibility. They suggested that the role of traits in the invasion process is to a very large extent invasion stage- and habitat-specific. Traits that confer an advantage at a given stage of the invasion process (naturalization, spread) and in a particular habitat may be neutral or even detrimental in another phase and/or a different habitat (Pyšek & Richardson, 2007). Quite surprisingly, few papers have built on this review to also investigate whether a successful invader exhibit different traits than the native communities, which allows the use of different resources and avoid competition ('niche gap-filling'), or exhibit more or less the same traits and directly compete for resources (Cahill *et al.*, 2008).

### **CONCLUSION AND TAKE HOME MESSAGE**

Biological invasions offer a unique opportunity to study the ecological and evolutionary mechanisms that drive species range expansion, species interactions and species coexistence within communities (Tilman, 2004; Callaway & Maron, 2006; Sax *et al.*, 2007). Ecologists have long separately studied species invasiveness and community invasibility. However, it is now increasingly recognized that both invader characteristics and community properties must be jointly accounted for if we ultimately want to predict invasion risks in the future (Richardson & Pyšek, 2006).

The long lasting quest for understanding community invasibility and resolving Darwin's invasion conundrum might be resolved if appropriate scales, metrics and statistical tests are thoroughly applied. Several considerations must be rigorously taken into account:

- Appropriate spatial scale: large spatial and grain scales are not appropriate to test co-existence mechanisms as drivers of community invasibility. An approach combining different scales is ultimately the best practice to disentangle all possible mechanisms driving species naturalization and spread (Fig. 2)

- Appropriate phylogenetic scale: the use of family and genera or more generally taxonomy-based relatedness metrics is not meaningful and tainted of strong unrealistic assumptions. Molecular phylogenies or supertree-based phylogenies should be favoured.
  - Appropriate metrics: the alpha niche concept offers a framework to link invader characteristics and community properties in order to test Darwin's naturalization hypotheses. We strongly suggest the use of four different metrics connected to this concept (Table 2).
  - Appropriate null models: past studies have not always used consistent statistical tests and null model hypotheses. Null models and the related species pool used to randomize should be carefully chosen based on the ecological mechanisms meant to be tested.
  - Appropriate measure of (dis)similarity: although phylogenetic relationships offer a solid background to test patterns of community invasibility, they are based on strong assumptions. We further suggest that, when available, functional traits related to resource acquisition and biotic interactions should also be considered in the quantification of the dissimilarity between invaders and native communities. Ultimately, measurement of functional relatedness via the alpha niche concept should be more heuristic than measurements of phylogenetic relatedness.
- The proper application of these above-mentioned criteria should foster the understanding of community invasibility and should generate comparable results from empirical tests of Darwin's naturalization hypothesis.

## ACKNOWLEDGEMENTS

We thank David Richardson for the invitation to write this opinion paper and for his insightful comments and editions on the manuscript. We also thank two anonymous reviewers who helped to significantly improve the paper and Pedro Jordano and Jordi Bascompte for providing literature references. This work was funded by ANR SCION (ANR-08-PEXT-03) and DIVERSITALP (ANR-07-BDIV-014) projects. We also received support from European Commission's FP6 ECOCHANGE (challenges in assessing and forecasting biodiversity and ecosystem changes in Europe, No. 066866 GOCE) project.

## REFERENCES

- Ackerly, D. (2009) Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19699–19706.
- Ackerly, D.D. & Cornwell, W.K. (2007) A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecology Letters*, **10**, 135–145.
- Albert, C.H., Thuiller, W., Yoccoz, N.G., Soudan, A., Boucher, F., Saccone, P. & Lavorel, S. (2010) Intraspecific functional

- variability: extent, structure and sources of variation within an alpine valley. *Journal of Ecology* (in press).
- Bininda-Emonds, O.R.P., Gittleman, J.L. & Steel, M.A. (2002) The (super)tree of Life: procedures, problems, and prospects. *Annual Review of Ecology and Systematics*, **33**, 265–289.
- 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.
- Braby, C.E. & Somero, G.N. (2006) Following the heart: temperature and salinity effects on heart rate in native and invasive species of blue mussels (genus *Mytilus*). *Journal of Experimental Biology*, **209**, 2554–2566.
- Brändle, M., Kühn, I., Klotz, S., Belle, C. & Brandl, R. (2008) Species richness of herbivores on exotic host plants increases with time since introduction of the host. *Diversity and Distributions*, **14**, 905–912.
- Brock, M.T. & Galen, C. (2005) Drought tolerance in the alpine dandelion, *Taraxacum ceratophorum* (Asteraceae), its exotic congener *T. officinale*, and interspecific hybrids under natural and experimental conditions. *American Journal of Botany*, **92**, 1311–1321.
- Brown, W.P. & Zuefle, M.E. (2009) Does the periodical cicada, *Magicicada septendecim*, prefer to oviposit on native or exotic plant species? *Ecological Entomology*, **34**, 346–355.
- Brown, B.J., Mitchell, R.J. & Graham, S.A. (2002) Competition for pollination between an invasive species (purple loosestrife) and a native congener. *Ecology*, **83**, 2328–2336.
- Bruno, J.F., Stachowicz, J.J. & Bertness, M.D. (2003) Inclusion of facilitation into ecological theory. *Trends in Ecology and Evolution*, **18**, 119–125.
- Burns, J.H. (2004) A comparison of invasive and non-invasive dayflowers (Commelinaceae) across experimental nutrient and water gradients. *Diversity and Distributions*, **10**, 387–397.
- Burton Hill, S. & Kotanen, P.M. (2009) Evidence that phylogenetically novel non-indigenous plants experience less herbivory. *Oecologia*, **161**, 581–590.
- Cahill, J.F., Kembel, S.W., Lamb, E.G. & Keddy, P.A. (2008) Does phylogenetic relatedness influence the strength of competition among vascular plants? *Perspectives in Plant Ecology, Evolution and Systematics*, **10**, 41–50.
- Callaway, R. & Maron, J.L. (2006) What have exotic plant invasions taught us over the past 20 years? *Trends in Ecology & Evolution*, **21**, 369–374.
- Cattin, M.-F., Bersier, L.-F., Banasek-Richter, C., Balten-sperger, R. & Gabriel, J.-P. (2004) Phylogenetic constraints and adaptation explain food-web structure. *Nature*, **427**, 835–839.
- Cavender-Bares, J., Keen, A. & Miles, B. (2006) Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. *Ecology*, **87**, S109–S122.
- Chase, J.M. & Leibold, M.A. (2003) *Ecological niches*. Chicago University Press, Chicago.
- Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, **31**, 343–366.

- Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., Mcglone, M.S., Weston, P.H., Westoby, M., Wilf, P. & Linder, H.P. (2009) Phylogenetic biome conservatism on a global scale. *Nature*, **458**, 754–758.
- Daehler, C.C. (2001) Darwin's naturalization hypothesis revisited. *American Naturalist*, **158**, 324–330.
- Darwin, C.R. (1859) *The origin of species*. John Murray, London.
- Davis, M.A., Grime, J.P. & Thompson, K. (2000) Fluctuating resources in plant communities: a general theory of invasibility. *Journal of Ecology*, **88**, 528–534.
- Dawson, W., Burslem, D.F.R.P. & Hulme, P.E. (2009) Herbivory is related to taxonomic isolation, but not to invasiveness of tropical alien plants. *Diversity and Distributions*, **15**, 141–147.
- De Bello, F., Thuiller, W., Leps, J., Choler, P., Clément, J.-C., Macek, P., Sebastia, M.T. & Lavorel, S. (2009) Partitioning of functional diversity reveals the scale and extent of trait convergence and divergence. *Journal of Vegetation Science*, **20**, 475–486.
- De Candolle, A.I. (1855) *Géographie botanique raisonnée*. Masson, Paris.
- Deng, X., Ye, W.H., Feng, H.L., Yang, Q.H., Cao, H.L., Hui, K.Y. & Zhang, Y. (2004) Gas exchange characteristics of the invasive species *Mikania micrantha* and its indigenous congener *M. cordata* (Asteraceae) in South China. *Botanical Bulletin of Academia Sinica*, **45**, 213–220.
- Diez, J.M., Sullivan, J.J., Hulme, P.E., Edwards, G. & Duncan, R.P. (2008) Darwin's naturalization conundrum: dissecting taxonomic patterns of species invasions. *Ecology Letters*, **11**, 674–681.
- Diez, J.M., Williams, P.A., Randall, J.M., Sullivan, J.J., Hulme, P.E. & Duncan, R.P. (2009) Learning from failures: testing broad taxonomic hypotheses about plant naturalization. *Ecology Letters*, **12**, 1174–1183.
- Duncan, R.P. & Williams, P.A. (2002) Darwin's naturalization hypothesis challenged. *Nature*, **417**, 608–609.
- Elton, C.S. (1958) *The ecology of invasions by animals and plants*. Methuen, London.
- García-Serrano, H., Sans, F.X. & Escarre, J. (2007) Interspecific competition between alien and native congeneric species. *Acta Oecologica-International Journal of Ecology*, **31**, 69–78.
- Gause, G.F. (1934) *The struggle for existence*. Williams & Wilkins, Baltimore, MD.
- Geng, Y.P., Pan, X.Y., Xu, C.Y., Zhang, W.J., Li, B. & Chen, J.K. (2006) Phenotypic plasticity of invasive *Alternanthera philoxeroides* in relation to different water availability, compared to its native congener. *Acta Oecologica-International Journal of Ecology*, **30**, 380–385.
- 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.
- Hawkes, C.V. (2007) Are invaders moving targets? The generality and persistence of advantages in size, reproduction, and enemy release in invasive plant species with time since introduction. *American Naturalist*, **170**, 832–843.
- Hubbell, S.P. (2001) *A unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, NJ.
- Hughes, C. & Eastwood, R. (2006) Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 10334–10339.
- Ives, A.R. & Godfray, H.C.J. (2006) Phylogenetic analysis of trophic associations. *American Naturalist*, **168**, e1–e14.
- Kandori, I., Hirao, T., Matsunaga, S. & Kurosaki, T. (2009) An invasive dandelion unilaterally reduces the reproduction of a native congener through competition for pollination. *Oecologia*, **159**, 559–569.
- Keane, R.M. & Crawley, M.J. (2002) Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology and Evolution*, **17**, 164–170.
- Kemmel, S.W. & Hubbell, S.P. (2006) The phylogenetic structure of a neotropical forest tree community. *Ecology*, **87**, S86–S99.
- Klironomos, J.N. (2002) Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature*, **417**, 67–70.
- Lambdon, P.W. & Hulme, P.E. (2006) How strongly do interactions with closely-related native species influence plant invasions? Darwin's naturalization hypothesis assessed on Mediterranean islands. *Journal of Biogeography*, **33**, 1116–1125.
- Levine, J.M., Vilà, M., D'Antonio, C.M., Dukes, J.S., Grigulis, K. & Lavorel, S. (2003) Mechanisms underlying the impacts of exotic plant invasions. *Proceedings of the Royal Society of London series B*, **270**, 775–781.
- Levine, J.M., Adler, P.B. & Yelenik, S.G. (2004) A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters*, **7**, 975–989.
- 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.
- Losos, J.B. & Glor, R.E. (2003) Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution*, **18**, 220–227.
- Macdougall, A.S., Gilbert, B. & Levine, J.M. (2009) Plant invasions and the niche. *Journal of Ecology*, **97**, 609–615.
- Maherali, H. & Klironomos, J.N. (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science*, **316**, 1746–1748.
- McDowell, S.C. (2002) Photosynthetic characteristics of invasive and noninvasive species of *Rubus* (Rosaceae). *American Journal of Botany*, **89**, 1431–1438.
- Memmot, J. & Waser, N.M. (2002) Integration of alien plants into a native flower–pollinator visitation web. *Proceedings of the Royal Society of London series B*, **269**, 2395–2399.
- Milton, S.J., Wilson, J.R.U., Richardson, D.M., Seymour, C.L., Dean, W.R.J., Iponga, D.M. & Procheş, Ş. (2007) Invasive alien plants infiltrate bird-mediated shrub nucleation processes in arid savanna. *Journal of Ecology*, **95**, 648–661.
- Mitchell, C.E., Agrawal, A.A., Bever, J.D., Gilbert, G.S., Hufbauer, R.A., Klironomos, J.N., Maron, J.L., Morris,

- W.F., Parker, I.M., Power, A.G., Seabloom, E.W., Torchin, M.E. & Vazquez, D.P. (2006) Biotic interactions and plant invasions. *Ecology Letters*, **9**, 726–740.
- Mooney, H.A. & Hobbs, R.J. (eds) (2000) *Invasive species in a changing world*. Island Press, Washington, D.C., USA.
- Morales, C.L. & Traveset, A. (2009) A meta-analysis of impacts of alien vs. native plants on pollinator visitation and reproductive success of co-flowering native plants. *Ecology Letters*, **12**, 716–728.
- Nentwig, W. (2007) *Biological invasions*. Springer-Verlag, Berlin, Heidelberg.
- Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2007) Niche dynamics in space and time. *Trends in Ecology & Evolution*, **23**, 149–158.
- Pearse, I.S. & Hipp, A.L. (2009) Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks. *Proceedings of National Academy of Sciences of USA*, **106**, 18097–18102.
- Peterson, A.T., Soberon, J. & Sanchez-Cordero, V. (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265–1267.
- Pickett, S.T.A. & Bazzaz, F.A. (1978) Organization of an assemblage of early successional species on a soil moisture gradient. *Ecology*, **59**, 1248–1255.
- Priddis, E., Rader, R., Belk, M., Schaalje, B. & Merkle, S. (2009) Can separation along the temperature niche axis promote coexistence between native and invasive species? *Diversity and Distributions*, **15**, 682–691.
- Prinzing, A., Durka, W., Klotz, S. & Brandl, R. (2001) The niche of higher plants: evidence for phylogenetic conservatism. *Proceedings of the Royal Society of London series B*, **268**, 2383–2389.
- Procheş, S., Wilson, J.R.U., Richardson, D.M. & Rejmanek, M. (2008) Searching for phylogenetic pattern in biological invasions. *Global Ecology and Biogeography*, **17**, 5–10.
- Pyšek, P. & Richardson, D.M. (2007) Traits associated with invasiveness in alien plants: where do we stand? *Biological invasions* (ed. by W. Nentwig), pp. 97–122, Springer-Verlag, Berlin Heidelberg.
- Rejmánek, M. (1995) What makes a species invasive? *Plant invasions: general aspects and special problems* (ed. by P. Pyšek, K. Prach, M. Rejmanek and M. Wade), pp. 3–13, SPB Academic Publishing, Amsterdam, the Netherlands.
- Rejmánek, M. (1996) A theory of seed plant invasiveness: the first sketch. *Biological Conservation*, **78**, 171–181.
- Rezende, E.L., Lavabre, J.E., Guimarães, P.R., Jordano, P. & Bascompte, J. (2007) Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*, **448**, 925–929.
- Rezende, E.L., Albert, E.M., Fortuna, M.A. & Bascompte, J. (2009) Compartments in a marine food web associated with phylogeny, body mass, and habitat structure. *Ecology Letters*, **12**, 779–788.
- Ricciardi, A. & Mottiar, M. (2006) Does Darwin's naturalization hypothesis explain fish invasions? *Biological Invasions*, **8**, 1403–1407.
- Ricciardi, A. & Atkinson, S.K. (2004) Distinctiveness magnifies the impact of biological invaders in aquatic ecosystems. *Ecology Letters*, **7**, 781–784.
- Richardson, D.M. & Pyšek, P. (2006) Plant invasions: merging the concepts of species invasiveness and community invasibility. *Progress in Physical Geography*, **30**, 409–431.
- Richardson, D.M., Rouget, M., Ralston, S.J., Cowling, R.M., Van Rensburg, B.J. & Thuiller, W. (2005) Species richness of alien plants in South Africa: environmental correlates and the relationship with native plant species richness. *Ecoscience*, **12**, 391–402.
- Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S., Grosberg, R.K., Hastings, A., Holt, R.D., Mayfield, M.M., O'connor, M.I. & Rice, W.R. (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution*, **22**, 465–471.
- Schierenbeck, K.A. & Marshall, J.D. (1993) Seasonal and diurnal patterns of photosynthetic gas-exchange for *Lonicera sempervirens* and *L. japonica* (Caprifoliaceae). *American Journal of Botany*, **80**, 1292–1299.
- Simberloff, D. & Von Holle, B. (1999) Positive interactions of nonindigenous species: invasional meltdown? *Biological Invasions*, **1**, 21–32.
- Smith, S.A., Beaulieu, J.M. & Donoghue, M.J. (2009) Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. *BMC Evolutionary Biology*, **9**, 37–48.
- Strauss, S.Y., Webb, C.O. & Salamin, N. (2006) Exotic taxa less related to native species are more invasive. *Proceedings of National Academy of Sciences of USA*, **103**, 5841–5845.
- Stubbs, W.J. & Wilson, J.B. (2004) Evidence for limiting similarity in a sand dune community. *Journal of Ecology*, **92**, 557–567.
- Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. & Zimmerman, J.K. (2006) The problem and promise of scale dependency in community phylogenetics. *Ecology*, **87**, 2418–2424.
- Takakura, K., Nishida, T., Matsumoto, T. & Nishida, S. (2009) Alien dandelion reduces the seed-set of a native congener through frequency-dependent and one-sided effects. *Biological Invasions*, **11**, 973–981.
- Thuiller, W. (2007) Biodiversity – climate change and the ecologist. *Nature*, **448**, 550–552.
- Thuiller, W., Richardson, D.M., Rouget, M., Procheş, Ş. & Wilson, J.R.U. (2006) Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology*, **87**, 1755–1769.
- Tilman, D. (2004) Niche tradeoffs, neutrality and community structure: a stochastic theory of resource competition, invasion and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, **30**, 10854–10861.
- Vacher, C., Piou, D. & Desprez-Loustau, M.-L. (2008) Architecture of an antagonistic tree/fungus network: the asymmetric influence of past evolutionary history. *PLoS ONE*, **3**, e1740.

- Vanparys, V., Meerts, P. & Jacquemart, A.L. (2008) Plant–pollinator interactions: comparison between an invasive and a native congeneric species. *Acta Oecologica-International Journal of Ecology*, **34**, 361–369.
- Vilà, M., Bartomeus, I., Dietzsch, A.C., Petanidou, T., Steffan-Dewenter, I., Stout, J.C. & Tscheulin, T. (2009) Invasive plant integration into native plant–pollinator networks across Europe. *Proceedings of the Royal Society of London series B*, **276**, 3887–3893.
- Violle, C., Navas, M.-L., Vile, D., Roumet, C., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of plant trait be functional! *Oikos*, **116**, 882–892.
- Walther, G.-R., Roques, A., Hulme, P.E. *et al.* (2009) Alien species in a warmer world – risks and opportunities. *Trends in Ecology & Evolution*, **24**, 686–693.
- Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution and Systematics*, **36**, 519–539.
- Willson, C.J., Manos, P.S. & Jackson, R.B. (2008) Hydraulic traits are influenced by phylogenetic history in the drought-resistant, invasive genus *Juniperus* (Cupressaceae). *American Journal of Botany*, **95**, 299–314.
- Zuefle, M.E., Brown, W.P. & Tallamy, D.W. (2008) Effects of non-native plants on the native insect community of Delaware. *Biological Invasions*, **10**, 1159–1169.

## BIOSKETCH

The EMABIO's team (Evolution, Modelling and Analysis of BIOdiversity) led by **Wilfried Thuiller** focuses at revealing the evolutionary mechanisms that have generated extant biodiversity and its spatial patterns, especially climatic niche evolution and its effect on species diversification. The EMABIO's team also aims to address the mechanisms governing the assembly of biotic communities, such as the genetic effects of keystone species on plant and microbial communities, but also the effects of dispersal limitation, environmental filtering, and resource competition. Finally, the EMABIO's team proposes new modelling tools and generate biodiversity assessments at various spatial scales (continental, national and regional).

Author contributions: All authors conceived the ideas and the structure of the paper. W.T. and S.L. led the writing.

---

Editor: David Richardson



## *Chapter 3.2*





# Darwin's naturalization hypothesis: scale matters in coastal plant communities

Marta Carboni, Tamara Münkemüller, Laure Gallien, Sébastien Lavergne, Alicia Acosta and Wilfried Thuiller

*M. Carboni (mcarboni@uniroma3.it), Dipartimento di Biologia Ambientale, Univ. degli Studi Roma Tre, V.le Marconi 446, IT-00146 Roma, Italy, and Laboratoire d'Ecologie Alpine, UMR 5553 CNRS – Univ. Joseph Fourier, FR-38041 Grenoble, France. – T. Münkemüller, L. Gallien, S. Lavergne and W. Thuiller, Laboratoire d'Ecologie Alpine, UMR 5553 CNRS – Univ. Joseph Fourier, FR-38041 Grenoble, France. – A. Acosta, Dipartimento di Biologia Ambientale, Univ. degli Studi Roma Tre, V.le Marconi 446, IT-00146 Roma, Italy.*

Darwin proposed two seemingly contradictory hypotheses for a better understanding of biological invasions. Strong relatedness of invaders to native communities as an indication of niche overlap could promote naturalization because of appropriate niche adaptation, but could also hamper naturalization because of negative interactions with native species ('Darwin's naturalization hypothesis'). Although these hypotheses provide clear and opposing predictions for expected patterns of species relatedness in invaded communities, so far no study has been able to clearly disentangle the underlying mechanisms. We hypothesize that conflicting past results are mainly due to the neglected role of spatial resolution of the community sampling. In this study, we corroborate both of Darwin's expectations by using phylogenetic relatedness as a measure of niche overlap and by testing the effects of sampling resolution in highly invaded coastal plant communities. At spatial resolutions fine enough to detect signatures of biotic interactions, we find that most invaders are less related to their nearest relative in invaded plant communities than expected by chance (phylogenetic overdispersion). Yet at coarser spatial resolutions, native assemblages become more invulnerable for closely-related species as a consequence of habitat filtering (phylogenetic clustering). Recognition of the importance of the spatial resolution at which communities are studied allows apparently contrasting theoretical and empirical results to be reconciled. Our study opens new perspectives on how to better detect, differentiate and understand the impact of negative biotic interactions and habitat filtering on the ability of invaders to establish in native communities.

Species transported far from their original range that spread and maintain viable populations (i.e. naturalized non-native species sensu Richardson and Pysek 2006) often pose significant challenges to conserving native biodiversity. Predicting which species can invade which communities is essential if control measures are to be successfully implemented (Marco et al. 2010). The composition of local native assemblages and the phylogenetic relatedness of an invader to these communities can influence invasion success and thus provide a predictive tool. Closely related species are more likely to be ecologically similar, provided that traits determining responses of species to environment and co-existence show a signal along the phylogeny (sensu Blomberg and Garland 2002; i.e. similar trait values between closely-related species). Under these conditions, species' phylogenetic distances can be used as a proxy for ecological similarity and have the advantage of combining multiple functional trait information.

There are two opposing hypotheses originally proposed by Darwin to link the phylogenetic relatedness between potential invaders and native communities with probabilities

of successful invasion (Darwin 1859). On the one hand, close relatedness is predicted to hamper local naturalization due to niche overlap and competition with native species (i.e. ecologically similar species compete more than dissimilar species; 'Darwin's naturalization hypothesis') (Elton 1958). The resulting pattern is commonly referred to as phylogenetic overdispersion in studies of community assembly rules. On the other hand, appropriate niche adaptation may instead favor the naturalization of closely-related introduced species due to habitat filtering, which leads to a spatial pattern of phylogenetic clustering or underdispersion of niches (Duncan and Williams 2002). Previous studies have found support for both of these hypotheses, leading to a fierce controversy in recent literature (Daehler 2001, Lambdon and Hulme 2006, Diez et al. 2009, Ricotta et al. 2010, Schaefer et al. 2011). Two aspects are likely to play a major role in explaining the discrepancy between theoretical predictions and among empirical studies: methodological differences and most importantly differences in the scale considered (Thuiller et al. 2010).

A standard methodological framework to address Darwin's naturalization hypothesis was outlined in a recent review paper (Thuiller et al. 2010). The authors suggest that the niche overlap between the invader and the members of the recipient community can be explored through a series of metrics based on functional or phylogenetic distances among species (when traits show a phylogenetic signal) and then tested with an appropriate null hypothesis and associated algorithm (Hardy 2008, Thuiller et al. 2010). As different metrics and methods have specific assumptions and may lead to different conclusions, adopting a combination of approaches both for the quantification of niche overlap and for the statistical test is a good method of corroborating results.

Ecological patterns are inherently scale-dependent and the resolution at which communities are sampled may have a major impact on the conclusions that can be drawn from data (Huston 1999, Willis and Whittaker 2002, Hanan and Ross 2010, Qian and Kissling 2010, Rocchini et al. 2010). Theoretically, we anticipate more niche dissimilarity among species (overdispersion) at finer resolutions where biotic interactions take place because of the effect of interspecific competition/facilitation or shared natural enemies. On the other hand, we cannot conclusively predict direct biotic interaction between co-occurring species at a coarser resolution, as species can segregate along environmental gradients encompassed within large sampling units. Therefore we rather anticipate greater similarity (clustering/underdispersion) among species because of shared resource requirements in this case. In fact, community assembly studies have shown that both phylogenetic clustering and overdispersion of native species within communities can appear within the same system, but at different spatial resolutions (Cavender-Bares et al. 2006, Swenson et al. 2006). In a similar way, spatial resolution may be the key to reconciling apparently contrasting hypotheses and empirical results in the field of invasion ecology (Stohlgren et al. 1997, 2002, Catford and Downes 2010, Jones et al. 2010).

Although the issue of spatial scale is clearly important in this context and has been addressed theoretically (Proches et al. 2008, Thuiller et al. 2010), there are few studies that have investigated the effect of scale on invaders' relatedness patterns using nested resolutions in the field (Cadotte et al. 2009, Davies et al. 2011, Schaefer et al. 2011). While promising evidence comes from an analysis conducted by Diez et al. (2008), which however did not include progressively finer sampling resolutions, to our knowledge no empirical study has been able to demonstrate the theoretically predicted spatial turning point from phylogenetic clustering to phylogenetic overdispersion of invaders. For example, Davies et al. (2011) found that native and non-native species were more distantly related than expected by chance not only at a fine resolution (plot scale), but also at a coarse one (hectare scale). Conversely, Cadotte et al. (2009) demonstrated phylogenetic clustering of invader success at the continental scale, but only found a random pattern at the smallest scale of analysis they considered (landscape). No study has yet included a comprehensive enough set of sampling resolutions to empirically reconcile the expected divergent patterns across sampling

resolutions. Consequently, there has been a call for cross-scale field-based approaches to tackle the issue of spatial scale in the context of phylogenetic patterns of biological invasions (Proches et al. 2008, Thuiller et al. 2010).

In this paper we explore the effect of sampling resolution on patterns of plant invasions with field data using a comprehensive set of metrics and statistical tests (Fig. 1). Focusing on Mediterranean coastal sand dunes, we test patterns of naturalized non-native species ('invaders' hereafter) in local communities at three spatial resolutions. We built a phylogenetic supertree to derive two complementary community scale measurements of the phylogenetic distance of invaders: the Mean Distance of the invader relative to Native Species (MDNS) and the Distance of the invader to its Nearest Native Species in the native community (DNNS). Finally, we assess the results by testing the hypothesis that relatedness of invaders is different in invaded communities from what it would be in non-invaded communities. To do so we rely both on randomization tests using an algorithm simulating 'random invasions' and on phylogenetic mixed effects models in a Bayesian framework. Specifically, we address the following crucial questions: 1) how phylogenetically distant should species be to successfully invade a native community? 2) Are the observed patterns different from random expectations? 3) Do the observed patterns change with increasing sampling resolution?

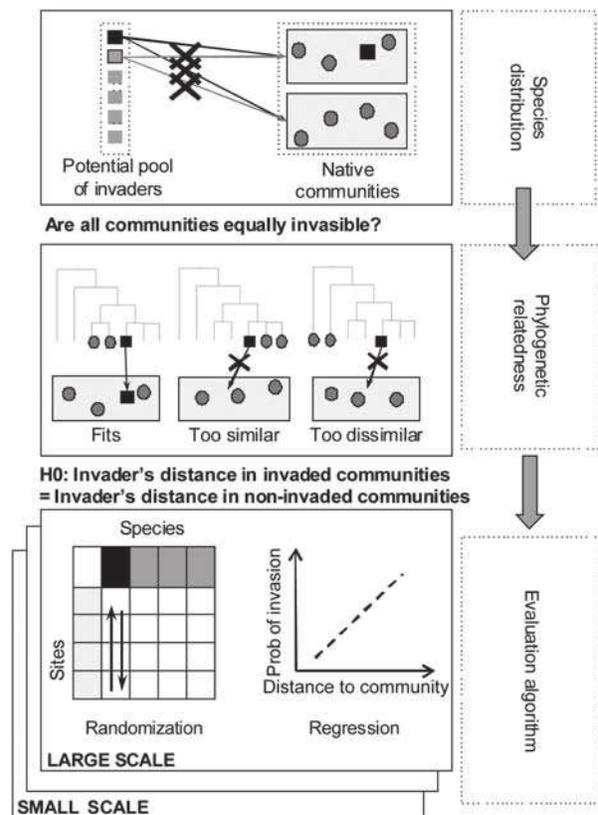


Figure 1. Conceptual diagram depicting the hypotheses relating naturalizations/invasions and phylogenetic relatedness/distance to the community and the testing procedure adopted in this paper. Non-native species are represented as squares and successful invaders in at least one community in the study area are in black.

## Methods

The system under study is located in Mediterranean coastal sand dunes known to be prone to invasions (Chytrý et al. 2009). The vegetation of sandy shores in central Italy has been extensively sampled in the past few years at nested sampling resolutions and (contrary to what is often the case with phytosociological surveys) with no bias towards native species (Acosta et al. 2008, 2009). The conditions were therefore ideal for testing phylogenetic structure in relation to invasion patterns at different sampling resolutions with a relevant number of invaders to generalize the findings.

### Study area and sampling resolutions

We specifically focused on recent coastal dunes (Holocene) extending over 400 km on the central Italian coasts: 250 km along the Tyrrhenian Sea (Lazio Region, from 42°23'N, 11°39'E to 41°11'N, 13°20'E) and 150 km along the Adriatic Sea (Molise and Abruzzi Regions). Holocene dunes represent ca 80% of the overall extent examined. Here we examined three different sampling resolutions: one very coarse and two progressively finer resolutions. For the coarse resolution, we relied on a survey of the vascular flora of central Italian coastal dunes carried out from 2004 to 2007 in 3' by 5' grid-cells (about 35 km<sup>2</sup>) which was limited to the geologic class of Holocenic dunes (Acosta et al. 2008, Carboni et al. 2010). Within each grid cell all vascular plant species (natives and introduced) were recorded wherever they occurred on recent dunes. 91 grid cells fall within the limits of the studied regions, however only 71 contained holocenic dunes. For the finer resolutions we used presence-absence data from a long-term (2002–2009) random sampling campaign of coastal dune vegetation in several study sites, comprising most of the best conserved remnant dune systems of the region (about 80 km on the western coast and 22 km on the eastern coast) (for more details see Acosta et al. 2009, Carboni et al. 2011). In this work we specifically consider two nested plot dimensions: 2 × 2 m (4 m<sup>2</sup>) as the finest resolution and an expansion of the plots to 8 × 8 m (64 m<sup>2</sup>) as the intermediate resolution of analysis. We considered 690 plots for each of these two dimensions. In these environments the finest plot size is compatible with the identification of homogenous plant communities, some habitat heterogeneity already occurs within the intermediate sized plots, and many different habitats occur within the grid cells at the coarsest resolution (sorted along a strong sea-inland environmental gradient; Carboni et al. 2011). Our sampling design is therefore arranged as to include decreasing environmental heterogeneity within sampling units from the coarse to the intermediate resolution, while we assume that environmental conditions are relatively homogenous within the fine resolution plots.

We classified only species introduced after the 15th century as non-natives in this study, but we included both invasive and naturalized species (Pyšek et al. 2004) according to the classification by Celesti-Grapow et al. (2009). From this list, we additionally excluded all species

that were clearly not naturalized on coastal dunes based on expert knowledge. See Supplementary material Appendix 1 for a list of all the non-native species (invasive and naturalized) sampled at each resolution. We refer to these as 'invaders' in a general sense, although we focus on all naturalized species, not only on those with high spread potential or with documented negative effects. Irrespective of the term used, we only make inferences on the naturalization process, not on the level of spread and further impact of introduced species.

### Supertree construction

We created a supertree of all the taxa in the communities sampled by combining a backbone tree based on the APG III phylogeny (<www.mobot.org/MOBOT/research/APweb/>), which was generated by Phylomatic (<www.phylodiversity.net/phyloomatic/phyloomatic.html>), along with subtrees that were created using other literature sources to include e.g. gymnosperms and ferns (Chaw et al. 1997, Frohlich and Chase 2007). We assigned branch lengths to the phylogenetic tree using the branch length adjustment algorithm (BLADJ) in Phylocom (Webb et al. 2008), based on the minimum age of nodes estimated from the fossil record (Wikstrom et al. 2001). To produce phylogenetic distance matrices and calculate distance-based metrics, we used the sum of branch lengths separating pairs of species. In the absence of more precise species phylogenies obtained by sequencing proper DNA regions these matrices provide a useful measurement expressing the phylogenetic relatedness for community analyses and have proved to be effective in other studies (Pillar and Duarte 2010, Ricotta et al. 2010, Davies et al. 2011). As our working phylogeny still contained many polytomies, we ran sensitivity analyses to test whether our limited tree resolution was substantially distorting our results (see Results section and Supplementary material Appendix 3).

We also checked for a phylogenetic signal (sensu Abouheif 1999, Blomberg and Garland 2002) in a set of functional traits known to relate to species strategies or resource use (Westoby 1998, Supplementary material Appendix 4). Tests of phylogenetic signal showed that most traits examined were more similar for closely-related species than under random expectations, corroborating our assumption that closely-related species shared more similar ecological characteristics than two species taken at random in the phylogeny. All details on traits, methods and results for these preliminary tests can be found in Supplementary material Appendix 4.

### Spatial structure of invaders

In order to better calibrate the subsequent randomization algorithms and regression analyses we performed a series of preliminary tests to check whether invaders were spatially clustered as a subgroup. To verify how invaders were spatially arranged at the three scales we measured co-occurrence patterns (species.dist function in the R package picante, R Development Core Team, Kembel et al. 2010).

Table 1. Spatial co-occurrence patterns of invaders and natives.  $C_{ij}$  co-occurrence Schoener's index at the three resolutions (by rows) compared with randomized indices to address (by columns): co-occurrence patterns assessed separately for invaders (1) and for natives (2), and (3) for invaders in comparison to natives.

Hypothesis: Data: Randomization:	(1) Some invaders tend to co-occur invaders (shuffle w/in sp)			(2) Some natives tend to co-occur native species (shuffle w/in sp)			(3) Invaders co-occur more often than natives all species (shuffle invaders w/random sp)		
	obs	mean rand	p	obs	mean rand	p	obs	mean rand	p
	4 m <sup>2</sup>	0.016	0.014	0.279	0.022	0.014	< 0.001	0.716	1.019
64 m <sup>2</sup>	0.014	0.007	< 0.001	0.018	0.009	< 0.001	0.784	1.016	0.692
35 km <sup>2</sup>	0.09	0.040	< 0.001	0.09	0.043	< 0.001	1.043	0.998	0.349

We calculated pairwise values of co-occurrence using Schoener's index ( $C_{ij}$ ), which is based on proportional similarity (Schoener 1970):  $C_{ij} = 1 - 0.5 \times \sum |p_{ih} - p_{jh}|$ , where  $C_{ij}$  is the co-occurrence of species  $i$  and  $j$  and  $p$  is the proportion of occurrences of the  $i_{th}$  species in the  $h_{th}$  plot. With presence/absence data,  $p_{jh}$  is zero if the species is absent from site  $h$ , otherwise it is the inverse of the number of sites where species  $i$  occurs. Mean observed  $C_{ij}$  at the three scales was compared with randomized indices. Co-occurrence patterns were assessed for the two separate subgroups of invaders and natives by using a null-model which maintains the overall frequency of each species in the study region, i.e. shuffling sites within each species (Gotelli 2000). This allowed us to see if there is clustering of invaders as a subgroup (Table 1, column 1) and then compare with patterns of natives (Table 1, column 2). Finally, to verify if there are differences between the two, clustering of invaders is assessed with respect to natives by comparing the observed ratio of ' $C_{ij}$  values for invaders/ $C_{ij}$  values for all species' with randomized values obtained by randomly selecting invaders from among the species in the species pool (Table 1, column 3).

### Testing Darwin's naturalization hypothesis

We considered two complementary distance-based metrics to quantify invaders' relatedness to the community: the Mean Distance of the invader relative to Native Species (MDNS) and the Distance of the invader to its Nearest Native Species in the native community (DNNS) (Thuiller et al. 2010). It has been shown that phylogenetic distance-based metrics can be confounded with species richness (Pavoine and Bonsall 2011) and that there can be associations between species richness and the presence of invaders (Stohlgren et al. 2002, Stachowicz and Tilman 2005). In order to partial out the effect of species richness and focus only on the effect of phylogenetic relatedness, we used residuals of MDNS/DNNS regressed against total plot richness rather than the observed MDNS/DNNS values in the following analyses (Davies and Buckley 2011). These residuals are measures of phylogenetic distance independent of species richness (for simplicity just MDNS<sub>resid</sub> and DNNS<sub>resid</sub> hereafter). Examples of regression plots of MDNS and DNNS vs species richness are reported in Supplementary material Appendix 2.

To test whether patterns of invasion measured with MDNS<sub>resid</sub> and DNNS<sub>resid</sub> were different from random expectations, we adopted two complementary approaches.

First, we defined an ad-hoc randomization scheme (Fig. 1), testing whether there are differences between invaded and non-invaded communities. We simulated 'random invasions' by manipulating the invaders in the species by site matrix and permuting, independently for each invader, local presences/absences among sites. At each examined resolution we thus generated null distributions of MDNS<sub>resid</sub> and DNNS<sub>resid</sub> averaged across sites for each invader recorded, to which the observed values could be compared. We concluded that the test was significant if the actual values were greater than 97.5% (overdispersion) of the generated values or lower than 97.5% of the values (underdispersion), i.e. if the overall two-tailed p-value was < 0.05. In other words, for each scale by metric combination ( $3 \times 2 = 6$  combinations) there are two one-tailed tests at  $\alpha = 0.025$  distinguishing phylogenetic clustering and overdispersion of each invader. To assess overall significance of patterns for each spatial scale but across species, we performed a Fisher's test that combined the p-values for each hypothesis (function 'combine.test' in package 'survcomp'; Haibe-Kains et al. 2008).

Second, we implemented mixed effects models to assess whether the probability of community invasion was related to the phylogenetic distance between invasive species and native communities. We independently modeled the effect of MDNS<sub>resid</sub> and DNNS<sub>resid</sub> on the binary response variable 'invaded/non-invaded' assuming a binomial distribution of the response. To account for the fact that more than one alien species could invade one plot and to allow for different intercepts for each invader, we included two random factors in each model: plot identity and invader identity. Additionally we accounted for non-independence among individual invaders with a matrix of phylogenetic relatedness among species. We took a Bayesian approach using the R package MCMCglmm, which enables both random factors and a correlation structure depending on species phylogenetic relationships to be included (Hadfield 2010). Each model was run for 250 000 Markov chain Monte Carlo (MCMC) steps, with a burn-in of 50 000 iterations. Uninformative prior distributions were used for parameters, with mean of 0 and residual variance-covariance matrices set to 1. We checked for convergence in the parameter estimation by inspecting trace plots of the MCMC iterations. We chose a thinning interval of 200 iterations, which resulted in posteriori distributions with 1000 samples. From these posteriori distributions we calculated mean parameter estimates, and 95% Highest Posterior Density (HPD) or Credible Intervals (CI). Significance of model parameters was estimated by

examining CIs: parameters with CIs overlapping with zero were considered not to be significant.

## Results

### Phylogeny

We obtained a phylogenetic supertree for our study system comprising a total of 798 species, of which 51 species were invaders (Supplementary material Appendix 1). A supertree constructed with Phylomatic is typically not fully resolved, with many species as polytomies within genera and some genera as polytomies within families. To test the influence of the polytomies on  $MDNS_{resid}$  and  $DNNS_{resid}$ , we performed a sensitivity analysis on the basis of randomly resolved trees (using the ‘polytomy resolver’ phylogenetic tool, Supplementary material Appendix 3). We found that the metrics from the unresolved tree were unbiased estimates and that the uncertainty was consistently moderate (Supplementary material Appendix 3). Hence, in a pragmatic way we decided to perform all analyses with the unresolved trees to avoid working with hundreds of randomly resolved trees. Visual inspection of the tree showed that invaders tended to be grouped in several independent clusters with likely different evolutionary histories and ecological niches.

### Spatial structure of invader distribution

In order to fine-tune the randomization approach we analyzed species’ spatial co-occurrence patterns. When a community is invaded by more than one species it is not obvious how all invaders of the community should be

treated in the randomization tests. It is indeed unknown whether a native species has been excluded by an invader, in which case this invader would reflect characteristics of the lost native. The spatial differences in co-occurrence patterns of invaders with respect to natives can be informative in this sense. We found that invaders as a subgroup did not tend to co-occur at fine resolutions, though they appeared to be somewhat spatially clustered at coarser resolutions (Table 1, column 1). However, they were never more spatially clustered than the native species in the species pool (compare columns 1 and 2 in Table 1; column 3). Given these results on spatial occurrences, we adopted the more conservative approach whereby other invaders occurring in the community were not excluded in permutations for generating ‘random invasions’.

### Testing Darwin’s naturalization hypothesis

When considering only the most related native species ( $DNNS_{resid}$  – Fig. 2a), results from randomization tests corroborated Darwin’s naturalization hypothesis (phylogenetic overdispersion) at the finest sampling resolution ( $2 \times 2$  m). For this resolution, a high proportion of invaders (ca 25%; Fisher’s  $p$ -value = 0.002) were more distantly related to the closest relative in the invaded local community than under random expectations. However, none of the invaders showed a  $DNNS_{resid}$  greater than expected by chance at the intermediate and coarse sampling resolutions. The trend was even inverted at the coarsest resolutions, with ca 10–15% of the invaders having smaller  $DNNS_{resid}$  values than expected by chance (Fisher’s  $p$ -value < 0.001 at intermediate and 0.003 at coarse resolution). In other words, at coarser resolutions, the invaders tended to preferentially invade communities where at least one close relative

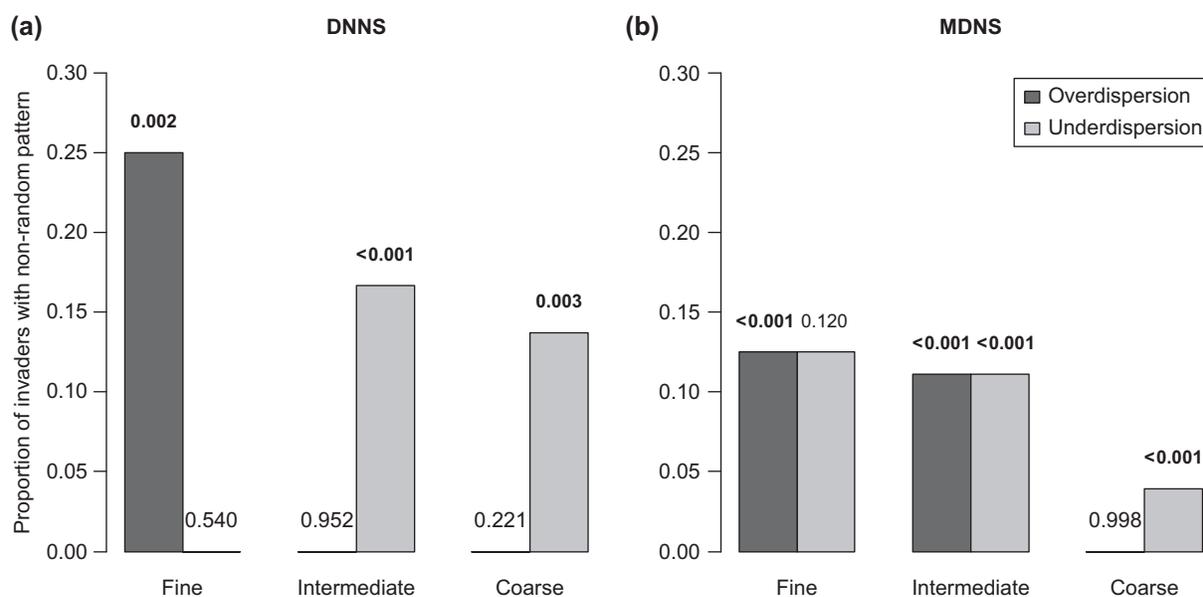


Figure 2. Proportion of invaders for which phylogenetic distance of the invader to the native community (after partialling out the effect of species richness) deviated from random patterns generated through randomizations. Significantly greater distances than expected by chance indicate overdispersion, whereas smaller distances indicate underdispersion. P-values on top of the bars are obtained by combining the p-values of randomization tests for each single invader through a Fisher’s test. Bold type indicates overall significant deviations from random expectations (i.e. across invaders). In each panel sampling resolutions from left to right are  $4 \text{ m}^2$ ,  $64 \text{ m}^2$  and ca  $35 \text{ km}^2$ .

already occurred (i.e. a species likely to share similar niche requirements).

Patterns were similar, but less clear when considering all the species in the community ( $MDNS_{resid}$ ). Fisher's test showed that invaders were significantly more distant from natives than expected by chance at the finest resolution ( $p < 0.001$ ) and closer than expected at the coarsest resolution ( $p < 0.001$ ). However at all scales examined, a relatively small proportion of invaders had  $MDNS_{resid}$  values that differed from random expectations, with approximately equal proportions of overdispersion and underdispersion at the fine and intermediate scales (Fig. 2b).

Mixed effect models supported the results obtained by comparing observed patterns with simulated 'random invasions' (Fig. 3a, b). Specifically, at the finest resolution we found a significant positive relation of the probability of invasion with the phylogenetic distance of the invader to the closest relative ( $DNNS_{resid}$ ). This relationship was inverted at the coarsest resolution, where the slope of the relationship was negative. The models with  $MDNS_{resid}$  as the explanatory variable showed the same trends at both fine and coarse resolutions, although the slope was not significantly different from zero in either case.

## Discussion

Recognition that community invasibility depends on the match between the characteristics of the invader and those of members of the recipient native community (Richardson and Pysek 2006) has been a major shift in the field of

invasion ecology. This has generated a growing interest in assessing the role that functional similarity and phylogenetic relationships play in biological invasions (Daehler 2001, Duncan and Williams 2002, Lambdon and Hulme 2006, Strauss et al. 2006, Winter et al. 2009). However, most studies have provided only partial or even diverging conclusions. Our study empirically investigated one of the main conceptual reasons put forward to explain conflicting results (Proches et al. 2008, Thuiller et al. 2010). Here we demonstrated using empirical data the crucial importance of spatial resolution for detecting phylogenetic patterns of invasion and we explored the implications of the choice of metrics and statistical tests.

When comparing  $DNNS_{resid}$  of potential invaders in invaded and non-invaded communities we found support for Darwin's naturalization hypothesis (phylogenetic overdispersion) at the fine sampling resolution at which competitive interactions take place on Mediterranean coastal dunes. In other words, invaders were more likely to be present in plots when they were more phylogenetically distant from their native relatives occurring in that plot (e.g. in the case of *C. acinaciformis*, one of the most invasive species in these environments. See Supplementary material Appendix 5 for a discussion on patterns of single invaders and their likely interactions with native species). This finding was reported at relatively small spatial scales in other contexts (Ricciardi and Atkinson 2004, Jiang et al. 2010). As we found that several traits tended to show a phylogenetic signal (Supplementary material Appendix 4), it is likely that relatedness of the invader indeed reflected high functional similarity and thus niche overlap with the

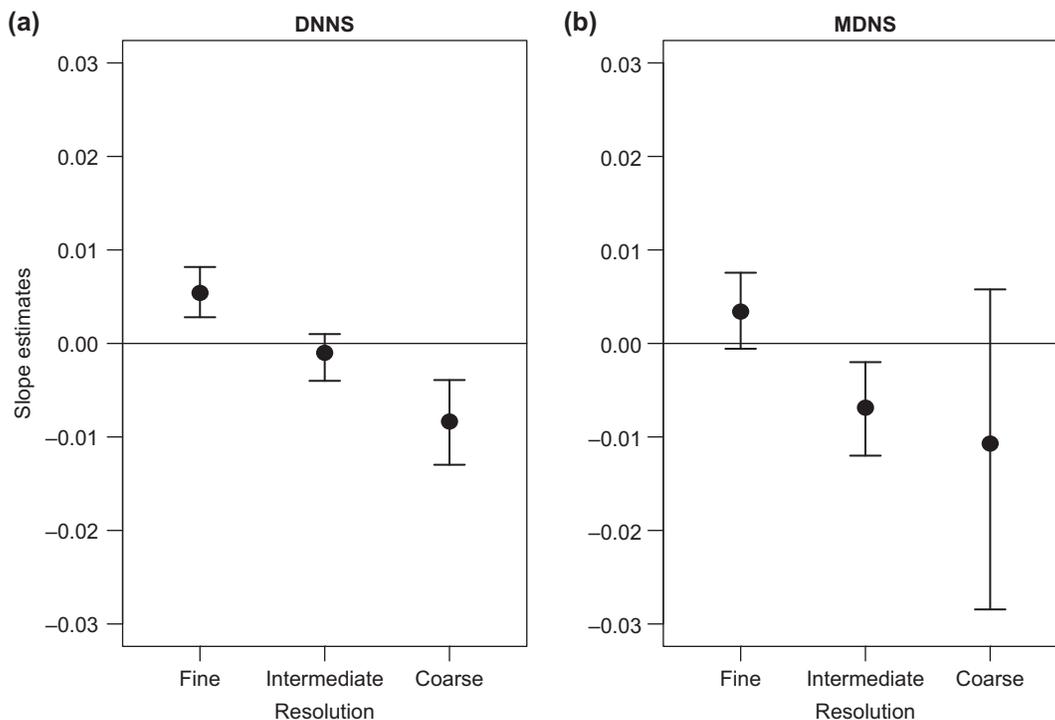


Figure 3. Results of MCMC mixed effects models for the probability of community invasion as a function of the phylogenetic distance of the invader to the community (after partialling out the effect of richness). Plots depict the mean posterior distributions for the slope parameter (with 95% credible intervals plotted as bars) for the effect of  $DNNS_{resid}$  and  $MDNS_{resid}$  at fine, intermediate and coarse spatial resolutions (i.e. 4 m<sup>2</sup>, 64 m<sup>2</sup> and ca 35 km<sup>2</sup>).

native species. Consequently, our results are in line with Darwin's and Elton's theoretical expectations on the biotic resistance of the native community to invasion. At coarser resolutions we instead found an opposite pattern suggesting a more dominant effect of habitat filtering (phylogenetic clustering). At the coarsest resolution (ca 35 km<sup>2</sup>) a high proportion of invaders was more related to the invaded communities than expected by chance. This trend mirrors patterns previously observed at regional and continental scales, for example in the floras of New Zealand and Australia (Duncan and Williams 2002, Diez et al. 2009). In fact, when considering a coarser resolution, species can co-occur while avoiding direct biotic interactions. The main reason is presumably that greater environmental variation is encompassed within larger sites, providing opportunities for species to sort across environmental gradients (Willis et al. 2010). It is therefore possible to reconcile apparently contrasting hypotheses and results for the patterns of relatedness of invading plants through the explicit consideration of the scale or resolution at which communities are sampled and defined. This is true in our system even though we restricted our analyses to coastal dunes, so that our coarse scale species pool is already quite filtered compared to earlier studies (Duncan and Williams 2002, Cadotte et al. 2009, Diez et al. 2009). However, given the strong sea-inland environmental gradient and strong zonation of the vegetation the species pool is still broad enough to detect the underdispersion in a high proportion of invaders.

Determining the spatial resolution at which the effects of biotic resistance are outbalanced by environmental filtering so that closely-related introduced species are no longer excluded has proven a difficult task. This is because competition is most plausible only at fairly fine spatial resolutions and among fairly related species (the 'Darwin-Hutchinson zone' according to Vamosi et al. 2009). The few studies explicitly searching for a turning point may have failed because they may have considered a range of scales inappropriate for the specific study system/taxon or because they may have missed the Darwin-Hutchinson zone (Vamosi et al. 2009). In contrast, our study picked up within a single area (central Italy) and ecosystem (coastal dunes) both phylogenetic overdispersion and phylogenetic clustering of invaders, effectively establishing the spatial turning point from one pattern to the other. Interestingly, in recent work on serpentine ecosystems, Davies et al. (2011) found that native and non-native species were more distantly related than expected by chance at a fine sampling resolution roughly comparable to the one used in this study. However, contrary to their expectations, they also found overdispersion at the rather coarse resolution of one hectare blocks. In contrast, we were able to detect the shift to phylogenetic clustering, presumably on two main grounds. First, our chosen system is characterized by marked habitat heterogeneity within relatively small extensions, as the strong sea-inland environmental gradient determines a compressed vegetation zonation across the dune profile. Second, we included a very broad range of sampling resolutions so that the grid cells used for our coarse scale analysis are much larger (35 km<sup>2</sup>) than the ones used by Davies et al. (2011). Intriguingly, Cadotte et al. (2009)

were able to show phylogenetic clustering of invader success as a result of habitat filtering only at an extremely large continental scale. Our results, forming a bridge between these two studies, seem to imply that the spatial turning point at which invaders become more similar to native species must be searched for at a resolution coarse enough to encompass a variety of different habitats (e.g. from annual beach communities to backdune Mediterranean macchia in our case). The exact plot size depends on the study system and on the amount of heterogeneity characteristic for the specific ecosystem examined, as well as on the taxa under consideration. Our study focusing on a single type of environment (coastal dunes) allows highlighting phylogenetic clustering at relatively finer scales than previous state wide or continental assessments (Strauss et al. 2006, Cadotte et al. 2009).

Quite surprisingly, in a recent study addressing 'Darwin's naturalization hypothesis' at different scales, Schaefer et al. (2011) found that introduced plant species were more likely to become invasive in the absence of close relatives in the overall native flora of the Azores, but could not confirm this trend with a fine sampling resolution. The authors argue that on these islands the exclusion of similar invaders seemed to be mostly driven by the clustering of common enemies, such as herbivores and pathogens rather than by competition. They argue that enemy release should act at all spatial scales in a system like the Azores and thus produce a signal of over-dispersion not only at small but also at large scales, explaining why their outcome partially contradicts theoretical expectations. A similar result is reported by van Wilgen and Richardson (2011) for reptiles in North America.

By using a combination of two metrics of community relatedness we were able to investigate whether biotic resistance to invaders was best predicted by a single closely-related species or by the community structure overall. Using the DNNS<sub>resid</sub> metric we could highlight clear patterns in accordance with theoretical expectations at all scales, whereas the MDNS<sub>resid</sub> measure of community relatedness gave less clear or non-significant results. In fact, within small homogenous plots we found a strong proportion of invaders, which tended to avoid the single most closely-related species (DNNS<sub>resid</sub> metric). It has previously been hypothesized that the biotic resistance of a given community would be mainly driven by the closest native species because one strong competitor is sufficient for competitive exclusion (Kraft et al. 2007). Furthermore, we found stronger support for the influence of habitat filtering at coarse resolutions when focusing specifically on the most related taxon (DNNS<sub>resid</sub>). This is in contrast with the common belief that total community relatedness tests should perform better with habitat filtering (Kraft et al. 2007). In large grained grid cells the non-limiting availability of resources and the included habitat heterogeneity may support a number of different environmental conditions and thus a high native diversity of phylogenetic lineages. The niche of the invader may need to be close to one of these suitable environmental conditions, but the mean distance to all species may be uninformative and obscure any pattern. Focusing only on closely-related taxa may instead provide the necessary precision to reveal the

availability of favorable environmental conditions for the invader within the heterogeneous site. In addition, the use of a not fully resolved phylogenetic tree (affecting 'phylogenetic scale') may also more strongly limit the analytical power of MDNS, given that randomly resolving the tree resulted in greater variation in MDNS than in DNNS (Thuiller et al. 2010; Supplementary material Appendix 3). The identified patterns were consistent among statistical tests (randomizations and regressions) and generally indicated that the closest relative was most informative for both the habitat preference of the invader and for the biotic resistance of the community. In contrast MDNS<sub>resid</sub> was a less effective proxy with few invaders that had significant patterns in randomization tests and large widely overlapping credible intervals for slope estimates in regressions. However, considering that these results may strongly depend on system-specific habitat heterogeneity or on the phylogenetic resolution available and that a collection of related species may have more impact than a single closely-related one (Strauss et al. 2006), it is generally advisable to check both metrics when analyzing invasion patterns.

As we have seen, non-random phylogenetic patterns of plant invasions may be detected either by a comparison with random expectations generated through various algorithms or through regression models. Using a combination of approaches enabled an assessment to be made as to whether our results were not dependent on methodological assumptions. Randomization techniques have been widely used in community ecology (Gotelli 2000) and phylogenetics (Cavender-Bares et al. 2006) as they are easy to implement and make relatively few a priori assumptions about expected relationships. Nevertheless, particular caution must be applied when formulating randomization schemes in order to avoid permutations that alter more patterns than the ones that are being specifically tested (Hardy 2008). We specifically chose a randomization procedure that breaks down the phylogenetic relationship between introduced and native species, but not the phylogenetic relationships between native species of recipient communities. However, restrictions in the randomization procedure reduce the statistical power. Alternatively, the regression approach has also often been adopted when studying phylogenetic patterns of invasions (Duncan and Williams 2002, Lambdon and Hulme 2006). While caution when formulating precise hypotheses and choosing adequate species pools should also be applied to regression models, these methods generally have higher testing power in comparison with randomizations. Moreover, in this paper we fit regression models including a random component that accommodates different intercepts for each invader. We can therefore evaluate patterns of more than one species in a single analysis, rather than only inferring trends from the proportion of invaders with significantly non-random patterns. Besides, we can take species phylogenetic relationships into account in the modeling process, thus compensating for statistical non-independence among invaders due to shared ancestry. In summary, although both approaches we employed, i.e. randomization and regression based, have shortcomings and specific assumptions, the convergence in the obtained results suggests that the patterns we highlight in this study are not dependent on methodological choices.

In conclusion, we found that the relationship between phylogenetic distance and probability of occurrence of an invader changes with spatial resolution and that we can confirm Darwin's naturalization hypothesis for fine resolutions where biotic interactions are also expected to be most important. Our results appear robust as both statistical tests applied supported the same conclusions. This paper therefore offers a new methodological framework for using the composition of local native species assemblages as a predictive tool for the new establishment of invaders. The specific resolution at which a community is no longer driven by biotic interactions, but rather by habitat filtering depends on habitat heterogeneity, and should therefore vary depending on the system being studied. In general, our results are promising for the perspective of incorporating information on the phylogenetic identity of resident native species into fine-grained predictive models for species invasions.

*Acknowledgements* – The research leading to these results received funding from the European Research Council under the European Community's Seven Framework Programme FP7/2007-2013 Grant Agreement no. 281422 (TEEMBIO). TM acknowledges support from the EraNet BiodivERsA project ANR-11-EBID-002 CONNECT. WT and SL acknowledge support from the funded by the French 'Agence Nationale de la Recherche' with the project SCION (ANR-08-PEXT-03). We thank Cristina Roquet for her help with the sensitivity test on polytomies. We are grateful to Jens-Christian Svenning for his insightful comments.

## References

- Abouheif, E. 1999. A method for testing the assumption of phylogenetic independence in comparative data. – *Evol. Ecol. Res.* 1: 895–909.
- Acosta, A. et al. 2008. Patterns of native and alien plant species occurrence on coastal dunes in central Italy. – In: Tokarska-Guzik, B. et al. (eds), *Plant invasions: human perception, ecological impacts and management*. Backhuys Publishers, pp. 235–248.
- Acosta, A. et al. 2009. Are there habitats that contribute best to plant species diversity in coastal dunes? – *Biodivers. Conserv.* 18: 1087–1098.
- Blomberg, S. P. and Garland, T. 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. – *J. Evol. Biol.* 15: 899–910.
- Cadotte, M. W. et al. 2009. Phylogenetic relatedness and plant invader success across two spatial scales. – *Divers. Distrib.* 15: 481–488.
- Carboni, M. et al. 2010. Disentangling the relative effects of environmental versus human factors on the abundance of native and alien plant species in Mediterranean sandy shores. – *Divers. Distrib.* 16: 537–546.
- Carboni, M. et al. 2011. Dealing with scarce data to understand how environmental gradients and propagule pressure shape fine-scale alien distribution patterns on coastal dunes. – *J. Veg. Sci.* 22: 751–765.
- Catford, J. A. and Downes, B. J. 2010. Using multi-scale species distribution data to infer drivers of biological invasion in riparian wetlands. – *Divers. Distrib.* 16: 20–32.
- Cavender-Bares, J. et al. 2006. Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. – *Ecology* 87: S109–S122.
- Celesti-Grapow, L. et al. 2009. Inventory of the non-native flora of Italy. – *Plant Biosystems* 143: 386–430.

- Chaw, S. M. et al. 1997. Molecular phylogeny of extant gymnosperms and seed plant evolution: analysis of nuclear 18S rRNA sequences. – *Mol. Biol. Evol.* 14: 56–68.
- Chytrý, M. et al. 2009. European map of alien plant invasions based on the quantitative assessment across habitats. – *Divers. Distrib.* 15: 98–107.
- Daehler, C. C. 2001. Darwin's naturalization hypothesis revisited. – *Am. Nat.* 158: 324–330.
- Darwin, C. 1859. *The origin of species*. – John Murray.
- Davies, K. F. et al. 2011. Native communities determine the identity of exotic invaders even at scales at which communities are unsaturated. – *Divers. Distrib.* 17: 35–42.
- Davies, T. J. and Buckley, L. B. 2011. Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. – *Phil. Trans. R. Soc. B* 366: 2414–2425.
- Diez, J. M. et al. 2008. Darwin's naturalization conundrum: dissecting taxonomic patterns of species invasions. – *Ecol. Lett.* 11: 674–681.
- Diez, J. M. et al. 2009. Learning from failures: testing broad taxonomic hypotheses about plant naturalization. – *Ecol. Lett.* 12: 1174–1183.
- Duncan, R. P. and Williams, P. A. 2002. Ecology – Darwin's naturalization hypothesis challenged. – *Nature* 417: 608–609.
- Elton, C. S. 1958. *The ecology of invasions by plants and animals*. – Methuen.
- Frohlich, M. W. and Chase, M. W. 2007. After a dozen years of progress the origin of angiosperms is still a great mystery. – *Nature* 450: 1184–1189.
- Gotelli, N. J. 2000. Null model analysis of species co-occurrence patterns. – *Ecology* 81: 2606–2621.
- Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. – *J. Stat. Softw.* 33: 1–22.
- Haibe-Kains, B. et al. 2008. A comparative study of survival models for breast cancer prognostication on microarray data: does a single gene beat them all? – *Bioinformatics* 24: 2200–2208.
- Hanan, E. J. and Ross, M. S. 2010. Across-scale patterning of plant–soil–water interactions surrounding tree islands in southern Everglades landscapes. – *Landscape Ecol.* 25: 463–476.
- 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. – *J. Ecol.* 96: 914–926.
- Huston, M. A. 1999. Local processes and regional patterns: appropriate scales for understanding variation in the diversity of plants and animals. – *Oikos* 86: 393–401.
- Jiang, L. et al. 2010. An experimental test of Darwin's naturalization hypothesis. – *Am. Nat.* 175: 415–423.
- Jones, C. C. et al. 2010. Combining local- and large-scale models to predict the distributions of invasive plant species. – *Ecol. Appl.* 20: 311–326.
- Kembel, S. W. et al. 2010. Picante: R tools for integrating phylogenies and ecology. – *Bioinformatics* 26: 1463–1464.
- Kraft, N. J. B. et al. 2007. Trait evolution, community assembly, and the phylogenetic structure of ecological communities. – *Am. Nat.* 170: 271–283.
- Lambdon, P. W. and Hulme, P. E. 2006. How strongly do interactions with closely-related native species influence plant invasions? Darwin's naturalization hypothesis assessed on Mediterranean islands. – *J. Biogeogr.* 33: 1116–1125.
- Marco, A. et al. 2010. From the backyard to the backcountry: how ecological and biological traits explain the escape of garden plants into Mediterranean old fields. – *Biol. Invasions* 12: 761–779.
- Pavoine, S. and Bonsall, M. B. 2011. Measuring biodiversity to explain community assembly: a unified approach. – *Biol. Rev.* 86: 792–812.
- Pillar, V. D. and Duarte, L. D. S. 2010. A framework for meta-community analysis of phylogenetic structure. – *Ecol. Lett.* 13: 587–596.
- Proches, S. et al. 2008. Searching for phylogenetic pattern in biological invasions. – *Global Ecol. Biogeogr.* 17: 5–10.
- Pyšek, P. et al. 2004. Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. – *Taxon* 53: 131–143.
- Qian, H. and Kissling, W. D. 2010. Spatial scale and cross-taxon congruence of terrestrial vertebrate and vascular plant species richness in China. – *Ecology* 91: 1172–1183.
- Ricciardi, A. and Atkinson, S. K. 2004. Distinctiveness magnifies the impact of biological invaders in aquatic ecosystems. – *Ecol. Lett.* 7: 781–784.
- Richardson, D. M. and Pyšek, P. 2006. Plant invasions: merging the concepts of species invasiveness and community invasibility. – *Progr. Phys. Geogr.* 30: 409–431.
- Ricotta, C. et al. 2010. Invasiveness of alien plants in Brussels is related to their phylogenetic similarity to native species. – *Divers. Distrib.* 16: 655–662.
- Rocchini, D. et al. 2010. Spectral variation versus species beta-diversity at different spatial scales: a test in African highland savannas. – *J. Environ. Monit.* 12: 825–831.
- Schaefer, H. et al. 2011. Testing Darwin's naturalization hypothesis in the Azores. – *Ecol. Lett.* 14: 389–396.
- Schoener, T. W. 1970. Nonsynchronous spatial overlap of lizards in patchy habitats. – *Ecology* 51: 408–418.
- Stachowicz, J. J. and Tilman, D. 2005. Species invasions and the relationships between species diversity, community saturation, and ecosystem functioning. – In: Sax, D. F. et al. (eds), *Species invasions: insights into ecology, evolution, and biogeography*. Sinauer, pp. 41–64.
- Stohlgren, T. J. et al. 1997. Multiscale sampling of plant diversity: effects of minimum mapping unit size. – *Ecol. Appl.* 7: 1064–1074.
- Stohlgren, T. J. et al. 2002. Assessing vulnerability to invasion by nonnative plant species at multiple spatial scales. – *Environ. Manage.* 29: 566–577.
- Strauss, S. Y. et al. 2006. Exotic taxa less related to native species are more invasive. – *Proc. Natl Acad. Sci. USA* 103: 5841–5845.
- Swenson, N. G. et al. 2006. The problem and promise of scale dependency in community phylogenetics. – *Ecology* 87: 2418–2424.
- Thuiller, W. et al. 2010. Resolving Darwin's naturalization conundrum: a quest for evidence. – *Divers. Distrib.* 16: 461–475.
- Vamosi, S. M. et al. 2009. Emerging patterns in the comparative analysis of phylogenetic community structure. – *Mol. Ecol.* 18: 572–592.
- van Wilgen, N. J. and Richardson, D. M. 2011. Is phylogenetic relatedness to native species important for the establishment of reptiles introduced to California and Florida? – *Divers. Distrib.* 17: 172–181.
- Webb, C. O. et al. 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. – *Bioinformatics* 24: 2098–2100.
- Westoby, M. 1998. A leaf–height–seed (LHS) plant ecology strategy scheme. – *Plant Soil* 199: 213–227.
- Wikstrom, N. et al. 2001. Evolution of the angiosperms: calibrating the family tree. – *Proc. R. Soc. B* 268: 2211–2220.
- Willis, C. G. et al. 2010. Phylogenetic community structure in Minnesota oak savanna is influenced by spatial extent and environmental variation. – *Ecography* 33: 565–577.
- Willis, K. J. and Whittaker, R. J. 2002. Ecology – species diversity – scale matters. – *Science* 295: 1245–1248.
- Winter, M. et al. 2009. Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. – *Proc. Natl Acad. Sci. USA* 106: 21721–21725.



## *Chapter 3.3*



# Choosing the appropriate indices and null model in invasion community ecology – A simulation study

Gallien, L., Münkemüller, T., Carboni, M., Chalmandrier, L., Lavergne, S. & Thuiller, W.

## ABSTRACT

In the last decade, the use of community ecology approaches to disentangle the mechanisms of invasion has seen a rising interest. However, the methods used differ across studies, limiting the comparison of their results and making difficult to draw general conclusions. Consequently, a set of indices based on similarity between the invaders and the native communities, and associated statistical tests specially designed for invasion ecology has been initially proposed based on theoretical expectations. However, they never have been properly tested. In this work, we developed a *mechanistic community assembly simulation model*, and tested the performance of the proposed methodology at detecting the true invasion drivers at the community scale. Our results demonstrate among the different ways of measuring the similarity between the invader and the native communities, that the most promising index is based on the functional distance between the invader and all species of the communities. Additionally, the best statistical approach for hypothesis testing consists at comparing invaded vs. resistant communities, which is far from being trivial in nature where resistant communities are usually unknown. We conclude that although some of the proposed indices and tests may not be appropriate to infer the causal mechanisms responsible for invasive success, some do show interesting behaviour.

## INTRODUCTION

*“What drives species naturalization at the scale of the community?”* has been a long-standing question in invasion ecology. In 1859, Darwin already developed two seemingly contradictory hypotheses: phylogenetic relatedness of invaders to native communities is predicted to promote naturalization because of appropriate niche-adaptation (i.e. environmental filtering) but is at the same time predicted to hamper naturalization because of niche overlap with native species (i.e. competition, also known as Darwin’s naturalization hypothesis; Daehler 2001).

During the last ten years, there have been successive calls for the development and use of indices able to

efficiently unravel the driving processes of invasion (Daehler 2001; Richardson & Pysek 2006; Proches *et al.* 2008; Thuiller *et al.* 2010). Today, among the plethora of studies investigating invasion processes, there is no unified methodological approach (e.g. Duncan & Williams 2002, Lambdon & Hulme 2006) and very few studies have been using indices and statistical tests specifically designed for invasions (e.g. Strauss *et al.* 2006, Carboni *et al.* 2012). As a consequence, most invasion studies using community plot surveys rely on methods that await theoretical validation.

To address this problem, Thuiller and colleagues (2010) have reviewed and proposed a set of indices and associated null models specifically designed to study invasion drivers. These indices estimate the niche similarity between the invader and all or a selected set of native species of the communities (either via functional trait differences or phylogenetic distances; see indices in Table 1), under the assumption that species similarity indicates the level of niche overlap and thus the intensity of competition. Null models are then usually used to identify whether the observed pattern are significantly different from the expected pattern under neutral community assembly (i.e. when species functional characteristics do not influence naturalization success). To go a step further, we decided to test whether successful and non-successful invasions show significantly different patterns, and thus use regression models (i.e. testing whether the success of invasion can be well explained by the indices). Two types of regression models can be applied: one comparing invaded communities vs. resistant communities (i.e. invasibility test), or one comparing successful vs. non-successful invaders (i.e. invasiveness test). It can be noted that the regression models can be influenced by the environmental characteristics of the region (i.e. all communities present in the dataset). In fact, when all communities of the dataset are experiencing the same environmental conditions: *environmental homogeneity* (e.g. all communities are sampled under the same climate; which is different from the intensity of the environmental filtering), model capacities for detecting environmental filtering decreases, making thus more easy the detection of competitive

Measure	Description	Assumptions
<b>MDNS</b> <i>Mean distance to the native species</i>	Mean functional distance between each native species and the invader	Each native species contribute equally to the overall biotic resistance of the community
<b>DNNS</b> <i>Distance to the nearest native species</i>	Distance of the invader to the most similar native species	Community's biotic resistance is driven by the native species that is the most similar to the invader
<b>WMDNS</b> <i>Weighted mean distance to the native species</i>	Mean functional distance of each native species to the invader, weighted by the abundances of native species	The contribution of each species to the overall biotic resistance of the community depends on its relative abundance
<b>DMAS</b> <i>Distance to the most abundant species</i>	Functional distance between the invader and the most abundant native species	Community's biotic resistance is driven by the most abundant native species

Table 1. The four tested indices and their underlying theoretical foundations (modified Thuiller *et al.* 2010)

mechanisms (Cavender-Bares *et al.* 2006; Swenson *et al.* 2006; Carboni *et al.* 2012).

Using the set of indices and associated null models proposed in a previous work to study invasion (Thuiller *et al.* 2010), we provide here a first analysis that permits to choose adequately the well performing indices in invasion community ecological studies. To do so, we built a simulation model of community assembly, with explicit definition of the assembly rules: environmental filtering and/or competition. With this model we had three objectives: (1) identify the best performing indices, (2) identify the best performing regression model, and (3) test the impacts of the level of environmental heterogeneity on the results.

## METHODS

### *1. The simulation model*

The simulation model can be described in four main steps (Fig. 1; for more details see Appendix 1). Note that we run a sensitivity analysis to appropriately select for the fixed parameters of the simulation model (see Appendix 2).

#### *Step 1: Create the species and community pools*

We generated 10 different species pools of 400 species by simulating phylogenies and trait evolution along these phylogenies (with a constant rate of trait evolution). Each species was characterized by a single trait value that defined its niche optimum; the species' niches were described by a bell-shaped curve centred on this trait optimum with a fixed variance (equal for all species). Then for each of 10 species pool, 200 species were selected to be "exogenous" species, and their position in the phylogeny was chosen to be over-dispersed.

For each species pool, two sets of communities were defined, a set containing only communities under the same environmental conditions, and a set containing an equal representation of three distinct environmental conditions (equally spaced along a single climatic gradient).

#### *Step 2: Assemble the communities*

We assembled native communities without letting the possibility to enter for "exogenous" species. Each community was initialized with 100 individuals randomly drawn from the species pool of native species only. For each simulation step, 100 random individuals were sequentially removed from the communities and replaced by individuals from the native species pool (asynchronous updating) according to their probability to enter deduced from the assembly rules using a lottery model. The probability of an individual from a given species to enter into the community depends on the specified assembly rules and their relative importance. The environmental filter attributes a higher probability to enter the community to the species with the closest niche optimum to the environmental condition of the community. The competition filter attributes a higher probability to enter the community to species with a low niche overlap with the niche of the individuals already present (symmetric competition). Note that as conspecifics had the same trait values, intra-specific competition was stronger than inter-specific competition (i.e. negative density-dependence). A recruitment filter, attributing higher probabilities to enter the communities to more abundant species already in the community, was also used to counteract the high intra-specific competition value generated by the competition filter (recruitment is fixed in this study).

#### *Step 3: Invasion of the native communities*

Once the species compositions of the native communities were stabilized, we allowed the possibility of invasion by the “exogenous” species of their respective species pools. The invasions followed exactly the same rules as the ones used for assembling the native communities. This lag of potential invasion follows the theoretical expectation that native communities assemble first during a large number of generations, and then long-distance dispersal or human transportation allows exogenous species to potentially invaded those native communities.

*Step 4: Indices and null expectations*

For each successful invader, in each successfully invaded community, we calculated the four indices proposed in Table 1. Then, to test the significance of the results we calculated the indices for unsuccessful invasion according to the two modelling viewpoints: invasibility and invasiveness. For the invasibility test, the same indices were calculated in resistant

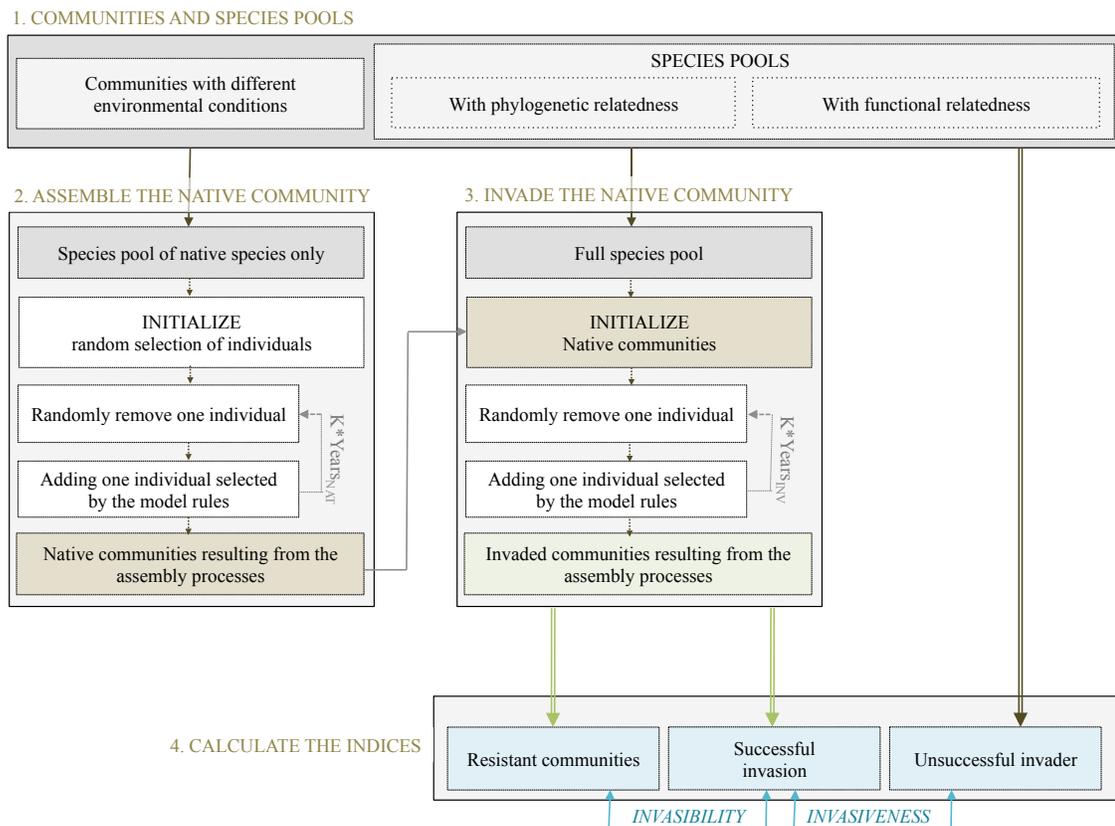
communities (non-invaded communities), while for the invasiveness test, indices of unsuccessful invasions were calculated for unsuccessful invaders (invaders that could not enter the communities).

2. The statistical models

Finally, to test whether each index can significantly predict the invasion success and underlying process, we applied mixed effect regression models. For each parameter combination (environmental and competition filters) and each species pool (10 in total), we fitted three models. The models were of the type:

$$\text{Invasion success (0/1)} \sim \text{Predictor} + (1|\text{SpeciesID}) + (1|\text{SiteID})$$

*Predictor* being either: the tested index as a linear predictor, the tested index as a quadratic predictor, or equal to 1 (intercept model). From the three models, the one with the lowest AIC was retained and considered best.



**Figure 1.** Schematic description of the simulation model.

The simulation model can be described in four steps: (1) first, the creation of species and community pools, (2) then the assembly of the native communities, (3) followed by the invasion of these communities, and (4) finally the calculation of the indices for both successful and unsuccessful invasions.

## RESULTS & DISCUSSION

Generally, we found that the significant detection of the true driving processes of community assembly (i.e. the rules explicitly used in the simulations) strongly varies across indices, regression models, and the level of environmental heterogeneity.

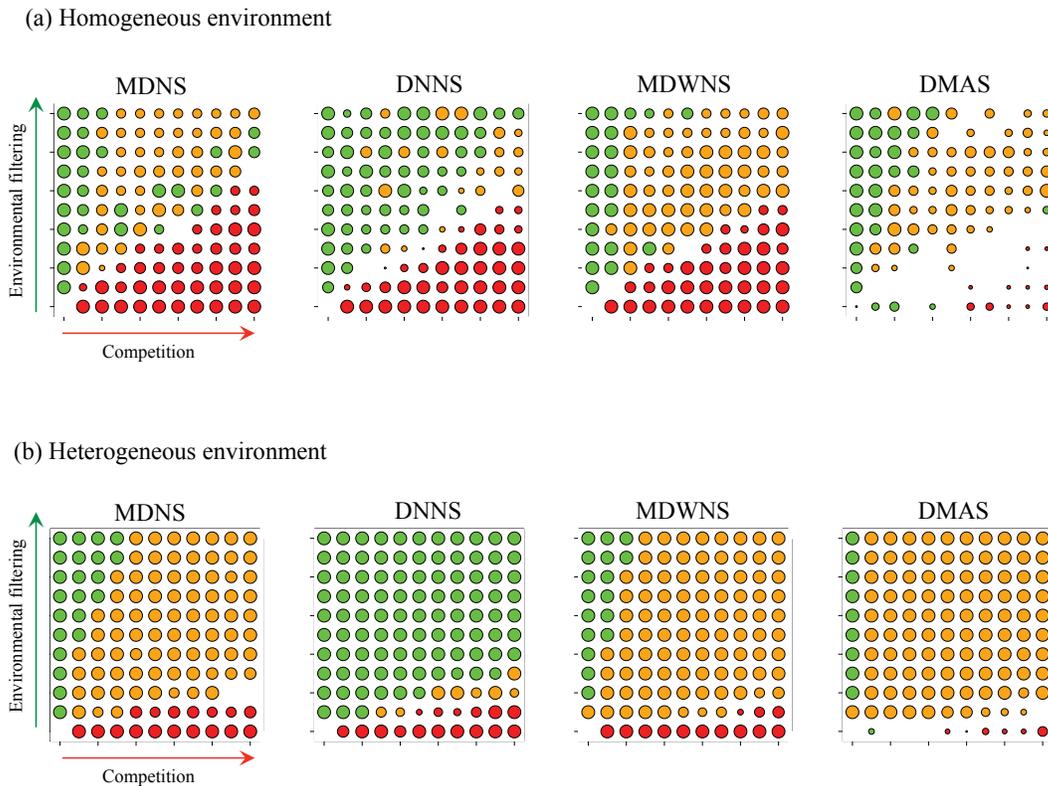
### 1. Robustness of the indices in detecting the right invasion mechanism.

The four indices appear to detect differently the two driving processes and their interactions (Fig. 2). Generally, it seems that the mean distance between the invader and the other species weighted (WMDNS) or not (MDNS) by species abundance perform relatively well. The distance between the invader and the most related species (DNNS)

generally failed to detect the interactions between the filters, and instead tend to wrongly detect signal of environmental filtering. Finally the distance to the most abundant species (DMAS) completely fails at detecting competitive interactions.

### 2. Tests of invasibility vs. invasiveness

When comparing the two null models: invasibility test (i.e. invaded vs. resistant communities; Fig. 2) and invasiveness test (i.e. successful vs. non-successful invaders; Fig. 3), we found that only invasibility model permits to identify the right invasion mechanisms. Indeed, when two processes are effectively driving the community assembly (environmental and competition filters), the test of invasibility permits to identify that both occur, while the invasiveness test always identifies only one or none of the two processes.



**Figure 2.** Performance of the indices under the *invasibility* null model.

The circles represent the results for each tested community assembly rules. The colours indicate the shape of the relationship between the invasion success and the index value: red for a positive relationship (i.e. more invasion success when native species are functionally different), green for a negative relationship (i.e. more success when native species are similar), and orange for a negative quadratic relationship (i.e. more success when the invader is not too similar and not too dissimilar from the natives: bell shape). The circle size represents the proportion of significant relationships.

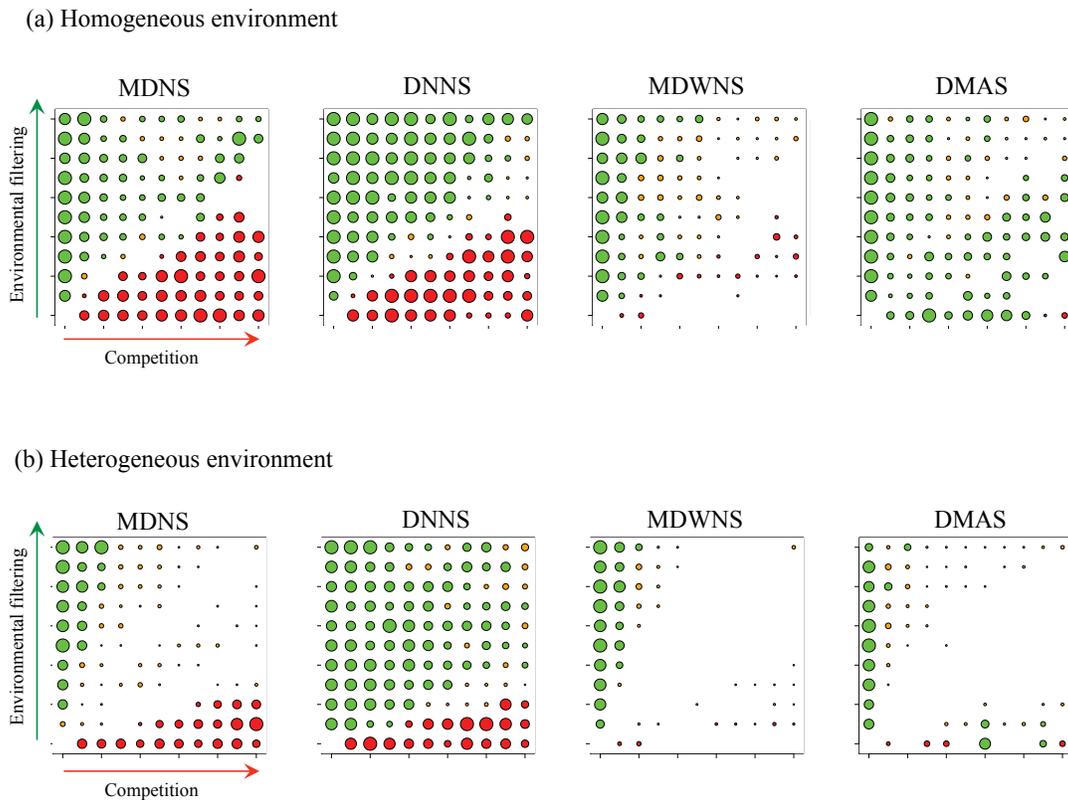
In theory if the index works as expected, when we move toward environmental filtering, the index should detect positive environmental filtering (green circle), while under competition, colours should be red. Under both environmental filtering and competition, circles should be orange.

### 3. Environmental heterogeneity vs. homogeneity

The performance of the four indices and regression models has been also found to vary according to the level of environmental heterogeneity in the dataset (e.g. Fig. 2). Indeed, when the environment is homogeneous across the considered communities, the detection of competition as the only driver of invasion success even though the two filters are interacting, is much more important than when the environment is heterogeneous. This is can be explained by the fact that, in heterogeneous environments (and when both environmental and competition filters interplay) there is more chance to find maladapted species for marginal environmental conditions. Thus, it is easier to detect both competition and environmental filtering in heterogeneous community pools.

### 4. General conclusions

Overall, we showed that the best performing methodology is to use (1) the *mean distance to the species* (MDNS), eventually weighted by species abundances (MDWNS) if good estimates of species abundances are available. At best, the interpretation of this index should be tested with (2) the *invasibility regression* model comparing invaded vs. resistant communities. Finally, if the objective of the study is only to detect whether there is or not competitive interactions involved in the invasion success: *environmental homogeneity* in the dataset will perform well; while if the objective of the study is to detect interactions between environmental and competition filtering: a *heterogeneous* dataset would be more suited.



**Figure 3.** Performance of the indices under the *invasiveness* null model.

The circles represent the results for each tested community assembly rules. The colours indicate the shape of the relationship between the invasion success and the index value: red for a positive relationship (i.e. more invasion success when native species are functionally different), green for a negative relationship (i.e. more success when native species are similar), and orange for a negative quadratic relationship (i.e. more success when the invader is not too similar and not too dissimilar from the natives: bell shape). The circle size represents the proportion of significant relationships.

In theory if the index works as expected, when we move toward environmental filtering, the index should detect positive environmental filtering (green circle), while under competition, colours should be red. Under both environmental filtering and competition, colours should be orange.

## 5. Perspectives

In this study we are using “perfect” datasets, which is probably never the case in analysis of field observations. Indeed, in the regression models we compare successful invasions to communities that we know are resistant, or to exogenous species that we know are non-successful. Therefore, in the next months further simulations will permit to test and quantify the sensibility of the methods to a gradual increase of proportion of “imperfect” data in the dataset. “Imperfect” data will be either (1) communities that we know could be invaded but were not (e.g. due to dispersal limitation of the invaders), or (2) exogenous species that we know would be successful but were not (e.g. not introduced yet).

## REFERENCES

- Carboni, M., Münkemüller, T., Gallien, L., Lavergne, S., Acosta, A. & Thuiller, W. (2012). Darwin’s naturalization hypothesis: scale matters in coastal plant communities. *Ecography*, 35, 001-009.
- Cavender-Bares, J., Keen, A. & Miles, B. (2006). Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. *Ecology*, 87, 109-122.
- Daehler, C.C. (2001). Darwin's naturalization hypothesis revisited. *American Naturalist*, 158, 324-330.
- Darwin, C.R. (1859). *The Origin of Species*. John Murray, London.
- Duncan, R.P. & Williams, P.A. (2002). Darwin’s naturalization hypothesis challenged. *Nature*, 417, 608.
- Lambdon, P.W. & Hulme, P.E. (2006). Predicting the invasion success of Mediterranean alien plants from their introduction characteristics. *Ecography*, 29, 853-865.
- Proches, S., Wilson, J.R.U., Richardson, D.M. & Rejmanek, M. (2008). Searching for phylogenetic pattern in biological invasions. *Global Ecology and Biogeography*, 17, 5-10.
- Richardson, D.M. & Pyšek, P. (2006). Plant invasions: merging the concepts of species invasiveness and community invasibility. *Progress in Physical Geography*, 30, 409-431.
- Strauss, Y.E., Webb, C.O. & Salamin, N. (2006). Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Sciences of the USA*, 103, 5841-5845.
- Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. & Zimmerman, J.K. (2006). The problem and promise of scale dependency in community phylogenetics. *Ecology*, 87, 2418-2424.
- Thuiller, W., Gallien, L., Boulangeat, I., de Bello, F., Muenkemueller, T., Roquet, C. & Lavergne, S. (2010). Resolving Darwin's naturalization conundrum: a quest for evidence. *Diversity and Distributions*, 16, 461-475.

## Appendix 1. Detailed description of the simulation model.

### **The simulation model**

We compared the performance of different indices and null models using a simulation model of community assembly using explicit assembly rules. Species information and the environmental conditions of the sites are the inputs for the simulation model. The observed species in the communities form the output of this model, and the species pools provide the basis for the calculation and analyses of the different indices. This approach can be described in four steps introduced in Figure 1.

#### *Step1: Create the species and community pools*

In a first step, we generated 10 different species pools of 400 species by simulating phylogenies and trait evolution along these phylogenies (with a fixed rate of trait evolution over time). Each species was characterized by a single trait value that defined its niche optimum, the species' niches were all Gaussians centred on this trait optimum with a fixed variance (equal for all species). Then within each of the species pool 200 species were sectioned to be "exogenous" species, and their position in the phylogeny was overdispersed. Because in real datasets invaders are not necessarily over-dispersed in the phylogeny (e.g. Carboni *et al.* 2013), we also tested for other distribution of the invaders in the phylogeny (clustered and random) but found no qualitative differences.

#### *Step 2: Assemble the communities*

In a second step, we assembled native communities without letting the possibility to enter for "exogenous" species. Each community was initialized with 100 individuals randomly drawn from the species pool of native species only. For each simulation step, 100 random individuals were sequentially removed from the communities and replaced by individuals from the native species pool (asynchronous updating) according to their probability to enter deduced from the assembly rules using a lottery model. The probability of an individual from species  $i$  entering the community  $k$ ,  $P_{all,i,k}$  depends on the specified assembly rules and their relative importance defined the factors  $B_{env}$  (environmental filtering),  $B_{comp}$  (competition) and  $B_{abun}$  (recruitment).

$$P_{all,i,k} = \exp(B_{env} \times \log(P_{env,i,k}) + B_{comp} \times \log(P_{comp,i,k}) + B_{abun} \times \log(P_{abun,i,k})) \quad (1)$$

$P_{env,i,k}$  modelled the environmental filter: the closer the species trait value (i.e. species' niche optimum) was to the environmental conditions of the community  $k$ , the higher was its probability to enter.  $P_{comp,i,k}$  modelled the competition filter: the closer the species trait value was

to those of the individuals already present, the lower was its probability of entering. In this way, competition between individuals was defined as symmetric. Note that as conspecifics had the same trait values, intra-specific competition was stronger than inter-specific competition.  $P_{abun,i,k}$  modelled the recruitment filter: the more abundant the species was in the community, the higher its probability of entering. This term counteracted the high intra-specific competition value generated by the competition filter.

The factors  $B_{env}$ ,  $B_{comp}$  and  $B_{abun}$  weighted the importance of the three filters in community assembly. In the special case of  $B_{env}$  and  $B_{comp}$  equalling one, the equation was comparable to a Lotka-Volterra equation with inter- and intra-specific competition and a maximal growth rate dependent on the environmental suitability.

Appendix 2. Fixed parameters and sensitivity analysis of the model.

**Sensitivity analysis:**

To perform the sensitivity analysis we assembled native communities while varying different parameter values (see Table S2.2). Then, for each set of parameter value, we calculated indices of community functional diversity (via the mean functional distance between all pair of species, weighted by their abundances or not). In order to identify whether the indices were revealing patterns significantly different from neutral patterns, we compared the observed indices from those obtained with a random re-attribution of the species trait values (Fig. S2).

Finally, we retained communities assembled for a carrying capacity of 100 individuals (more would also be fine but slower), a niche breath of 10, a weight of regeneration of 10, and a distribution of the invasive species under-dispersed (also tested for indices of invasion processes, and showed not affecting the results).

Table S2.1. List of the parameters fixed and tested in the simulation of the study.

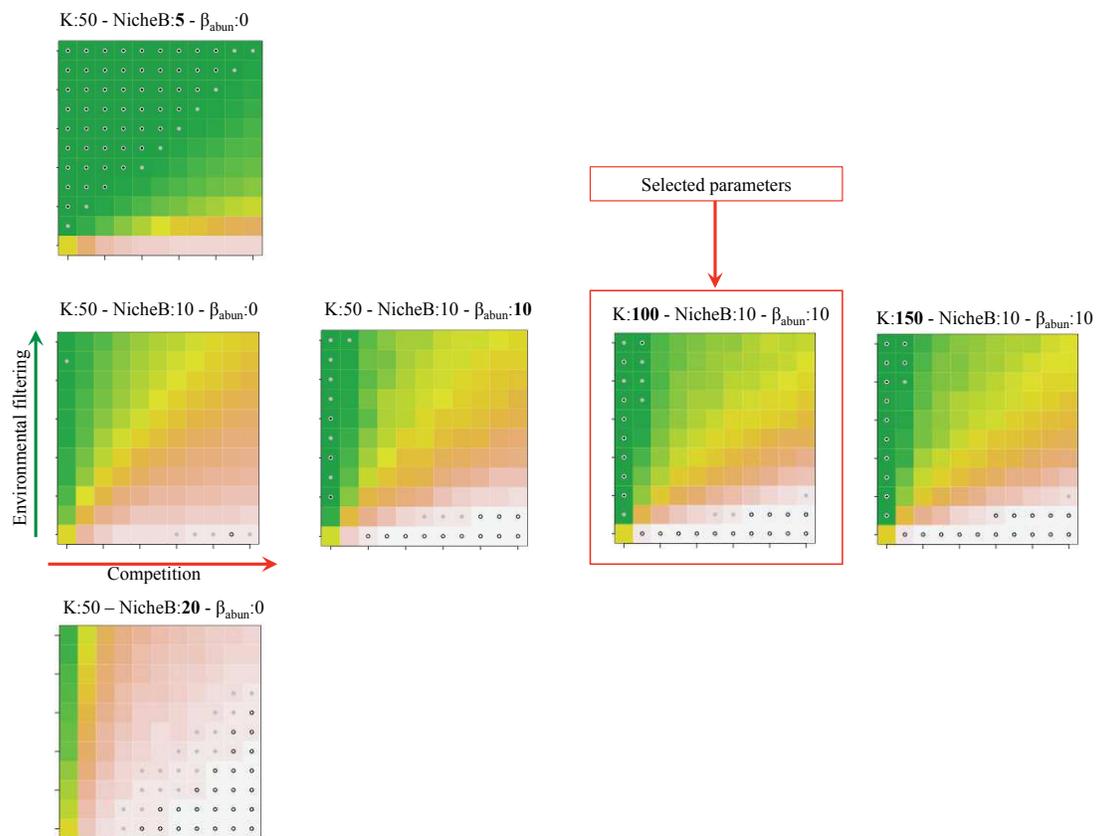
Parameters	Acronyms	Fixed	Tested in the study
Number of Native species pool	Pool <sub>NAT</sub>	200	
Number of Invasive species pool	Pool <sub>INV</sub>	200	
Native specie assembly time	Years <sub>NAT</sub>	10	
Invasion time	Years <sub>INV</sub>	10	
Carrying capacity	K	100	
Niche breath	NicheB	5	
Number of species pool	NbPools	10	
Number of community per species pool	NbCommunities	100	
Invader position in the phylogeny	Position <sub>INV</sub>	Over-dispersed	
Weight of the regeneration	$B_{abun}$	10	
Weight of the environmental filtering	$B_{env}$		0 to 1
Weight of the competition	$B_{comp}$		0 to 10
Dataset Environments	Heterogeneity		Equal proportion of environments = 20, 50, 80 All environments = 50

Table S2.2. List of the parameters tested in the sensitivity analyses.

Parameters	Acronyms	Fixed	Tested in the sensitivity analysis
Native specie assembly time	Years <sub>NAT</sub>		5 – 10 – 20
Carrying capacity	K		50 – 100 – 150
Niche breath	NicheB		5 – 10 – 20
Invader position in the phylogeny	Position <sub>INV</sub>		Under-dispersed – random - clustered
Weight of the regeneration	$B_{abun}$		0 – 1 – 10

Figure S2. Performance of the community assembly model

The colours indicate the rank of the indices in the null distributions (from red: over-dispersed, to green: clustered), and dots indicate the level of the p-values: grey dots for 0.05 tests and black-white dots for 0.025 tests.



## *Chapter 4*



DISENTANGLING THE RELATIVE IMPORTANCE  
OF DIFFERENT COMMUNITY ASSEMBLY RULES  
TO UNDERSTAND INVASION SUCCESS IN  
ALPINE PLANT COMMUNITIES

*COMMENT DEMELER L'IMPORTANCE RELATIVE  
DES DIFFERENTES REGLES D'ASSEMBLAGE POUR  
COMPRENDRE L'INVASION DES COMMUNAUTES  
DE PLANTES DANS LES ALPES*



## PRESENTATION DU CHAPITRE

*Malgré les efforts considérables qui ont été développés pour enquêter sur les processus d'assemblage des communautés qui déterminent le succès d'invasion des plantes, peu de conclusions générales ont vu le jour. Trois processus principaux, généralement agissant comme des filtres successifs, sont anticipés comme étant d'importance capitale. L'invasive doit d'abord disperser dans un site (1<sup>er</sup> filtre) aux conditions environnementales favorables (2<sup>ème</sup> filtre) et réussir à s'établir dans la communauté locale malgré les interactions compétitives (3<sup>ème</sup> filtre). Les interactions compétitives peuvent passer par différentes stratégies: soit par la capacité de l'envahisseur à éviter la compétition (l'opportunité de niche), soit par sa capacité à remplacer une espèce native (exclusion compétitive). De façon surprenante, malgré le consensus général sur le fait que ces trois processus sont important et qu'il faille les étudier ensemble avec leurs interactions, ils sont habituellement étudiés indépendamment. Ici, nous présentons une approche qui permet de relever ce défi, en utilisant un grand jeu de données de communautés de plantes. Dans un premier temps, nous essayons de démêler l'importance relative de ces trois processus dans le succès d'invasion d'espèces végétales. Dans un deuxième temps, nous avons particulièrement étudié les différentes stratégies d'interaction compétitives, et comment elles interagissent avec les deux autres processus.*

*Comme cas d'étude, nous utilisons une base de données contenant plus de 5'000 communautés végétales sur l'ensemble des alpes françaises, pour lesquelles nous avons également des données climatiques et de traits fonctionnels pour la majorité des espèces. Nous avons développé un jeu d'indices permettant de décrire les différentes stratégies d'interactions compétitives. Ensuite grâce à un ensemble de modèles du type «random forest», nous avons mesuré l'importance relative de ces trois processus, ainsi que leurs interactions respectives afin d'expliquer la distribution de 7 espèces herbacées invasives dans les alpes. Pour la majorité des espèces, nous avons montré que (1) les indices d'interaction biotiques sont de bons prédicteurs de la présence des invasives, mais (2) qu'ils sont surtout des indicateurs de stress environnementaux cachés, qui (3) agissent préférentiellement au cœur de la niche abiotique des espèces*



# Disentangling the relative importance of different community assembly rules to understand invasion success in Alpine plant communities

Laure GALLIEN<sup>1\*</sup>, Florent MAZEL<sup>1</sup>, Sébastien LAVERGNE<sup>1</sup>, Julien RENAUD<sup>1</sup>, Roland DOUZET<sup>2</sup>, Wilfried THUILLER<sup>1</sup>.

<sup>1</sup> *Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, Grenoble I, BP 53, 38041 Grenoble Cedex 9, France.*

<sup>2</sup> *Station Alpine Joseph Fourier, CNRS UMS 2925, Université Joseph Fourier, Grenoble I, BP 53, 38041 Grenoble Cedex 9, France.*

## ABSTRACT

Despite considerable efforts to investigate the assembly processes that determine plant invasion success, few general conclusions have so far been drawn. Three main processes, generally acting as successive filters, are thought to be of prime importance. The invader has to disperse (1<sup>st</sup> filter) in a suitable environment (2<sup>nd</sup> filter) and succeed in establishing itself in a given community through competitive interactions (3<sup>rd</sup> filter). Competitive interactions can reflect different strategies, either the invader's ability to avoid competition (niche opportunity), or its ability to replace a native species (competitive exclusion). Surprisingly, despite the general consensus on the importance of investigating these three processes and their interplay, they are usually studied independently. Here, we present an approach that meets this challenge using a large plant community data set. We first aim to disentangle the relative importance of these processes in determining plant invasion success. Secondly, we considered how competitive interactions influence the success of plant invasions and how they interact with the other two processes.

As a case study, we used a vegetation-plot database for the French Alps containing 5,000 plots for which the species functional traits and environmental information were available. We developed a suite of indices to depict the different components of competitive interactions. Then, using a set of random forest models we measured the relative importance of

the three processes and their respective interactions to explain the distribution of seven herbaceous invaders in natural communities. For most species, we showed that adding dispersal and biotic indices significantly improved model performance, and that the main drivers of invader presence were: temperature, land use, human footprint, dispersal, and invader height compared to the average community height. Overall, we show that combining environmental, dispersal and biotic information to model invader presence has excellent potential for improving our understanding of their success.

## INTRODUCTION

Understanding the processes that favour or hamper biological invasions in natural communities is crucial to better anticipating their potential spread and impacts (Rejmánek *et al.* 2005). Even though accurately modelling and predicting invasive species distribution remains challenging (Gallien *et al.* 2010), it is generally well accepted that three major ecological processes acting as successive filters influence the outcome of invasion. The species has to disperse (1<sup>st</sup> filter) in a suitable environment (2<sup>nd</sup> filter) and succeeds in establishing itself in a given community through biotic interactions (3<sup>rd</sup> filter). Previous works have demonstrated that dispersal (e.g. propagule pressure; Simberloff 2009), abiotic

environment (e.g. climate matching), and biotic interactions (e.g. enemy release; Mitchell *et al.* 2006) are important determinants of biological invasions (Higgins, *et al.* 2000; Roura-Pascual *et al.* 2009; Thuiller *et al.* 2012). However, it remains unclear whether interspecific competitive interaction, in particular, plays a prominent role in driving invasion success and how it may interact with dispersal and environmental gradients (Von Holle & Simberloff 2005; Simberloff 2006; Seastedt & Pysek 2011).

It is our understanding that some of the divergences between previous studies focussing on competition result from conceptual confusion regarding the expected outcomes of this competition process. Indeed, there are only two possible ways for an invader to succeed in the context of competitive interactions. First, the invader occupies an empty niche (e.g. an unused resource at the community level) regardless of whether it is a good or bad competitor. This strategy, also called *niche opportunity*, is made possible due to niche differences between the invader and native species. Secondly, the invader is a better competitor than at least one other native species and it will tend to replace it in the communities (i.e. *competitive exclusion*). Although the two strategies are relatively straightforward to explain, there is an issue as to how to measure and distinguish between them using adequate community-level indices (HilleRisLambers *et al.* 2012).

The use of an empty niche by the invader can be estimated from two possible viewpoints. Firstly, indices measuring the diversity of the native community (e.g. taxonomic or functional diversity) can be considered as proxies for the level of niche filling in a local community. These proxies are based on the hypothesis that more diverse communities have a higher level of niche filling than less diverse ones, which should mean that there is a lower probability of invasion by a poor competitor (Fargione & Tilman 2005). However, community diversity does not only reflect the level of niche filling, as it can also reveal the quantity of resources available in the community (Shea & Chesson 2002). We therefore developed an index that measures the level of competition between species in a community, based on the assumption that an increase in competition intensity should confer a certain degree of biotic resistance on this local community (Herben 2005). In order to do so, we estimated the proportion of the regional species pool that is effectively observed in these communities (community richness) relative to the number of species that could occur in the community according to their environmental requirements (potential community richness). Similar types of indices have proved to be useful to estimate the intensity of competition over large environmental gradients (Boulangeat *et al.* 2012). For a given level of potential community richness (and assuming homogeneous spatial isolation over the study region),

we expect that, (i) when the observed richness is comparatively low, it indicates a high interspecific competition intensity because the low number of species already present in the community have the ability to prevent more species from entering, while (ii) when the observed richness equals the potential one (i.e. when all species that could survive in local climatic conditions are present in the community), it indicates that the competition intensity is not high enough to drive species to exclusion. Secondly, indices based on measures of functional differences between the potential invader and the native species of a given community can also be used to identify whether the invader tends to occupy empty niches unused by native species in local communities (Thuiller *et al.* 2010). Two predictions are typically proposed. In highly stressful environments, a successful invader needs to be pre-adapted to survive and therefore tends to be functionally similar to the natives. In a non-stressful environment, where invasion success is mostly limited by resource competition, invaders which are functionally dissimilar from native species will be favoured (Darwin 1859; Duncan & Williams 2002; Richardson & Pysek 2006). Interestingly, these indices are assumed to indicate whether competition is driving invasion, whereas in reality, they only reflect niche differences and cannot detect when the invader is a superior competitor (Mayfield & Levine 2010).

Some functional or morphological traits are known to predict species' *competitive ability* and are therefore useful in community ecology to delimitate the species which are more likely to win resource competition. Competitive ability can be estimated by ranking the invader's competitive value compared to the competitive values for native species (asymmetric competition; Chesson 2000; Mayfield & Levine 2010). Indeed, it has been repeatedly shown that some key morphological and functional traits correctly reveal species' competitive abilities (Westoby *et al.* 2002), and that these traits can also reveal competitive interactions in natural communities (Kunstler *et al.* 2012). According to Chesson's framework (2000), in the absence of niche differences between species but marked differences in their competitive abilities, only the best competitor is expected to survive. However, several good competitors can also co-exist locally in a fluctuating environment or due to demographic stochasticity (i.e. neutral-like dynamics of functionally similar species; Chesson 2000; Hubbell 2001; Thuiller *et al.* 2007).

The lack of consistency in the identified assembly rules that drive invasion success across a large number of empirical studies (see examples in Levine & D'Antonio 1999; Lonsdale 1999; Shea & Chesson 2002; Mitchell *et al.* 2006; Proches *et al.* 2008) can be explained by three methodological confusions. Firstly, in most published studies only one of the two of invasion strategies detailed above (i.e. niche

opportunity vs. competitive exclusion) is properly tested, while in fact the interaction between them may determine the distribution of each invader according to its stage of invasion, the environment and the characteristic community. Indeed, it is possible that niche opportunity and environmental filtering prevail in the early stages of invasion when the invader still occurs in disturbed habitats (i.e. transport and colonisation stages), and that the importance of competitive ability increases when the invader establishes in natural communities (i.e. establishment and spread stages, Theoharides & Dukes 2007). The second confusion arises when selecting the interacting species within the communities. Indeed, not all species in a community necessarily compete for the same resources. For example for light interception, *niche opportunity* is probably of prime importance for the coexistence of different growth forms (e.g. herbaceous vs. woody species), while *competitive ability* drives the outcome of assembly between woody species (e.g. Kunstler *et al.* 2012; but see also Fargione *et al.* 2003). Finally, the position of the community within the environmental niche of the invasive species is often ignored, while the importance of competition may vary according to the strength of the environmental stress (Körner 1999; Callaway *et al.* 2002, see also Welden & Slauson 1986; Kikvidze *et al.* 2011). Indeed, the competition filter is thought to be less important at the niche edge where physiological constraints limit species distribution (e.g. in cold alpine environments; Callaway & Walker 1997; *et al.* 2013), while it can be more important at other niche edges where the distribution of the focal species is limited by negative biotic interaction with another species (e.g. presence of a predator; Wisz *et al.* 2013).

In this paper we present an approach for investigating the relative importance of the three general processes thought to drive plant invasion success: environmental filtering, dispersion and competitive interactions. As a case study, we use an extensive dataset of 5,000 plant community sites over the French Alps area, and model the invasion success of seven herbaceous plant invaders in these communities. Building on previous works demonstrating that environmental filtering is an important driver of biological invasions, we evaluate whether including dispersal and community-wide indices, depicting competitive interactions, in predictive models increases their performance in identifying invaded sites. In order to distinguish between the two competitive interaction strategies (niche opportunity vs. competitive ability), we use a set of different metrics based on observed and potential richness, as well as functional trait differences between the invaders and native species, accounting for different species life-forms. Finally, we test whether the effects of environmental gradients and competition indices do interact or are actually independent.

## MATERIAL & METHODS

### Data

#### Community data

We made use of an extensive database of vegetation survey plots spanning the French Alps, collected by the Alpine Botanical Conservatory (Boulangeat *et al.* 2012). Each vegetation plot ('community' hereafter) consists of an exhaustive survey of homogeneous plant assemblages (about 10x10m) with species relative abundance. The total number of available communities was 15,931 including around 3,030 species. Given our interest in studying both (1) biotic interactions between the invader and all other species, but also (2) biotic interactions between the invader and the species of the same life forms (i.e. herbaceous species), we only kept 5,141 communities for which functional trait data (see below) were available for at least 70% of species cover across all species and for herbaceous species only (according to the 'biomass ratio hypothesis', Grime 1998).

Among the 142 exogenous species of the French Alps, we focused on the invaders that were either classified as "colonising", "establishing" or "spreading" (*sensu* Theoharides & Dukes 2007), recorded at least 30 times within the 5,141 communities, and for which the functional traits were available (see below). Furthermore, in order to consider biotic interactions between species of the same life form we chose to work only with herbaceous invaders. We finally retained 7 herbaceous invaders (which occurred between 30 and 142 times in community plots): *Ambrosia artemisiifolia* L., *Bidens frondosa* L., *Conyza canadensis* (L.) Cronquist, *Erigeron annuus* (L.) Desf., *Panicum capillare* (L.) , *Solidago canadensis* L., and *Solidago gigantea* Aiton.

#### Environmental data

A set of five environmental variables (three bioclimatic, one human footprint, and one land cover) known to be important for species establishment and spread was used to estimate invader distribution. Three climatic variables originating from the French meteorological model Aurelhy (Bénichou & Le Breton 1987) based on interpolated measurements at a 100m resolution and summarising climatic information over the last 30 years (here 1971-2000) were considered. These variables were: the annual sum of degree-days with a 5.56°C threshold (ddeg556) as a measure of the environmental heat vital to plant biomass production, the mean annual level of solar radiation as an important driver of plant growth rate, and mean annual precipitation as an indicator of water stress either for water deficiency or water overload. Land cover information was

extracted from the CORINE Land Cover map for Europe (CORINE 2006, <http://www.epa.ie/whatwedo/assessment/land/corine>) at 250m resolution, and indicates the type of ecosystem in which the communities have been recorded.

Using the same methodology as Sanderson *et al.* 2002, we built a human footprint map of the French Alps at a 250 m resolution based on CORINE Land cover 2006, including the population density (Gallego 2010; <http://www.eea.europa.eu/data-and-maps/data/population-density-disaggregated-with-corine-land-cover-2000-2>), the sum of road length per pixel (BD CARTO®, Institut Géographique National), and a normalized value of light pollution (<http://www.ngdc.noaa.gov/dmsp/downloadV4composites.html>). This human footprint variable ranging from 0 (wild) to 1 (highly disturbed) is used to indicate the degree of disturbance in the communities.

### Functional trait data

We focused on three key functional traits that can play a role in both niche differentiation and competitive ability: the specific leaf area (SLA; light-capturing area deployed per unit of leaf dry mass), the height of the plant's canopy at maturity (Height) and the seed mass (SeedM), which are well-known components of the leaf-height-seed (LHS) syndrome of plant traits (Westoby 1998). SLA reflects species' relative growth rate, differences among species in terms of water use efficiency and competitive abilities for nitrogen (Grime 1998; Suding *et al.* 2005, Angert *et al.* 2009), SeedM reflects the dispersal distance and establishment success (ability to colonise new sites), and Height captures each species' ability to intercept light (Westoby *et al.* 2002). These traits were extracted from the trait database ANDROSACE (AlpiNe Database ResOurce for Species And eCosystems fEatures, Thuiller *et al.* unpublished data). The database includes trait information for Alpine plants from individual projects and freely available databases such as LEDA (Knevel *et al.* 2003), BioFlor (Kühn, Durka & Klotz 2004), Ecoflora (Fitter & Peat 1994) and CATMINAT (Julve 1998). Although other functional traits may influence species co-existence (e.g. the level of leaf nitrogen content, the degree of plasticity), they were not available for all species occurring in the 5,141 selected communities.

### Statistical analyses

To estimate the relative influence of the three processes (environment, dispersal and competitive interaction) on the presence and the absence of each selected invaders, for each species we built three nested sets of species distribution models (using

random forest algorithms): (i) environmental variables only (5 variables), (ii) environmental and dispersal variables (6 variables), and (iii) environmental, dispersal and biotic interaction variables (13 variables). The importance of each variable was then evaluated by a randomisation procedure and the model performances were compared (see below for more details). All statistical analyses were carried out in R 2.15.2 (Development Core Team 2012; <http://www.R-project.org>) with the package *party* (v.1.0-2; Hothorn *et al.* 2006).

### Random Forest models

Given that we had no expectations regarding either the form of the relationships between the invader's presence and the set of variables, or the type and degree of interactions between the variables, we needed a data-driven modelling technique with internal optimisation procedure. In addition, we needed a modelling technique that makes it possible to identify the variable relative importance when a correlation is found between a set of variables. These two requirements meant it was not possible to use traditional stepwise logistic regressions. Instead, we used Random Forests (RF) from the R package {*party*}. A random forest is a bootstrap aggregation (known as “bagging”, Breiman 2001) of classification trees (here 1,000 trees) from subsamples of the original calibration data (here 80% of the full dataset is used to calibrate the model and 20% to evaluate it). All the trees in a forest are different for two reasons: each tree is based on a random subset of observations, and each split within each tree is created with a random subset of variables. Then, the predictions of the RF are a weighted average over the predictions of each single tree, which has been shown to highly outperform single tree predictions (Breiman, 2001). We then ran a 10-fold cross-validation procedure for each species to ensure consistent results across the different subsets of data used for calibration.

We estimated the variable importance using the permutation accuracy importance algorithm, as suggested by Strobl and colleagues (2007; we used 500 permutations). The principle was to apply a random permutation of each variable value in order to mimic the absence of the variable in the model. Then, importance is estimated according to the difference in prediction accuracy (i.e. mean squared error) without and with the permutation (Strobl *et al.* 2007). For each species we extracted the median variable importance across the ten cross-validation models.

The quality of the RF for each species was measured using an Area under the Relative Operating Characteristic curve (AUC). The AUC ranges from 0.5 (prediction not better than random) to 1 (perfect prediction). A model with an AUC that is higher than 0.8 is usually classified as relatively good (Swets 1988). For each species, the ten cross-validated

models were evaluated on the remaining 20% of the data not used for calibration. To test model performance between the different sets of variables we used pairwise Wilcoxon tests.

### Dispersal index

It has been shown that dispersal limitations lead to the spatial clustering of invader presences, also called spatial autocorrelation (hereafter SAC; Dorman *et al.*, 2007). Unsuitable sites can indeed be invaded when they are close enough to source sites (false presences), while suitable sites can be spared from colonisation if they are too distant (false absences). However, apparent spatial autocorrelations of individuals also emerge when the abiotic, and even the biotic, environments are themselves spatially autocorrelated (Legendre & Legendre 2012). As the effects of biotic and abiotic environment are directly modelled in our framework, we suggest that the remaining unexplained spatial autocorrelations are most likely due to dispersal limitations (although they may also emerge from unknown local environmental configurations). To model this unexplained spatial autocorrelation, we used an autocovariate variable (Augustin *et al.* 1996). This modelling approach is the simplest and least computationally intensive method to account for SAC. Although other methods (e.g. Eigenvector Mapping, Dorman *et al.* 2007) are likely to provide a better estimation, the use of an autocovariate term was the only option here because of computational limitation for more than 5,000 sites. The autocovariate term  $D_i$  was then defined as (Augustin *et al.* 1996):

$$D_i = \frac{\sum_{j=1}^{k_i} f(d_{ij}) y_j}{\sum_{j=1}^{k_j} f(d_{ij})}$$

where  $D_i$  at site  $i$  depends of  $f(d_{ij})$  a function of the Euclidian distance between sites  $i$  and  $j$  (here :  $f(d_{ij})=1/d_{ij}$ ),  $k_i$  is the number of sites within the vicinity of the site  $i$ , and  $y_j$  is the response variable (here presence and absences) of the site  $j$ . This index measures, for each site, the potential effect, weighted by the inverse of the distance, of the surrounding sites (up to 10km, that is an approximation of the maximal natural long distance dispersal of the seed rain; see anthropochory in Engler *et al.* 2009).

### Biotic interaction indices

For each of the 5,141 communities we computed 4 indices to describe niche opportunity, and 3 indices of competitive ability (with pair-wise correlations lower than 0.65).

### Niche opportunity indices

To estimate the proportion of available niches in the communities, we calculated the ratio between the observed richness (i.e.  $Rich_{obs}$ , number of species) and the “potential richness” ( $Rich_{pot}$ ) for each community ( $Rich_{obs/pot}$ ). The potential richness was calculated by summing all species that could potentially occur in these communities according to their climatic requirements (Boulangeat *et al.* 2012). This was done by estimating, for each species (occurring more than 20 times in the dataset), their observed climatic ranges, delimited by the minimum and maximum observed values for all abiotic variables simultaneously. A  $Rich_{obs/pot}$  close to one indicates that most species for which the community is climatically suitable have managed to enter the community, and thus the competition intensity should not be too high. Conversely, a low  $Rich_{obs/pot}$  value indicates that only a few species managed to enter the community compared to the total number that theoretically could in terms of climate. This suggests very high levels of interspecific competition.

We also calculated three indices of niche difference based on functional distances (i.e. Gower distances between species using the three traits simultaneously; Thuiller *et al.* 2010). These metrics were used to calculate the functional relatedness between invaders for both all natives and with herbaceous natives only. We calculated (1) the Mean Distance of the invader to *all* Native Community (MDNC<sub>all</sub>), (2) the Mean Distance of the invader to *herbaceous* Native Community (MDNC<sub>herb</sub>), and (3) the Distance of the invader to the Nearest Native species (MDNN; note that this index is the same when considering all natives or herbaceous natives only). These indices do not overlap but are complementary as they suggest different hypotheses as to how biotic interactions drive the integration of a given invader into native communities. MDNC<sub>all</sub> assumed that the invaders compete with all native species with the same strength, MDNC<sub>herb</sub> assumed that the invaders only compete with native species with the same growth form, and MDNN assumed that biotic resistance is mainly driven by the species most similar to the invader because they are supposed to compete for the same resource (i.e. share the same niche in the community).

### Competitive ability indices

The differences in competitive ability between species were calculated separately for each of the three functional traits (SLA, Ht and SeedM) because these traits represent different competitive strategies. SLA contrasts fast growing vs. slow growing or stress resistant species (Wright *et al.* 2004, Westoby & Wright 2006). SeedM separates species with high

seed production but low establishment rates vs. species with low seed production but high establishment rates, and Ht is a key component of competition for light (Westoby et al 2002). Competitive ability indices were calculated as the standardised mean differences in trait value between the invader and each herbaceous species in the community (Kunstler *et al.* 2012). Only herbaceous native species were used to calculate the indices to avoid the comparison of traits between trees and herbaceous species, which may reflect broad life histories and niche differences more than competitive ability. When the index equals 0, it means the invader has the same trait value as the average trait value for the community; it is negative when the invader has a lower value than the community mean, and it is positive when the invader has a higher value than the community mean. These indices differ from those for niche differences because the invader's hierarchical position in the trait gradient of the community is emphasised rather than just the difference in trait values between the invader and the community (Kunstler *et al.* 2012). In other words, because the index is directional it can describe asymmetric competition: establishing for instance whether the

invader is taller or shorter than the other species on average. Indeed, it has recently been shown that for tree species, these indices are better proxies for competitive interaction than niche differences (Kunstler *et al.* 2012).

## RESULTS

### 1. Model performance

For all modelled invasive species, the median model performance evaluated on the validation datasets ranged from good (AUC between 0.80 and 0.90) to very good (AUC > 0.90) for all sets of variables. In general, when considering the full model (with abiotic, dispersal, and biotic variables) we found that the three types of variables were important for most species (Fig. 2). However, in most cases, the inclusion of biotic interaction indices improved model accuracy, while the benefit of adding dispersal to environmental variables was moderate and variable across invasive species (Figure 1).

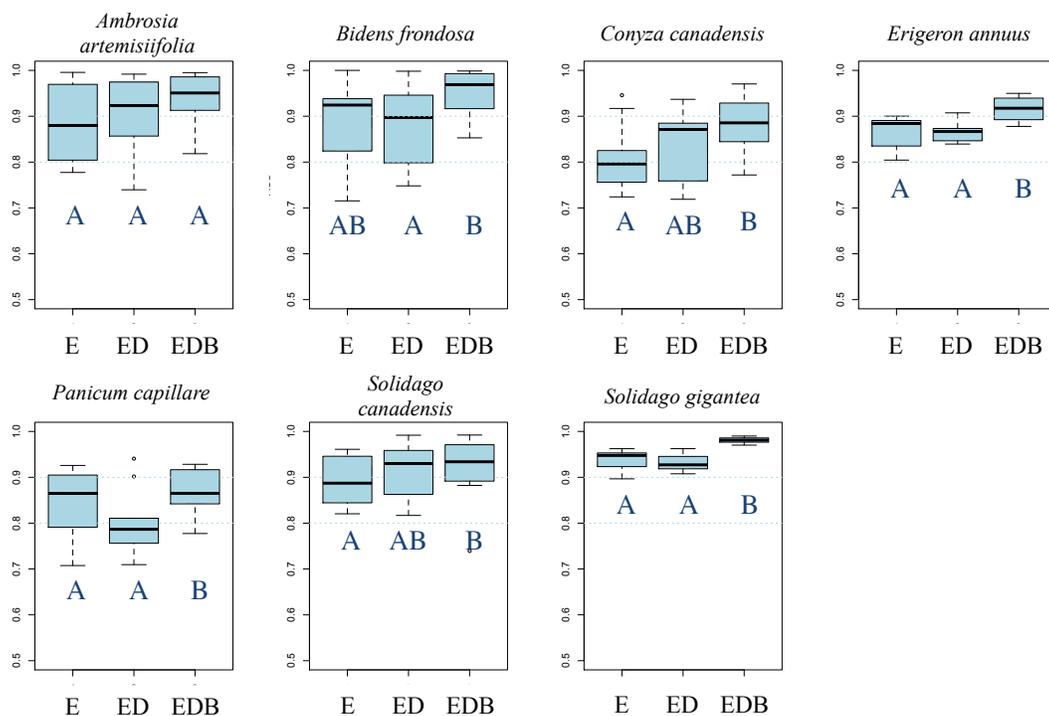


Figure 1. For each species, the boxplots represent the distribution of model accuracy (measured with AUC scores) when using (E) environmental variables only, (ED) environmental and dispersal variables, and (EDB) environmental, dispersal and biotic variables. The pairwise differences between the distributions were tested with Wilcoxon ranked test and letters represents the results.

## 2. Environment and spatial autocorrelation

Amongst the environmental variables, the most influential were land use type, human footprint and the sum of degree-days. The spatial autocovariate was important for all species except *Bidens frondosa* and *Panicum capillare*.

## 3. Biotic interaction indices

Amongst the biotic interaction indices, competitive ability measured by the height difference between the invader and all other herbaceous species was an important predictor of the presence and absence of invaders. Intriguingly, the invaders mostly occurred in communities with a similar average height, while they were generally predicted absent when they were much taller than the average height of the community. However, the two other indices of competitive ability on seed mass and SLA were not

very important in determining invader presence, except for *Panicum capillare* which tended to occur more in sites where the other species have a lower seed mass on average. Indices measuring niche differences showed that *Panicum capillare* and *Erigeron annuus* occurred more in sites where the other native species have similar traits (“clustering” of MDNC.h). *Conyza canadensis* and *Bidens frondosa* occurred in sites where there is at least one very similar species (clustering of MDNN). Interestingly, *Solidago canadensis* showed a clustering signal when considering only herbaceous native species (MDNC.h), but an overdispersion signal (i.e. a preference for being dissimilar) when considering all natives in the communities (MDNC.t). Finally the relationship between observed vs. potential species richness was important for *Erigeron annuus* that tended to occur in sites where the Rich<sub>obs/pot</sub> index was high.

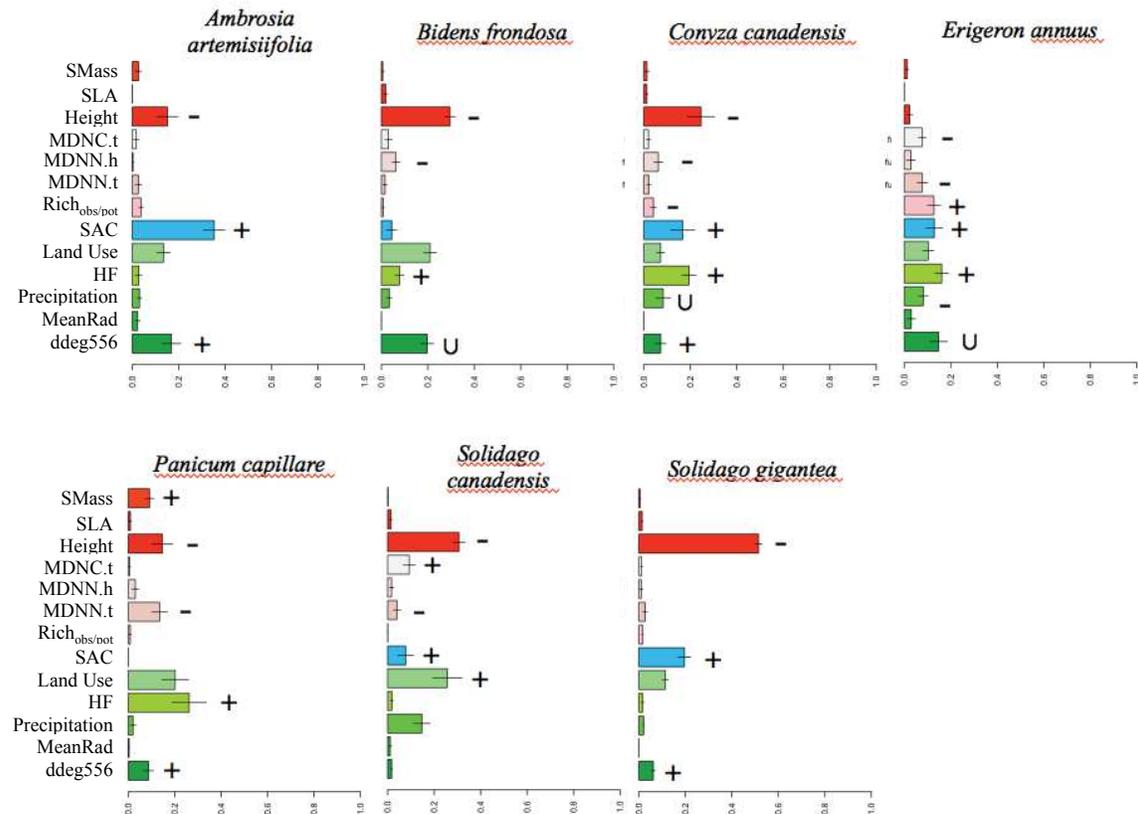


Figure 2. Relative variable importance in predicting invasion success for 7 invasive plants in the French Alps (averaged across 10 models per species). The variables were sorted by general categories of filters: environmental (in green), dispersal (in blue), biotic interaction indices (in red). The positive and negative symbols indicated the sign of the relationship between invader probability of presence and each variable (e.g. when a competitive ability for height index has a “-”, the invader was predicted to invade communities where it is of equal or smaller size than the native species), and when it is a “U” it means that the relationship is not linear.

#### 4. Interactions between environmental and biotic variables

Finally, we looked at the interaction patterns between the different variables, especially between the biotic indices and the environmental variables, to identify whether the position of the communities along climatic gradients could influence the effects of different indices of competitive interaction (Figure 3). Overall, it was clear that for every interaction between important variables (identified in Fig. 2), the biotic interaction indices were systematically more influential towards the core of the invader's niches (i.e. when the probability of invader presence was high in the abiotic gradient).

For instance, the effect of competitive ability of native communities, as captured by height differences with invaders, turned out to be more important when the abiotic conditions were not at the cold (e.g. *Ambrosia artemisiifolia*), warm (e.g. *Bidens frondosa*) or dry edge of the environmental gradients in the Alps (e.g. *Solidago canadensis*). Similarly, in sites with a greater human footprint which are therefore highly disturbed, the probability of occurrence of some invaders increased with their functional similarity to the native community, as

captured by both competitive ability and niche difference indices (e.g. *Panicum capillare*, *Solidago gigantea*). Note that the patterns were very similar across habitat types (data not shown). We have therefore only presented the response curves for the grassland habitats in Figure 3.

#### DISCUSSION

The general aim of our work was to identify the respective influence of the three potential processes driving invasion success in the French Alps: environmental filtering, dispersal, and competitive interactions. Using a palette of different variables and indices, we were able to refine our understanding of the relative influence of these processes and their interactions, which is of prime importance for improving predictive models of invasion success and thus facilitate management efforts. Overall, we show that including both dispersal and biotic interaction indices can significantly improve the predictive ability of traditional distribution models (Figure 1) but that the relative importance of the three types of variables may vary across species, thus requiring finely tuned model parameterisation.

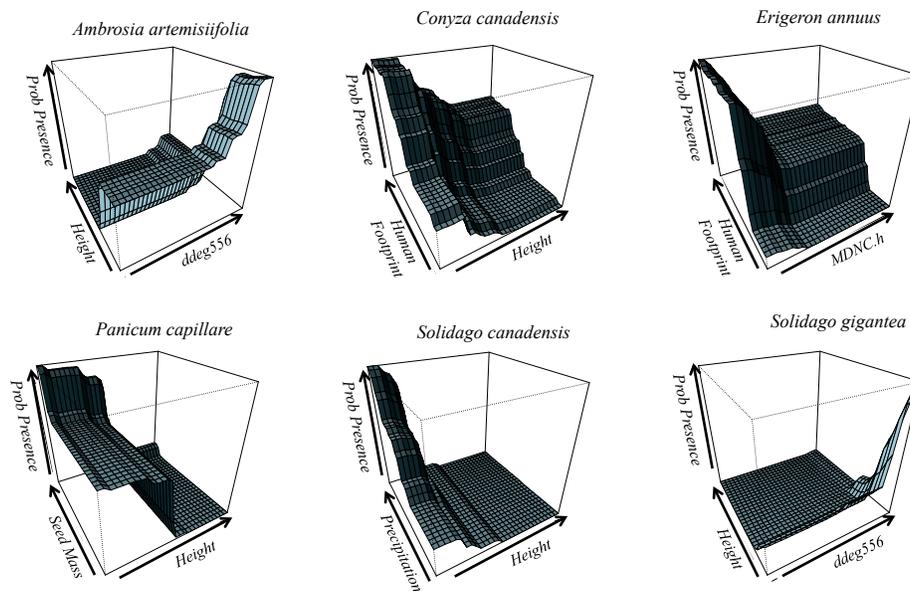


Figure 3. 3-dimensional response curves of invasion success relating to two important predictors identified in Fig. 2.

### Environmental variables

In general, the type of land use, the level of human footprint and the sum of degree-days were the most important variables for our studied herbaceous invaders. This outcome corresponds to the commonly identified abiotic drivers of invasion (e.g. Vila & Pujadas 2001; Pysek *et al.* 2010) and especially indicates that *Bidens frondosa*, *Conyza canadensis*, *Erigeron annuus* and *Panicum capillare* are currently at an establishment stage as they are greatly influenced by the level of human footprint. Indeed, the human footprint includes two components which are particularly important for alien species in the early stages of invasion. It is both an indicator of the intensity of site disturbance and a proxy for the level of propagule pressure (Sanderson *et al.* 2002). However, it is important to keep in mind that the level of human footprint in the alpine landscape is partly linked to the temperature gradient, as human populations tend to settle in lowland areas. Nonetheless, this variable is still highly predictive. Concerning the importance of the land use type, it is not surprising that species have particular habitat preferences, with some invaders preferring wetlands (e.g. *B. frondosa*) while others occur more in arable lands (e.g. *C. canadensis*, *E. annuus*) or grasslands (e.g. *P. capillare*). We would encourage its use in invasive species distribution models, as it is freely available at continent and global scales (e.g. CORINE 2006; [Global Land Cover, http://ionial.esrin.esa.int/index.asp](http://ionial.esrin.esa.int/index.asp)).

### Dispersal variable

Interestingly, the dispersal variable, estimated here in the simplest manner (with an autocovariate term), is shown to be quite an important predictor of the invader presence. It is particularly important for the species that have the highest seed production, as for example *A. artemisiifolia* (2,000 seeds per individual on average), *C. canadensis* (15,000 seeds per individuals on average) or *S. gigantean* (2,500 seeds per individuals on average; Douzet personal communication). Invaders producing lower seed numbers are less influenced by this dispersal variable, such as *B. frondosa* (500 seeds per individuals on average) or *P. capillare* (100 seeds per individuals on average; Douzet personal communication). This difference between species with high vs. low seed production can be explained by the fact that they are both involved in a colonisation process where the spatial autocorrelation is essentially driven by species fecundity (Levin *et al.* 2003). However, our measure of dispersal could also reflect a missing spatially autocorrelated abiotic or biotic driver operating at the same spatial scale (Legendre & Legendre, 2012). In any case, if this variable is shown to be an important predictor, either as a dispersal limitation indicator or

as a proxy for a missing variable, it is important to include it in the modelling process, especially for predictive purposes at local scales. Propagule pressure is certainly a major driver of plant invasions (Simberloff 2009), and dispersal processes, or spatial autocorrelation therefore need to be accounted for in predictive models. It is then important to parameterise such models with life history and ecological data in order to identify the species for which propagule pressure will have an important or unimportant effect on invasion dynamics.

### Competitive interaction variables

Among the different indices developed to test the two strategies of competitive interactions (niche opportunity vs. competitive ability), the indices of individual height differences were the most influential in determining invader presence (for all species except *E. annuus*). In all cases, height difference indicated that these invaders always co-occur with native herbaceous species of similar height but very rarely with native species of much lower height. It reveals high levels of filtering (abiotic or biotic) despite the use of several abiotic variables in the model. This conclusion is corroborated by the other index responses (except for *P. capillare* and *S. canadensis*, detailed below), although these are less important for model performance. For instance, MDNC<sub>herb</sub> (e.g. *E. annuus*, *P. capillare*) and MDNN indices (e.g. *B. frondosa*, *C. Canadensis*, *E. annuus*) show that invaders occur more in communities where they have the same traits as the native species. Furthermore, additional information can also be deduced for the two species that show positive correlations between some biotic variables and their probability of presence: *P. capillare* and *S. canadensis*. Firstly, for *S. canadensis* we can see that the species tends to occur in communities where it is similar to the native herbaceous species (negative relationship between its presence and MDNC<sub>herb</sub>) but at the same time where there are also woody species with very dissimilar traits (positive relationship with MDNC<sub>all</sub>). This result corroborates findings in the literature showing that this species particularly appreciates forest edges (Jakobs *et al.* 2004), probably because tree species modify the local abiotic environment toward more favourable conditions for this invader (e.g. less evapotranspiration, less solar radiation, etc.). This finding shows the importance of accurately defining the different organisms that can interact through competitive exclusion, in order to not misinterpret the patterns of dissimilarity. Secondly, *P. capillare* shows a preference for sites where its seed mass (up to 0.30 mg, Clements *et al.* 2003) is greater than for the native species. This difference is probably linked to the fact that this invader has an extremely persistent seed bank (up to 47 years; Clements *et al.* 2003), which gives it the advantage of growing in

sites where it was dormant, as soon as the conditions become favourable. In contrast, other species that have a lower seed mass, hence probably less seed dormancy are less likely to be already present in the site when the conditions become favourable. Finally, for *E. annuus*  $R_{\text{obs}/\text{pot}}$  was positively correlated with invader presence, indicating that the intensity of competition in the communities where it occurs is low.

Overall, our results showing that competitive interactions do not hamper invasions for our set of species in the French Alps contradict some studies (e.g. using phylogeny as a proxy for functional traits: Davies *et al.* 2011; Schaefer *et al.* 2011; Carboni *et al.* 2012) but corroborate others (e.g. Daehler 2001; Diez *et al.* 2009; Van Wilgen & Richardson 2011). We believe that this discrepancy reflects the fact that our seven herbaceous invaders are still in the establishment phase in the region which make any potential signal of biotic resistance in communities still quite low. In addition, the fact that we detected high levels of trait similarity (notably for individual height) between the invader and the native species indicates that an important variable was probably missing from the models. This variable could be abiotic, such as disturbance regime or fine-scale soil nutrient content (Lake & Leishman 2004), or it could be biotic, such as facilitation or shared pollinators (Morale & Traverest 2009). Finally, since most of the invaders had Mediterranean affinities, our alpine dataset did not contain their full range of environmental niches, and therefore could not capture the warmer edge of their geographic range. There was, however, a visible increase in the importance of biotic indices towards the centre of species' niches (for each temperature class, from cold to warm, the influence of the biotic interaction index increased, Figure 3). This result cannot corroborate our expectation that stressful environments decrease the relative importance of competition (Körner 1999; Callaway *et al.* 2002; Graham *et al.* 2009) as we did not detect any. Nevertheless it appears that the clustering pattern is more pronounced in suitable environments. If the clustering patterns were the results of facilitation processes, it would suggest that facilitation is more important towards niche core than towards niche boundaries, while if the clustering is the result of a hidden abiotic factor, this would indicate a simple interaction between these factors. Further investigation into the interplay between environmental stress and competitive interactions is necessary and promises to provide very interesting results.

#### Concluding remarks

Overall it seems that competitive interactions are not of vital importance to alien herbaceous species in the

French Alps as they are essentially at an establishment stage. Identifying that these species are not yet *stricto sensu* invasives is important as it indicates that management and eradication efforts remain feasible, while these invaders still have no apparent impact on native community compositions. Moreover, we have shown that proxies of biotic interactions are useful for detecting other important factors and consequently, we recommend constructing these proxies when developing distribution models for both invasive species and native species.

#### ACKNOWLEDGEMENTS

We would like to thank T. Münkemüller for the thought-provoking discussions we have had. The research leading to these results received funding from the European Research Council under the European Community's Seven Framework Programme FP7/2007-2013 Grant Agreement no. 281422 (TEEMBIO). WT and SL acknowledge support from the French "Agence Nationale de la Recherche" under the SCION project (ANR-08-PEXT-03).

#### REFERENCES

- Angert, A. L. et al. 2009. Functional tradeoffs determine species coexistence via the storage effect. — *Proceedings of the National Academy of Sciences of the United States of America* 106: 11641-11645.
- Augustin, N. H. et al. 1996. An autologistic model for the spatial distribution of wildlife. — *Journal of Applied Ecology* 33: 339-347.
- Benichou, P. and Le Breton, O. 1987. Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. — *La Météorologie* 7: 23-34.
- Boulangeat, I. et al. 2012a. Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. — *Ecology Letters* 15: 584-593.
- Boulangeat, I. et al. 2012b. Niche breadth, rarity and ecological characteristics within a regional flora spanning large environmental gradients. — *Journal of Biogeography* 39: 204-214.
- Breiman, L. 2001. Statistical modeling: The two cultures. — *Statistical Science* 16: 199-215.
- Callaway, R. M. et al. 2002. Positive interactions among alpine plants increase with stress. — *Nature* 417: 844-848.
- Callaway, R. M. and Walker, L. R. 1997. Competition and facilitation: A synthetic approach to interactions in plant communities. — *Ecology* 78: 1958-1965.

- Carboni, M. et al. 2012. Darwin's naturalization hypothesis: scale matters in coastal plant communities. — *Ecography* 35: 001-009.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. — *Annual Review of Ecology and Systematics* 31: 343-366.
- Clements, D. R. et al. 2004. The biology of Canadian weeds. 127. *Panicum capillare* L. — *Canadian Journal of Plant Science* 84: 327-341.
- Daehler, C. C. 2001. Darwin's naturalization hypothesis revisited. — *American Naturalist* 158: 324-330.
- Davies, K. F. et al. 2011. Native communities determine the identity of exotic invaders even at scales at which communities are unsaturated. — *Diversity and Distributions* 17: 35-42.
- De Roy, K. et al. 2013. Environmental conditions and community evenness determine the outcome of biological invasion. — *Nature communication*
- Diez, J. M. et al. 2009. Learning from failures: testing broad taxonomic hypotheses about plant naturalization. — *Ecology Letters* 12: 1174-1183.
- Dormann, C. F. 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. — *Global Ecology and Biogeography* 16: 129-138.
- Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. — *Ecography* 30: 609-628.
- Duncan, R. P. and Williams, P. A. 2002. Darwin's naturalization hypothesis challenged. — *Nature* 417: 608.
- Engler, R. et al. 2009. Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? — *Ecography* 32: 34-45.
- Fargione, J. et al. 2003. Community assembly and invasion: An experimental test of neutral versus niche processes. — *Proceedings of the National Academy of Sciences of the United States of America* 100: 8916-8920.
- Fargione, J. E. and Tilman, D. 2005. Diversity decreases invasion via both sampling and complementarity effects. — *Ecology Letters* 8: 604-611.
- Fitter, A. H. and Peat, H. J. 1994. The ecological flora database. — *Journal of Ecology* 82: 415-425.
- Gallego, F. J. 2010. A population density grid of the European Union. — *Population and Environment* 31: 460-473.
- Gallien, L. et al. 2010. Predicting potential distributions of invasive species: where to go from here? — *Diversity and Distributions* 16: 331-342.
- Graham, C. H. et al. 2009. Phylogenetic structure in tropical hummingbird communities. — *Proceedings of the National Academy of Sciences of the United States of America* 106: 19673-19678.
- Grime, J. P. 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. — *Journal of Ecology* 86: 902-910.
- Herben, T. 2005. Species pool size and invasibility of island communities: a null model of sampling effects. — *Ecology Letters* 8: 909-917.
- Higgins, S. I. et al. 2000. Using a dynamic landscape model for planning the management of alien plant invasions. — *Ecological Applications* 10: 1833-1848.
- HilleRisLambers, J. et al. 2012. Rethinking Community Assembly through the Lens of Coexistence Theory. — *Annu. Rev. Ecol. Evol. Syst.* 43: 227-248.
- Hothorn, T. et al. 2006. party: A Laboratory for Recursive Part(y)itioning. — <http://CRAN.R-project.org> R package version 0.9-96.:
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. — Princeton University Press.
- Jakobs, G. et al. 2004. Introduced plants of the invasive *Solidago gigantea* (Asteraceae) are larger and grow denser than conspecifics in the native range. — *Diversity and Distributions* 10: 11-19.
- Julve, P. 1998. Baseflor. Index botanique, écologique et chorologique de la flore de France.
- Kikvidze, Z. et al. 2011. Importance versus intensity of ecological effects: why context matters. — *Trends in Ecology & Evolution* 26: 383-388.
- Knevel, I. C. et al. 2003. Life-history traits of the northwest European flora: The LEDA database. — *Journal of Vegetation Science* 14: 611-614.
- Körner, C. 1999. *Alpine Plant Life*. — Springer-Verlag.
- Kühn, I. et al. 2004. BioFlor: a new plant-trait database as a tool for plant invasion ecology. — *Diversity and Distributions* 10: 363-365.
- Kunstler, G. et al. 2012. Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic or functional similarity: implications for forest community assembly. — *Ecology Letters* 15: 831-840.
- Lake, J. C. and Leishman, M. R. 2004. Invasion success of exotic plants in natural ecosystems: the role of disturbance, plant attributes and freedom from herbivores. — *Biological Conservation* 117: 215-226.
- Legendre, P. and Legendre, L. 1998. *Numerical Ecology*.
- Levin, S. A. et al. 2003. The ecology and evolution of

- seed dispersal: a theoretical perspective. — *Annu. Rev. Ecol. Evol. Syst.* 34: 575-604.
- Levine, J. M. and D'Antonio, C. M. 1999. Elton revisited: a review of evidence linking diversity and invasibility. — *Oikos* 87: 15-26.
- Lonsdale, W. M. 1999. Global patterns of plant invasions and the concept of invasibility. — *Ecology* 80: 1522-1536.
- Mayfield, M. M. and Levine, J. M. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. — *Ecology Letters* 13: 1085-1093.
- Mitchell, C. E. et al. 2006. Biotic interactions and plant invasions. — *Ecology Letters* 9: 726-740.
- Morales, C. L. and Traveset, A. 2009. A meta-analysis of impacts of alien vs. native plants on pollinator visitation and reproductive success of co-flowering native plants. — *Ecology Letters* 2009:
- Proches, S. et al. 2008. Searching for phylogenetic pattern in biological invasions. — *Global Ecology and Biogeography* 17: 5-10.
- Pysek, P. et al. 2010. Disentangling the role of environmental and human pressures on biological invasions across Europe. — *Proceedings of the National Academy of Sciences of the United States of America* 107: 12157-12162.
- R Development Core Team (ed.) 2012. R: A Language and Environment for Statistical Computing. — R Foundation for Statistical Computing, Vienna, Austria. Available online at <http://www.R-project.org>.
- Rejmánek, M. et al. 2005. Ecology of invasive plants: state of the art. — In: Mooney, H. A. et al. (eds), *Invasive alien species: a new synthesis*. pp. 104-161.
- Richardson, D. M. and Pyšek, P. 2006. Plant invasions: merging the concepts of species invasiveness and community invasibility. — *Progress in Physical Geography* 30: 409-431.
- Roquet, C. et al. in review. Building megaphylogenies for macroecology: taking up the challenge. — *Ecography*
- Roura-Pascual, N. et al. 2009. From introduction to equilibrium: reconstructing the invasive pathways of the Argentine ant in a Mediterranean region. — *Global Change Biology* 15: 2101-2115.
- Sanderson, M. J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. — *Bioinformatics* 19: 301-302.
- Schaefer, H. et al. 2011. Testing Darwin's naturalization hypothesis in the Azores. — *Ecology Letters* 14: 389-396.
- Schuettpelz, E. and Pryer, K. M. 2009. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. — *Proceedings of the National Academy of Sciences of the United States of America* 106: 11200-11205.
- Seastedt, T. R. and Pysek, P. 2011. Mechanisms of Plant Invasions of North America and European Grasslands. — In: Futuyma, D. J. et al. (eds), *Annual Review of Ecology, Evolution, and Systematics*, Vol 42. pp. 133-153.
- Shea, K. and Chesson, P. 2002. Community ecology theory as a framework for biological invasions. — *Trends in Ecology and Evolution* 17: 170-176.
- Simberloff, D. 2006. Invasional meltdown 6 years later: important phenomenon, unfortunate metaphor, or both? — *Ecology Letters* 9: 912-919.
- Simberloff, D. 2009. The Role of Propagule Pressure in Biological Invasions. *Annual Review of Ecology Evolution and Systematics*. *Annual Reviews*, pp. 81-102.
- Smith, S. A. et al. 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. — *Proceedings of the National Academy of Sciences of the United States of America* 107: 5897-5902.
- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. — *Bioinformatics* 22: 2688-2690.
- Stohlgren, T. J. et al. 2003. The rich get richer: patterns of plant invasions in the United States. — *Frontiers in Ecology and Environment* 1: 11-14.
- Strobl, C. et al. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. — *Bmc Bioinformatics* 8: 25.
- Suding, K. N. et al. 2005. Functional- and abundance-based mechanisms explain diversity loss due to N fertilization. — *Proceedings of the National Academy of Sciences of the United States of America* 102: 4387-4392.
- Swenson, N. G. 2006. Gis-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. — *Journal of Evolutionary Biology* 19: 717-725.
- Swets, K. A. 1988. Measuring the accuracy of diagnostic systems. — *Science* 240: 1285-1293.
- Theoharides, K. A. and Dukes, J. S. 2007. Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. — *New Phytologist* 176: 256-273.
- Thuiller, W. et al. 2010. Resolving Darwin's naturalization conundrum: a quest for evidence. — *Diversity and Distributions* 16: 461-475.

- Thuiller, W. et al. 2012. Ecological niche and species traits: key drivers of regional plant invader assemblages. — *Biological Invasions* 14: 1963-1980.
- Thuiller, W. et al. 2011. Consequences of climate change on the tree of life in Europe. — *Nature* 470: 531-534.
- Thuiller, W. et al. 2007. Stochastic species turnover and stable coexistence in a fire-prone plant community. — *PloS One* 2: e938. doi:10.1371/journal.pone.0000938.
- van Kleunen, M. et al. 2011. Preadapted for invasiveness: do species traits or their plastic response to shading differ between invasive and non-invasive plant species in their native range? — *Journal of Biogeography* 38: 1294-1304.
- van Wilgen, N. J. and Richardson, D. M. 2011. Is phylogenetic relatedness to native species important for the establishment of reptiles introduced to California and Florida? — *Diversity and Distributions* 17: 172-181.
- Vila, M. and Pujadas, J. 2001. Land-use and socio-economic correlates of plant invasions in European and North African countries. — *Biological Conservation* 100: 397-401.
- Vittoz, P. and Engler, R. 2007. Seed dispersal distances: a typology based on dispersal modes and plant traits. — *Botanica Helvetica* 117: 109-124.
- Von Holle, B. and Simberloff, D. 2005. Ecological resistance to biological invasion overwhelmed by propagule pressure. — *Ecology* 86: 3212-3218.
- Webb, C. O. et al. 2002. Phylogenies and community ecology. — *Annual Review of Ecology and Systematics* 33: 475-505.
- Welden, C. W. and Slauson, W. L. 1986. The intensity of competition versus its importance - an overlooked distinction and some implications. — *Quarterly Review of Biology* 61: 23-44.
- Westoby, M. 1998. A leaf-height-seed (LHS) plant ecology strategy scheme. — *Plant and Soil* 199: 213-227.
- Westoby, M. et al. 2002. Plant ecological strategies: Some Leading Dimensions of Variation Between Species. — *Annu. Rev. Ecol. Syst.* 33: 125-159.
- Westoby, M. and Wright, I. J. 2006. Land-plant ecology on the basis of functional traits. — *Trends in Ecology & Evolution* 21: 261-268.
- Wisz, M. S. et al. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. — *Biological reviews of the Cambridge Philosophical Society* 88: 15-30.
- Wright, I. J. et al. 2004. The worldwide leaf economics spectrum. — *Nature* 428: 821-827.



## *Chapter 5*



EVIDENCE FOR RAPID – GENETICALLY-BASED –  
CLIMATIC NICHE EXPANSION IN AN INVASIVE  
WEED: *AMBROSIA ARTEMISIIFOLIA* L.

*PREUVE DE L'EXPANSION RAPIDE DE LA NICHE  
CLIMATIQUE – SUR DES BASES GENETIQUES – CHEZ  
UNE MAUVAISE HERBE INVASIVE: AMBROSIA  
ARTEMISIIFOLIA L.*



## PRESENTATION DU CHAPITRE

*Est ce que les espèces invasives changent leurs niches climatiques durant leur colonisation? Cette question est au cœur d'une forte controverse. En effet, plusieurs témoignages ont montré soit de fortes différences climatiques entre les régions occupées de l'espace natif et de l'espace envahi des invasives, soit un fort conservatisme des niches quelque soient les régions. Dans ce chapitre nous apportons le premier exemple d'expansion de la niche climatique d'une espèce invasive dans la région envahie sur des bases génétiques. Pour ce faire, nous avons étudié l'ambrosie à feuille d'armoise (*Ambrosia artemisiifolia* L.), une espèce de forte préoccupation pour la santé humaine à cause de son pollen fortement allergisant, invasive dans les alpes françaises. Dans un premier temps, nous avons estimé la niche de l'invasive à la fois à l'échelle globale et à l'échelle régionale, et nous les avons utilisé pour stratifier la collecte de données génétiques et de graines (ensuite semées pour une expérience de jardin commun). Ensuite, nous avons relié la structure génétique de l'espèce au long de sa niche avec les variations de ses traits impliqués dans l'adaptation au climat. Ce faisant, nous avons put montrer que cette espèce n'est pas seulement en train de s'adapter localement aux conditions climatiques, mais qu'elle est de surcroît en train d'étendre sa niche vers les conditions plus froides des alpes françaises.*



# Evidence for rapid, genetically-based, climatic niche expansion in an invasive (annual) weed

Laure Gallien<sup>1</sup>, Wilfried Thuiller<sup>1</sup>, Noémie Fort<sup>2</sup>, Marti Boleda<sup>1</sup>, Florian Alberto<sup>1</sup>, Delphine Rioux<sup>1</sup>, Juliette Lainé<sup>1</sup> & Sébastien Lavergne<sup>1</sup>.

<sup>1</sup> *Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, Grenoble I, BP 53, 38041 Grenoble Cedex 9, France.*

<sup>2</sup> *Conservatoire Botanique National Alpin, Domaine de Charance, 05000 Gap, France.*

## ABSTRACT

Whether invasive species shift their climatic niche when colonizing a new area remains a challenging issue with conflicting results. Here, we provide the first example of a genetic-based, climatic niche expansion of an invader in its adventive range. Our study is based on the common ragweed (*Ambrosia artemisiifolia* L.), a species of great concern for human health due to its allergenic pollen, invasive in Europe and particularly in the French Alps. We estimated the invader's niche at both global and regional scales to stratify the sampling design of genetic and seed sample collection (sown in a common garden experiment). By linking the species' genetic structure over the niche and the traits involved in adaptation, we show that the species is adapting to local conditions and currently expands its niche toward colder climates in the French Alps. Such results should pave the way for a better eco-evolutionary understanding of biological invasions.

*Keywords:* local adaptation, *Ambrosia artemisiifolia* L., climatic niche expansion,  $F_{ST}$ ,  $G$  matrix, invasion, population genetics,  $Q_{ST}$ .

## INTRODUCTION

Biological invasions, a major component of global changes, alter the structure of native communities and can disturb ecosystem functioning and associated goods and services (Walther *et al.* 2009). The ever increasing spread of invasive species have thus stimulated an important body of research, leading to significant discoveries about the fundamental ecological mechanisms shaping species ranges, community structure, or food-web

stability (Sax *et al.* 2007). Among the long-lasting issues that can be investigated with invasive species are the understanding of the niche dynamic drivers, and ultimately the investigation of niche conservatism vs. niche lability theories (Alexander & Edwards 2010; Lavergne *et al.* 2010). Indeed, it becomes of increasing interest to understand how evolution shapes species' climatic niches, as it will shed light on the determinants of invasions under current and future climate, which is particularly important for developing forecasting tools for biological invasions (Gallien *et al.* 2010). It is however, not clear whether climatic niches of invasive species can be considered as fixed species features or whether they can experience rapid evolution, as shown for other species characteristics (e.g. Dlugosch & Parker 2008 for life history traits).

Climatic niche shifts have been demonstrated in a number of invasive species, by comparing the native and adventive climatic conditions in which they occur (e.g. Broennimann *et al.* 2007; Gallagher *et al.* 2010). These shifts represent changes in the realized climatic niches of the species, and they may not necessary be driven by adaptive changes in climatic affinities (Gallien *et al.* 2012). Observed shifts between native and invasive climate niches may only reflect sorting of populations or lineages that are pre-adapted to particular climatic conditions (Lachmuth *et al.* 2010). Alternatively, some introduced plant species have quickly evolved clines of local adaptation along climatic and topographic gradients (Alexander & Edwards 2010). However, local adaptation and clinal differentiation can differ in their strength between species or regions (Alexander & Edwards 2010), suggesting that niche expansions towards marginal climatic

conditions are much dependent on introduction history, genetic make up of introduced populations, gene flows and relevant trait variation.

The niche expansion of an invasive species will necessarily happen upon (1) *genetic innovation* that is the make-up of new genotypes, and (2) *functional evolvability* that is the genetically-based variation in phenotypic traits with adaptive significance.

*Genetic innovation* leading to adaptation may be enhanced by repeated introductions, important propagule pressure and gene flows, besides the overcome of population bottlenecks (Holt 2005). Repeated introductions followed by post-introduction admixture have the potential for increasing genetic variation in selective traits and/or increase individual fitness through hybrid advantage (Facon *et al.* 2006). However, it remains unknown whether such increased adaptive potential or fitness occur are the forefront of climatic niche limits. Increased propagule pressure and gene flows may, nonetheless, also provoke gene swamping in climatically marginal populations and prevent them from developing local adaptations (Kawecki 2008; Sexton *et al.* 2009). Important gene flows may also increase the probability of allele surfing and create the possibility for spurious allele-environment relationships, which would be wrongly attributed to local adaptations (Lachmuth *et al.* 2010).

*Functional evolvability* involved in local adaptation depends on the level of genetic variation for the traits involved, their genetic covariance, and the direction of the selection pressure. When several traits are considered together, the number of combination that can respond to selection can be much smaller than the number of traits, limiting or preventing the evolution of an optimal combination of traits (Pigliucci & Kaplan 2000; Stepan *et al.* 2002). Consequently, to understand climatic niche expansion it is fundamental to contrast adaptive from non-adaptive genetic variation to tease apart the effects of adaptive processes from random changes due to the colonization process, but also to validate genome scan detection of allele-environment relationships with a multivariate framework of quantitative traits analysis.

Here we propose a novel integrative approach that combines climatic niche models at different geographic scales, population genetics and

common garden experiment to identify and localize adaptive evolution in the common ragweed (*Ambrosia artemisiifolia* L., Asteraceae) in the French Alps region. We first inferred the climatic niche of the invader in the French Alps (*regional niche*), and its global climatic niche that combines all native and introduced range to get close to the species' physiological limits (*global niche*; Gallien *et al.* 2010). This allows identifying the populations in the French Alps that occur outside of the global climatic niche and that could be suspected of adaptation toward niche expansion (as shown in Gallien *et al.* 2012). The regional niche estimates is used to stratify the sampling of genetic data and trait collection (Albert *et al.* 2010), notably in identifying populations at the most extreme climatic niche edges. Second, we study the determinants of the invader genetic structure across the climatic space to infer both the neutral genetic structure due to dispersal and drift processes, and the genetic signatures of natural selection across different climatic environments. Third, we measure phenotypic traits in a common garden to relate them to molecular signatures of selection (based on seeds collected across the whole regional niche). These traits and their variation across the populations are linked to the genetic structures and provide hints on what consist this adaptation. Then, the analysis of genetic traits covariances (**G** matrices; Cheverud 1996) defines whether the combination of traits responds independently to selection. If the traits are not independent, then a test can be applied to identify whether adaptation will be able to pursue according to trade-offs between the multiple trait (selection skewers; Calsbeek & Goodnight 2009).

Using this methodology, we answer (1) whether the suspected niche expansion toward cold mountainous conditions is confirmed? And (2) whether this niche expansion can pursue toward more extreme conditions?

## MATERIAL & METHOD

### 1. Study region and species

The common ragweed (*Ambrosia artemisiifolia* L., Asteraceae) is a North American annual weed that was transported with seed crops and forages about 150 years ago, into several distinct locations across Eurasia, Australia and South America, is continuously spreading since then (Chauvel *et al.* 2006). It is a pioneer species of open semi-arid

ruderal habitats (e.g. road sides, cultivated fields, or river banks, Essl *et al.* 2009), wind-pollinated and described as self-incompatible, although recent population genetic evidence suggests that it may have shifted towards inbreeding during its colonization process (Gaudeul *et al.* 2011). Dispersal is essentially human-assisted through crop trade, soil transportation or via agricultural machines, as well as flowing water as the achenes are able to float (Fumanal *et al.* 2007). Interestingly, this species usually present at low elevations in its native range has been recorded at unexpectedly high elevations in the French Alps (National Botanical Conservatory of the Alps, CBNA, unpublished data, Mont Genève, 1600 m.a.s.l. in 2009), suggesting an adaptation to cold conditions (see also Gallien *et al.* 2012).

The French Alps represent an approximate area of 60,000 km<sup>2</sup> of highly fragmented and heterogeneous habitats over wide environmental gradients. Mean annual temperatures and sum of precipitations range from -6.91°C to +15.04°C and from 462mm to 2,895mm respectively.

## 2. Climatic niche estimation

Both the global and the regional niches (i.e. French Alps) were estimated. Result for the global niche was extracted from a former work (Gallien *et al.* 2012). To summarize, it was based on an ensemble forecast (Marmion *et al.* 2009) using 4,803 occurrence of *Ambrosia artemisiifolia* over the world from the Global Biodiversity Information Facility (GBIF; <http://data.gbif.org>) at a minimum resolution of 2.5' (c. 4.5 km), and 20,000 background data from realistically reachable locations (within a buffer zone of 20 km around any presence record). Five uncorrelated climatic variables were extracted from WorldClim (Hijmans *et al.* 2005; <http://www.worldclim.org>): maximal temperature in the warmest month, annual temperature range, mean temperature in the coldest quarter, precipitation in the wettest month, and precipitation in the driest month.

The regional niche for the French Alps was build using 3800 observations of *A. artemisiifolia* (source CBNA). To stratify our sampling design, we selected the 2 variables that most segregated the sites occupied by the species with a Principal Component Analysis (PCA) run over 32 pedo-topo-climatic variables. These variables were obtained from the meteorological model Aurelhy, based on interpolated measurements at a resolution of 100 m summarizing climatic information over

1971 and 2000 (Bénichou & Le Breton 1987). The selected gradients were: the mean solar radiation and the mean summer temperature (highly correlated with first two axes of the PCA and explaining 70% of the inter-sites differences).

## 3. Measuring the genetic structure across the niche

### *3.1. Molecular genetic data*

During summer 2010 we collected leaf samples within 27 populations selected to be representative of the regional niche (see Fig. 1). Genomic DNA was extracted using DNeasy 96 Plant kits (Qiagen) according to the manufacturer's instructions for 10 individuals in each of the 27 populations. Three blanks and 12 repetitions (~5%) were included to further assess the repeatability and reliability of the markers. We used an Amplified Fragment Length Polymorphism (AFLP) protocol based on standard methods (Vos *et al.* 1995) with minor modifications (see Appendix 1 for more details). Allele scoring was performed first via PeakScanner© v2.0 (Applied Biosystems) then RawGeno (Arrigo *et al.* 2009) with a visual checking of the bins, and manually using R v.2.15.0 software (R Development Core Team 2012): picks occurring in the blank samples were removed, the mismatch error rate acceptance was of 0% (and the reproducibility was high: 95%), pick size outliers were removed, as well as singleton picks and ubiquitous ones. In the end, from the initial set of 830 AFLP markers only 240 were kept for analysis (29% of the initial set).

### *3.2. Identify putative markers linked to regions under selection*

We first tested whether there were loci that could potentially be under selection, or linked to regions under selection, related to the two main niche gradients using logistic regressions (Manel *et al.* 2012 MolEcol). For each locus we tested for a significant association (linear and/or quadratic) with the two regional niche variables and their interactions using a stepwise procedure based on AIC. For the selected markers, we tested whether the residuals were spatially-autocorrelated using Moran's I (R package *spdep*; Cliff & Ord 1973) but none of was significant.

### *3.3. Description of the neutral genetic structure*

To analyse the neutral genetic structure, we removed the putative markers under selection as

they will bias the spatial neutral genetic patterns and blur the fingerprint of neutral processes and dispersal. The remaining AFLP makers (204 markers) were used to analyse the distribution of the neutral genetic diversity among populations.

We first estimated three population-specific genetic parameters:  $F_{IS}$ ,  $F_{ST}$  and  $H_e$  (Wright 1951) using well-validated approaches (see Appendix 1). The relationship between  $F_{IS}$ ,  $F_{ST}$  and  $H_e$  and the two regional niche gradients (temperature and radiation) were then tested using linear regression models (allowing for linear and/or quadratic relationships) within a stepwise AIC procedure. To investigate the genetic structure of the populations we used the Bayesian clustering program Structure 2.3 (Prichard *et al.* 2000) from which we extracted the optimal number clusters.

We also tested genetic isolation by geographic  $IBD_{geo}$  and environmental distances  $IBD_{env}$  with two AFLP marker pools: the full set of markers, and the reduced set of markers after removal of the ones putatively under selection (see the above section 3.1). We used mantel tests with 999 randomizations (R package *adeigenet*; Jombart 2008) under the expectation that the signal of  $IBD_{env}$  would be lost when markers under selection following the gradients are removed.

#### 4. Measuring the phenotypic structure across the niche

##### *4.1. Common garden*

Among the 27 populations used for the genetic analyses, we chose 18 that were not mown and the most distant in the climatic space. In each

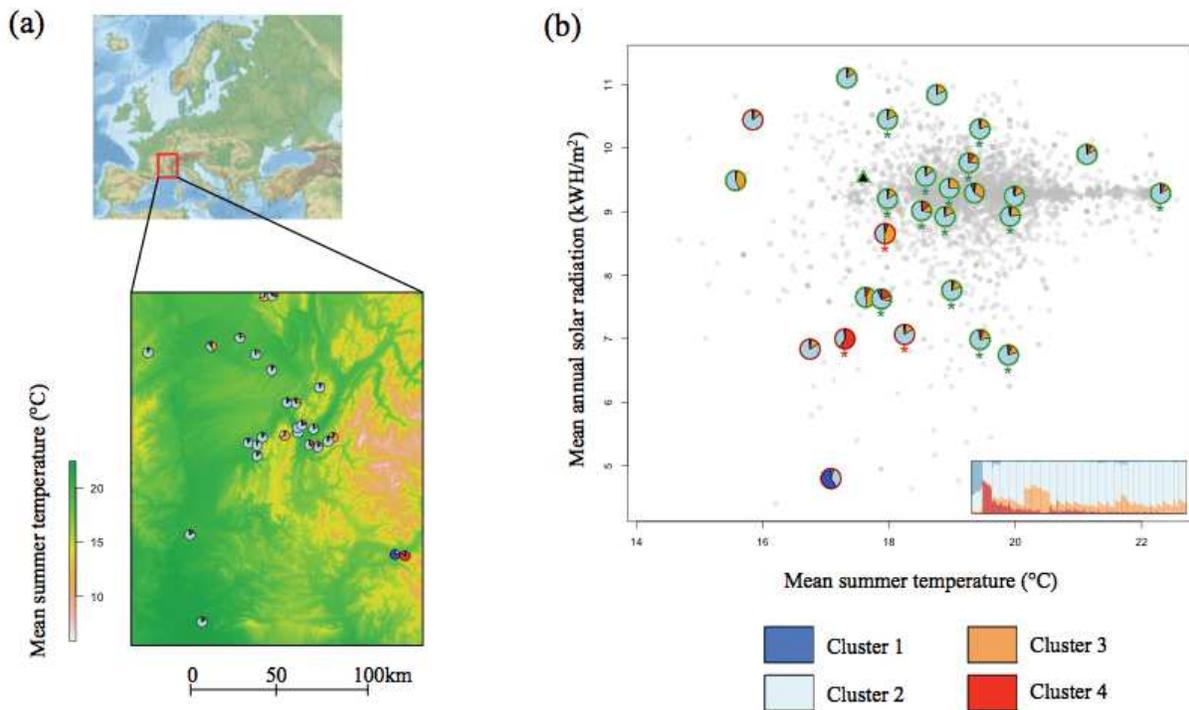


Figure 1. (a) Location of the 27 sampled populations for the genetic analysis in the geographic space, over a map of the mean summer temperature. Each population is represented with a pie chart showing the average proportion of the genetic cluster present in the populations. (b) Location of the 27 sampled population for the genetic analysis in the regional niche space, indicated with the pie charts. The grey dots indicate the position of the 3,800 populations of *A. artemisiifolia* recorded in the French Alps, used to estimate the regional niche of the species. When the pie chart is circled of green it means that the population is situated within the global niche of the species, and when it is of red it means that the population is located outside of the global niche. A star indicates when the population was also used in the common garden experiment, and a black triangle represents the location of the common garden experiment. Note that the figure in the bottom left panel present the probability of belonging to each cluster for every individual.

population, 30 seeds from 6 randomly chosen mother plants were collected in September-October 2010, dried at room temperature and stratified at 4°C during one month. Then we conducted a random block designed common garden experiment in Gap (French Alps; see climatic position in the Fig. 1 and Appendix 1 for the protocol).

#### 4.2 Trait measurement

We measure four quantitative traits: the total dry biomass (above + below ground biomass), the plant height, the shoot-root dry biomass ratio, and the leaf dry matter content (LDMC, ratio of fresh to dry biomasses). The biomass and height were chosen as proxies for plant fitness, while shoot-root ratio and LDMC were expected to represent stress resistance for soil content (e.g. water, nutrients) and climatic stresses (e.g. freezing temperature, the level of solar radiation) respectively (Violle *et al.* 2007). We measured the initial height of the seedlings from the ground to the top, and then recorded it every two weeks during two months. We stopped the measurements before blooming to avoid pollen spread. Given the size of the experiments, only half of the plants were randomly chosen to be measured (a total of 1134 plants). We then collected separately five mature leaves closest to the top, the shoot and the roots after washing in water pots to measure their biomasses. Leaves were weighted immediately to obtain their fresh biomass. These three plant compartments were then dried separately for 48h in an oven at 60°C, to weight their dry biomasses.

#### 4.3. Quantitative genetic analysis

The block design effect was tested using a likelihood ratio test (with Restricted Maximum Likelihood and a bootstrap test) between the full (Eq. 1) and all possible reduced models (R package *lme4*). The random effects of population and family provenance were always included; see Appendix 2). When significant, block-design was included in the final model.

$$\text{Trait} \sim 1 + 1|\text{Population} + 1|\text{Population/Family} + 1|\text{Block}$$

##### 4.3.1. Estimating $Q_{ST}$ and $CV_G$

We estimated  $Q_{ST}$  from phenotypic variance components among populations and families, estimated from the Eq. 1.  $Q_{ST}$  help identifying traits under selection when compared to  $F_{ST}$ . When

the traits are not under selection:  $Q_{ST} = F_{ST}$ , when there is stabilizing selection:  $Q_{ST} < F_{ST}$ , and when there is divergent selection  $Q_{ST} > F_{ST}$  (Merila & Crnokrak 2001). In a common garden the phenotypic variance equals the genetic variance plus the maternal effects that we assumed here negligible or at least constant across populations and families. The genetic variance component ( $VG$ ) can then be decomposed into an additive ( $VG_{add}$ : the additive effect of the transmitted alleles) and a “non-additive” component (including dominance and interaction effects). The general  $Q_{ST}$  of the species can then be estimated as  $Q_{ST} = V_{pop} / (V_{pop} + 2 * VG_{add})$ , where  $V_{pop}$  is the trait variance inter-population, and  $VG_{add} = N * V_{fam}$ , where  $V_{fam}$  is the trait variance inter-family within the population and  $N$  is the degree of kinship between individuals of the same family (here individuals of the same family were estimated has half-sibs,  $N=4$ ). Confidence intervals were calculated with 99 bootstraps of families. We calculated the genetic coefficient of variation  $CV_G$  ( $CV_G = VG_{add} / \text{trait mean}$ ; Houle 1992) to provide a relative measure of evolvability in describing how much phenotypic change may occur given a unit of selection. We could not estimate  $CV_G$  per population as  $V_{fam}$  was not accurately estimated within each population, but we divided the populations into two groups: *out of the global niche in particularly cold and shadow conditions* vs. *the rest of the populations* (with arbitrary cuts in the temperature [18°C] and radiation [9 kWh/m<sup>2</sup>] gradients based on visual inspections, see Fig. 1). For these two groups we calculated  $CV_G$ .

##### 4.3.3. Trait trends across the regional niche

We tested whether the traits changed over the regional niche of the species using generalized linear mixed models (R package *nlme*; Pinheiro *et al.* 2013). Both temperature and radiation were included in both simple and quadratic forms as fixed effects as well as their interactions (to allow for a quadratic response of the trait in the niche). Each combination of fixed effect parameters was tested and for each trait we retained the model with the lower AIC scores and tested whether such models were different from an intercept model with ANOVA (Figure 4; Appendix 3).

#### 4.4. Multi-dimensional quantitative genetic variation

To estimate whether the genetic correlations limit the ability of some populations to respond to

selection, we studied in each population the “lines of least resistance” in the additive genetic variance covariance matrix  $\mathbf{G}$  (Schluter 1996). From  $\mathbf{G}$ , we identified three major descriptors: (1) the *total genetic variance* measured as the sum of all trait variance (i.e. the trace of  $\mathbf{G}$ ), the *genetic correlation strength* taken as the proportion of the total genetic variance due to  $P_{max}$  (e.g. the largest eigenvector), and the *direction of the genetic correlations* (‘line of least resistance’; Cheverud 1996; Kirkpatrick 2009) estimated as  $P_{max}$  direction.  $P_{max}$ ’s direction is particularly interesting for comparing different populations, where the difference between populations is estimated as one minus the correlation between their  $P_{max}$  (i.e. the cosine between two  $P_{max}$ ). For the first two indices, we tested if they were varying along the niche gradients, using generalized linear models. For the direction of  $P_{max}$ , we tested whether the angle between each pair of populations was linked to their climatic differences, i.e. populations in similar conditions presented similar  $P_{max}$  directions, using a correlogram (including 9999 matrix randomizations and Bonferroni correction for multiple testing with the R package *vegan*; Oksanen *et al.* 2012). Note that all traits were normalized so that they had an equal weight in the analysis. We also took into account the fact that all individuals belonging to one family are not independent, by building  $\mathbf{G}$  matrices with MANOVAs as they can incorporate the family

effect (code modified from Martin *et al.* 2008).

We finally applied the sectioned skewers method (hereafter called SSM; Calsbeek & Goodnight 2009) to estimate the potential of evolution of the traits at the niche edges, based on the shape of the trait relationship with the niche gradients. SSM uses the breeder’s equation to estimate the response to selection of a  $\mathbf{G}$  matrix. The SSM is an extension of the random skewers method (Revell 2007), which compares the response of different  $\mathbf{G}$  matrices to a selection vector (with a given type and intensity of selection) to test whether they will result in alternative evolutionary responses. Our goal here was to test whether populations at the edge of niche could pursue adaptation toward colder conditions. We applied one selection scenario for the populations occurring in sites with cold temperatures and low levels of solar radiation, as they correspond to the populations that we detect as potentially adapting out of the species global niche. This scenario tested the population responses to selection toward colder and lower levels of solar radiation conditions, and was estimated as the coefficient of each trait-temperature regression (using standardized traits for adequate comparison). The multivariate response to selection is specified by the multivariate equivalent of the breeder’s equation:  $z = \mathbf{G}\beta$  where  $z$  is the vector of mean trait response to selection,  $\beta$  is the vector of selection gradients.

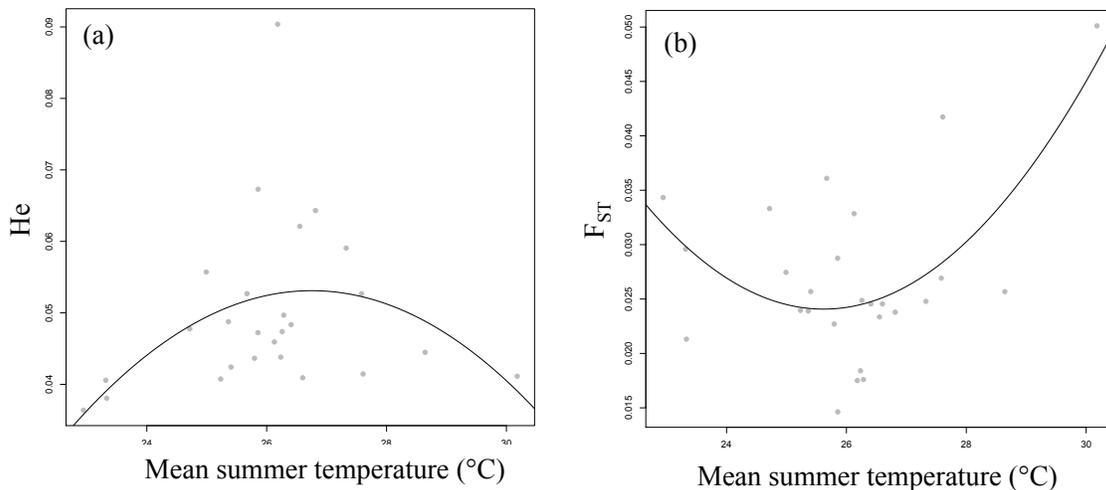


Figure 2. Neutral genetic differences of the 27 sampled populations represented on one of the regional niche axis: mean summer temperature. (a) Shows the population genetic diversity, while (b) shows the population’  $F_{ST}$  (here estimated as the level of genetic differentiation from an ancestry population).

	Biomass	Height	Shoot/Root	LDMC
$CV_G$ total	16.9	63.9	2.9	0.01
$CV_G$ in cold-shade conditions	3.19	33.3	3.3	0.01
$CV_G$ out cold-shade conditions	20.8	76.0	2.8	0.01
$Q_{ST}$ [ <i>half-sibs</i> – <i>full-sibs</i> ]( $\pm$ sd)	0.074 (0.012) – 0.138 (0.021)	0.078 (0.010) – 0.144 (0.018)	0.069 (0.011) – 0.130 (0.018)	0.207 (0.071) – 0.343 (0.095)

**Table 1.** Estimation of the genetic coefficient of variation ( $CV_G$ ) for the four traits considered in this study, considering all population together, only the populations in conditions of cold temperatures and low levels of solar radiation “in cold-shade”, and only the populations “out” of the cold-shade conditions. The  $Q_{ST}$  is also estimated for the four traits, considering either that the individuals of the same family are half- or full-sibs.

#### 4.5. Linking the functions to niche and genetics

To estimate whether the functional traits could be linked with some of the 36 putative markers under selection (identified in section 3.1.) we calculated the spearman rank correlation coefficient between each trait predicted values per population (estimated in section 4.3.2.) and the predicted allelic frequency in the same populations (estimated in section 3.1.).

## RESULTS

### 1. Genetic structure across the niche

#### Identification of putative markers under selection

Among the 240 AFLP markers, we identified 36 (15%) that could potentially be under selection. These markers had at least one significant relationship with temperature or radiation (linear or quadratic). As we did not detect significant  $IBD_{geo}$  and  $IBD_{env}$  on the overall dataset (see Appendix 4) we considered that these relationships reflect natural selection. These markers were thus removed from the analyses on the neutral genetic structure (204 neutral markers remaining).

#### Description of the neutral genetic structure

The general level of genetic differentiation among the 27 populations was low but significant ( $F_{ST}=0.0214$ ;  $p\text{-val}<0.001$ ). Genetic diversity  $H_e$  was maximal at the core of the temperature gradient and reduced at the two edges (Fig. 2a), varying between 0.035 and 0.09. This corroborated with higher population-specific  $F_{ST}$  values at the edges of the niche, ranging between 0.015 and 0.05 ( $p\text{-val}=0.030$ ; Fig. 2b), but not significantly correlated with *in situ* population size. Moreover,  $F_{IS}$  was generally high (0.441 on average, ranging

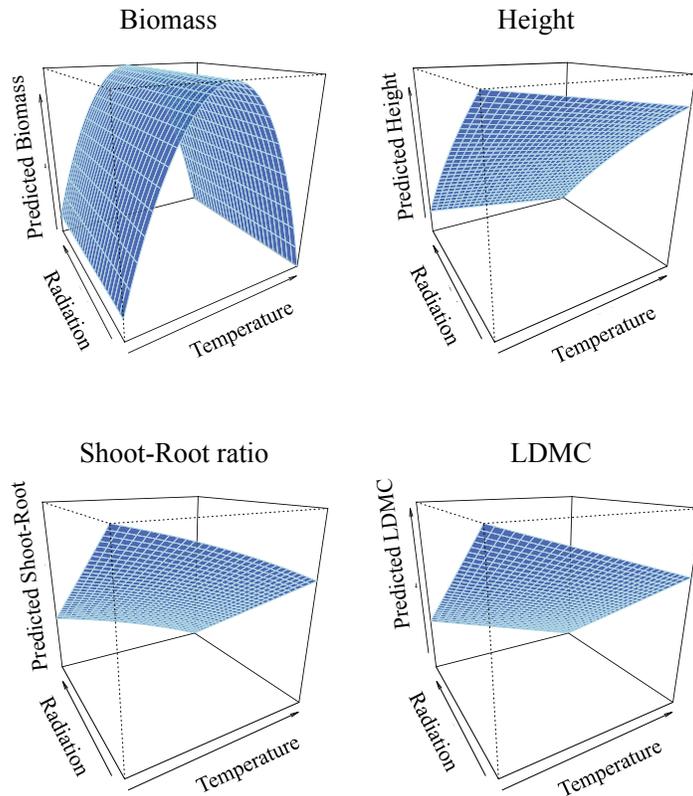
from 0.11 to 0.70) but not significantly related to the niche gradients. Four clusters were detected with Structure, with high but homogeneous levels of admixture within each population (Fig. 1). Globally, individuals in most of the populations were assigned mainly to cluster 2 and then to cluster 3. In populations 33, 63 and 138, individuals were assigned to cluster 3 with higher assignment probabilities (on average 43 %). Two populations (7 and 8) were assigned distinctly to respectively cluster 1 and cluster 4 with only 38% of assignment probability for cluster 2. The assignment probability to these clusters was not statistically linked with the environmental niche of the species.

### 2. Functional structure across the niche

We found that the block design had a significant effect on the measured traits and included it in all further analysis (see Appendix 2).

#### Trait variances across the niche

When comparing the genetic differentiation of these traits at the species level (see  $Q_{ST}$  in Table 1) with the general  $F_{ST}$  (0.0214),  $Q_{ST} > F_{ST}$  for all traits. This result indicates that it is likely that these traits are under divergent selection at the regional scale. Additionally, species level  $CV_G$  showed that in general biomass and height have the highest evolvability. Interestingly, when comparing the  $CV_G$  calculated with populations occurring outside of the global niche (in cold and shadow conditions) versus the other populations, we found that these stressful conditions strongly reduced the  $CV_G$  estimates (Table 1). This substantially corroborates our previous hypothesis on the bottleneck due to maladaptation followed by adaptation in marginal populations (see upper section), as  $CV_G$  should decrease after strong selection. Further, when identifying the trait



**Figure 3.** Trait variation across the niche.

Representation of the relationship between each of the four traits considered and the two axis of regional niche gradients: mean summer temperature and annual solar radiation.

response to climatic niche gradients we found that biomass was related only to the temperature gradient with a quadratic relationships, while in the three other traits were negatively correlated with both temperature and radiation (Figure 3 and see Appendix 3 for estimated parameters).

We found that the total genetic variance ( $G$ 's volume) significantly decreased with decreasing temperature ( $R^2=0.41$ ; Figure 4(a)), indicating that the overall potential to respond to selection for this combination of trait diminishes at the cold edge of *A. artemisiifolia* niche. Second, we detected that  $P_{max}$  was strongly linked with biomass and that the percentage of variance explained by  $P_{max}$  ( $G$ 's shape) was lower for cold conditions, suggesting that the selection already reduced the evolutionary potential on the plant biomass axes ( $R^2=0.24$ , Figure 4(b)). Third, we observed that the direction the genetic integration  $P_{max}$  varied according to the temperature gradient, and that the populations of particularly cold conditions respond in very similar directions (notably toward higher LDMC). Note that we presented only the results on the temperature gradient, as the intensity of solar radiation in the sampled sites did not significantly impact any of the descriptors of the  $G$  matrices.

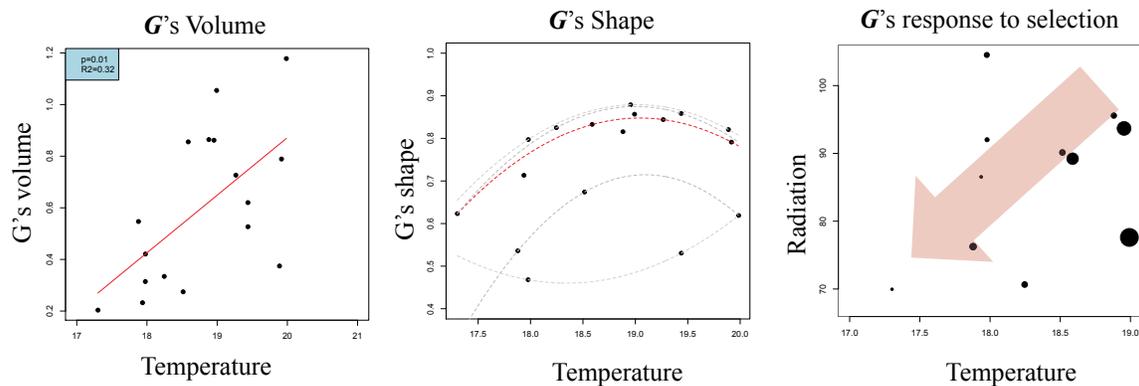
Furthermore, we applied the selection skewer method, and found that when the direction of the selection pressure was directed toward colder and lower levels of solar radiation, populations already at the niche edge have reduced potentials for more marginal conditions.

#### *Linking traits and allelic frequencies*

Finally, the relationship between the predicted trait values and the predicted AFLP marker distributions showed that among the 36 markers potentially under selection 7 were highly correlated with quantitative trait values (correlation coefficient  $> 0.8$ ; Figure 5). Four allele frequencies were correlated with plant height and three with plant biomass.

## DISCUSSION

Over the last decade, several studies have demonstrated that rapid adaptation could occur on short time-scales and fuel invasive species range expansion into new regions (Sax *et al.* 2007). In plants, main evidence have so far concerned adaptive changes in response to new biotic conditions, especially the lack of natural enemies



**Figure 4.** Trait integration across the niche

The three panels represent different aspects of the trait genetic variances and co-variances:  $\mathbf{G}$  matrices. (a) The relation between  $\mathbf{G}$ 's volume (i.e. total genetic variance) per population over the temperature gradient. (b) The relationship between  $\mathbf{G}$ 's shape (proportion of variance explained by  $P_{max}$ ) and the temperature gradient. (c) The relationship between the population potential response to selection toward colder and low levels of solar radiation, and two regional niche axis: temperature and solar radiation (only considering population of the coldest half of the gradient). The size of the dots represents the mean absolute trait displacement after application of the Selection Skewer Method, and the red arrow indicates the direction of the selection applied.

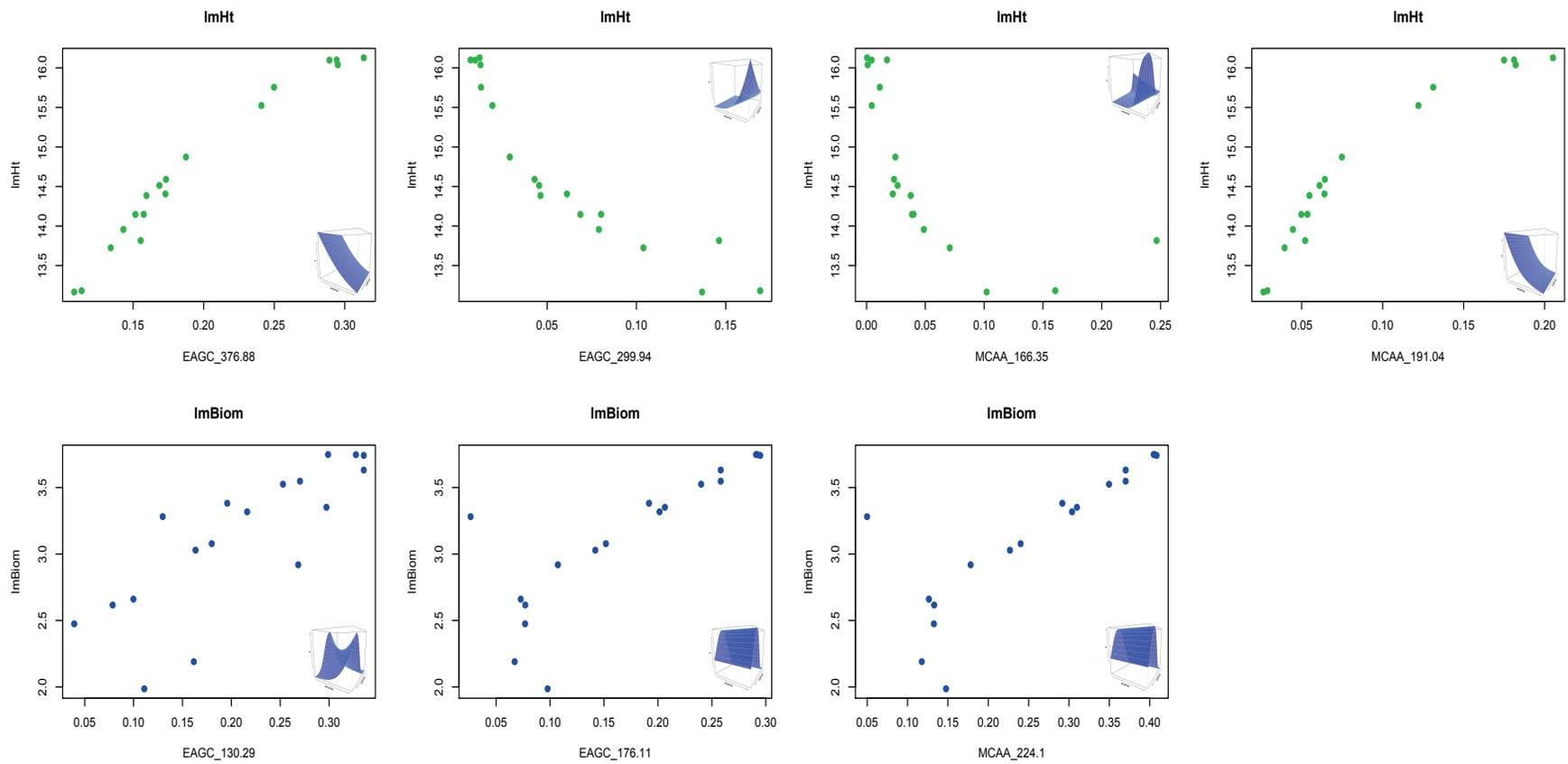
(reviewed by Bossdorf *et al.* 2005), and on life-history traits favouring colonization potential (e.g. Lavergne & Molofsky 2007). Here, we add the first evidence that invasive plants could also evolve towards new climatic environments and thus expand their climatic niche. Our results have important implications for understanding whether and how evolution can foster species invasions along climatic gradients and thus amplify their adverse effects on native biodiversity (Lankau *et al.* 2009).

#### ***Local adaptation despite extensive gene flow***

We built our sampling in a highly heterogeneous alpine region, along strong climatic gradients (from lowlands to sub-alpine habitats) and across different valleys separated by mountains chains. Our data show an important genetic diversity and admixture within all study populations, as well as a limited spatial neutral genetic structure and an absence of isolation by geographic distance. It corroborates previous findings that the species has been introduced several times, favouring post-introduction admixture and high within population genetic diversity (Gaudel *et al.* 2011; Gladieux *et al.* 2011). We also found an important gene flow that tends to homogenize populations across the environmental, despite strong environmental

heterogeneity and the apparent barriers to dispersal. Road developments and high traffic must have contributed to this pattern.

A primary result of our study is that common ragweed populations exhibit local adaptation along the temperature gradient, despite extensive homogenizing gene flows. Indeed, we found both: more phenotypic differentiation (as far as this is captured by  $Q_{ST}$  estimates) than expected from the genetic differentiation inferred from neutral markers ( $F_{ST}$ ), and significant phenotypic clines across the temperature and the radiation gradients. Focusing on AFLP loci that showed non-neutral patterns, we found molecular signature of local adaptation, where the allelic frequency of 36 alleles depended on climatic gradients (e.g. involved in adaptation, hitchhiked via selective sweeps). This confirms the results of previous studies that found local adaptations in this species (Chun *et al.* 2010; Hodgins & Rieseberg 2011), but here we identify the environmental gradients driving this local adaptation. The theoretical implication of this finding is that, after a certain level of environmental steepness, gene flows are not enough to hamper local adaptations (Kawecki 2008).



**Figure 5.** Relationship between AFLP marker frequency and functional trait, with a correlation coefficient above 0.8. The 3D figures in the lower left panels represent the predicted distribution of the genetic markers over the two niche gradients. The first line shows the correlations with Height, and the second line shows the correlations with the plant biomass

The most variable traits were the total plant biomass and plant height, while the shoot-root ratio and LDMC had less evolvability. Interestingly, the plant biomass, which is a fitness-trait, has a unimodal distribution over the temperature gradient suggesting a Gaussian response of fitness. Instead, plant height, a more integrated trait indirectly affected by natural selection, was linearly negatively related to the climatic gradients. These results suggest that divergent selection acts on these traits over the niche of the species, and more specifically in the population located out of the species' global niche (cold and low solar radiation conditions) where the trait evolvability is reduced by more than 30%. However, as the maternal effect was ignored, the exact values of the  $Q_{ST}$  and  $CV_G$  should be taken cautiously. Interestingly, our results are in disagreement with Chun and colleagues (2011) who identified a significant diversifying selection only for the reproductive allocation but not for height and total biomass. In our work, we however try to have a consistent representation of the abiotic niche of the species, which can explain why we detect higher  $Q_{ST}$ . Moreover the work by Hodgins and Rieseberg (2011) showed genetic differentiation in life-history traits between native and introduced populations in Europe (mainly France), notably for growth, biomass and plant width, thus supporting our conclusions.

The full analysis of genetic, functional and niche data provide strong arguments for the rapid adaptation of *A. artemisiifolia*. Such rapid phenomenon has probably been enhanced by the combination of the multiple introduction history of the species, increasing the genetic diversity in the region, an active gene flow, and the opportunity to reach colder climates. It has been shown that the recent populations in France have greater allelic and genetic diversity than historical populations, with a lower level of population differentiation and less structured (Chun *et al.* 2010).

### **Genetically based climatic niche expansion**

To evaluate whether local adaptation has permitted niche expansion, we use an innovative stratified sampling-design based on the climatic niche of the species inferred both at the global and regional scale. This allowed us locating populations currently experiencing novel climatic space compared to the species' worldwide climatic niche. We found that the species is currently colonizing areas of cold climate and that populations at the

forefront of this migration have lower genetic diversity and reduced phenotypic variation than non-marginal populations, as expected due to founder effects (e.g. Boucher *et al.* 2012). These populations also present a significant decrease in the biomass genetic variance (very correlated with  $P_{max}$  of the  $G$  matrices) with cold environment, leading to a strong decrease in their response capacities to selection toward colder and more shaded environments. In other words, these populations allowing niche expansion via local adaptation will have little possibilities for further adaptations to more marginal conditions, slowing down the niche expansion process.

Although it might appear surprising that populations at niche edges show continued adaptation towards marginal climatic conditions despite evidence for genetic bottlenecks, it was already observed for other invasive organisms (Dlugosch & Parker 2008). Two mechanisms could explain this result. First, marginal populations have a stronger genetic isolation than lowland populations, which may favour local adaptations and niche novelties by limiting genet swamping from lower-land populations. Second, niche spread may have been favoured by a shift towards more selfing (as suspected from Gaudeul *et al.* 2011), which conserves co-adapted alleles. Testing this hypothesis will require the use of co-dominant markers to analyse progenies produced in natural conditions (e.g. measuring population kinship and breeding systems), in order to better understand the nature and extent of genetic variation in phenotypic traits.

### **Conclusion**

Our study provides a timely case study of how molecular tools, when combined with distribution modelling, laboratory and field studies, provide essential glimpses into the nature of adaptive constraints during invasion. Three key messages can be taken home:

- Global and regional niche estimation are crucial to identify niche expansion.
- When regional niche gradients are carefully selected, AFLP markers can provide insightful signatures of selection.
- *A. artemisiifolia* has already expanded its niche toward stressful mountainous conditions, but this expansion will be slowed down in the future due to functional trait correlations.

## ACKNOWLEDGMENTS

We are in F. Boucher's debt for his crucial explanations on *G* matrix analysis and more. We particularly want to thank L. Zupan, F. Boucher, S. Pironon, T. Münkemüller, C. Roquet, D. George, L. Chalmandrier, C. Gallouet, L. Garraud and J. Van Es for their help when collecting ambrosia in the field and in the common garden. This common garden would not have been possible without the help of S. Fillacier, F. Houard, B. Bertrand, A. Pichet, M. Girard, M. Molinatti and M.H. Marie. We also would like to thank P. de Villemereuil for exciting discussion on  $Q_{ST}$  and  $H^2$  estimations.

## REFERENCES

- Albert, C.H., Yoccoz, N.G., Edwards, T.C., Graham, C.H., Zimmermann, N.E. & Thuiller, W. (2010). Sampling in ecology and evolution - bridging the gap between theory and practice. *Ecography*, 33, 1028-1037.
- Alexander, J.M. & Edwards, P.J. (2010). Limits to the niche and range margins of alien species. *Oikos*, 119, 1377-1386.
- Arrigo, N., Tuszyński, J.W., Ehrich, D., Gerdes, T. & Alvarez, N. (2009). Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *Bmc Bioinformatics*, 10.
- Benichou, P. & Le Breton, O. (1987). Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. *La Météorologie*, 7, 23-34.
- Bossdorf, O., Auge, H., Lafuma, L., Rogers, W.E., Siemann, E. & Prati, D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia*, 144, 1-11.
- Boucher, F.C., Thuiller, W., Arnoldi, C., Albert, C.H. & Lavergne, S. (2013). Unravelling the architecture of functional variability in wild populations of *Polygonum viviparum* L. *Functional Ecology*, 27, 382-391.
- Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007). Evidence of climatic niche shift during biological invasion. *Ecology Letters*, 10, 701-709.
- Calsbeek, B. & Goodnight, C.J. (2009). Empirical comparison of *G* matrix test statistics: finding biologically relevant change. *Evolution*, 63, 2627-2635.
- Chauvel, B., Dessaint, F., Cardinal-Legrand, C. & Bretagnolle, F. (2006). The historical spread of *Ambrosia artemisiifolia* L. in France from herbarium records. *Journal of Biogeography*, 33, 665-673.
- Cheverud, J.M. (1996). Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S-fuscicollis*) tamarins. *Journal of Evolutionary Biology*, 9, 5-42.
- Chun, Y.J., Fumanal, B., Laitung, B. & Bretagnolle, F. (2010). Gene flow and population admixture as the primary post-invasion processes in common ragweed (*Ambrosia artemisiifolia*) populations in France. *New Phytologist*, 185, 1100-1107.
- Chun, Y.J., Le Corre, V. & Bretagnolle, F. (2011). Adaptive divergence for a fitness-related trait among invasive *Ambrosia artemisiifolia* populations in France. *Molecular Ecology*, 20, 1378-1388.
- Cliff, A.D. & Ord, J.K. (1973). *Spatial autocorrelation*. Pion, London.
- Dlugosch, K.M. & Parker, I.M. (2008). Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology Letters*, 11, 701-709.
- Essl, F., Dullinger, S. & Kleinbauer, I. (2009). Changes in the spatio-temporal patterns and habitat preferences of *Ambrosia artemisiifolia* during its invasion of Austria. *Preslia*, 81, 119-133.
- Facon, B., Genton, B.J., Shykoff, J., Jarne, P., Estoup, A. & David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution*, 21, 130-135.
- Fumanal, B., Chauvel, B., Sabatier, A. & Bretagnolle, F. (2007). Variability and cryptic heteromorphism of *Ambrosia artemisiifolia* seeds: What consequences for its invasion in France? *Annals of Botany*, 100, 305-313.
- Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010). Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology*, 98, 790-799.
- Gallien, L., Douzet, R., Pratte, S., Zimmermann, N.E. & Thuiller, W. (2012). Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography*, 21, 1126-1136.
- Gallien, L., Munkemüller, T., Albert, C.H., Boulangéat, I. & Thuiller, W. (2010). Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, 16, 331-342.
- Gaudeul, M., Giraud, T., Kiss, L. & Shykoff, J.A. (2011). Nuclear and Chloroplast Microsatellites Show Multiple Introductions in the Worldwide Invasion History of Common Ragweed, *Ambrosia artemisiifolia*. *Plos One*, 6.
- Gladieux, P., Giraud, T., Kiss, L., Genton, B.J., Jonot, O. & Shykoff, J.A. (2011). Distinct invasion sources of common ragweed (*Ambrosia artemisiifolia*) in Eastern and Western Europe. *Biological Invasions*, 13, 933-944.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International*

*Journal of Climatology*, 25, 1965-1978.

Hodgins, K.A. & Rieseberg, L. (2011). Genetic differentiation in life-history traits of introduced and native common ragweed (*Ambrosia artemisiifolia*) populations. *Journal of Evolutionary Biology*, 24, 2731-2749.

Holt, R.D. & Keitt, T.H. (2005). Species' borders: a unifying theme in ecology. *Oikos*, 108, 3-6.

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, 130, 195-204.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403-1405.

Kawecki, T.J. (2008). Adaptation to Marginal Habitats. *Annual Review of Ecology Evolution and Systematics*, 39, 321-342.

Kirkpatrick, M. (2009). Patterns of quantitative genetic variation in multiple dimensions. *Genetica*, 136, 271-284.

Lachmuth, S., Durka, W. & Schurr, F.M. (2010). The making of a rapid plant invader: genetic diversity and differentiation in the native and invaded range of *Senecio inaequidens*. *Molecular Ecology*, 19, 3952-3967.

Lankau, R.A., Nuzzo, V., Spyreas, G. & Davis, A.S. (2009). Evolutionary limits ameliorate the negative impact of an invasive plant. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 15362-15367.

Lavergne, S. & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 3883-3888.

Lavergne, S., Mouquet, N., Thuiller, W. & Ronce, O. (2010). Biodiversity and Climate Change: Integrating Evolutionary and Ecological Responses of Species and Communities. In: *Annual Review of Ecology, Evolution, and Systematics, Vol 41* (eds. Futuyma DJ, Shafer HB & Simberloff D). Annual Reviews Palo Alto, pp. 321-350.

Manel, S., Gugerli, F., Thuiller, W., Alvarez, N., Legendre, P., Holderegger, R., Gielly, L., Taberlet, P. & IntraBioDiv, C. (2012). Broad-scale adaptive genetic variation in alpine plants is driven by temperature and precipitation. *Molecular Ecology*, 21, 3729-3738.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, 15, 59-69.

Martin, G., Chapuis, E. & Goudet, J. (2008). Multivariate Q(st)-F(st) Comparisons: A Neutrality Test for the Evolution of the G Matrix in Structured Populations. *Genetics*, 180, 2135-2149.

Merila, J. & Crnokrak, P. (2001). Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, 14, 892-903.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2012). vegan: Community Ecology Package. In: *R package version*.

Pigliucci, M. & Kaplan, J. (2000). The fall and rise of Dr Pangloss: adaptationism and the Spandrels paper 20 years later. *Trends in Ecology & Evolution*, 15, 66-70.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Team, t.R.D.C. (2013). nlme: Linear and Nonlinear Mixed Effects Models. In: (ed. version Rp).

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959.

Revell, L.J. (2007). The G matrix under fluctuating correlational mutation and selection. *Evolution*, 61, 1857-1872.

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S., Grosberg, R.K., Hastings, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I. & Rice, W.R. (2007). Ecological and evolutionary insights from species invasions. *Trends in Ecology and Evolution*, 22, 465-471.

Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution*, 50, 1766-1774.

Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009). Evolution and Ecology of Species Range Limits. In: *Annual Review of Ecology Evolution and Systematics*. Annual Reviews Palo Alto, pp. 415-436.

Steppan, S.J., Phillips, P.C. & Houle, D. (2002). Comparative quantitative genetics: evolution of the G matrix. *Trends in Ecology & Evolution*, 17, 320-327.

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.

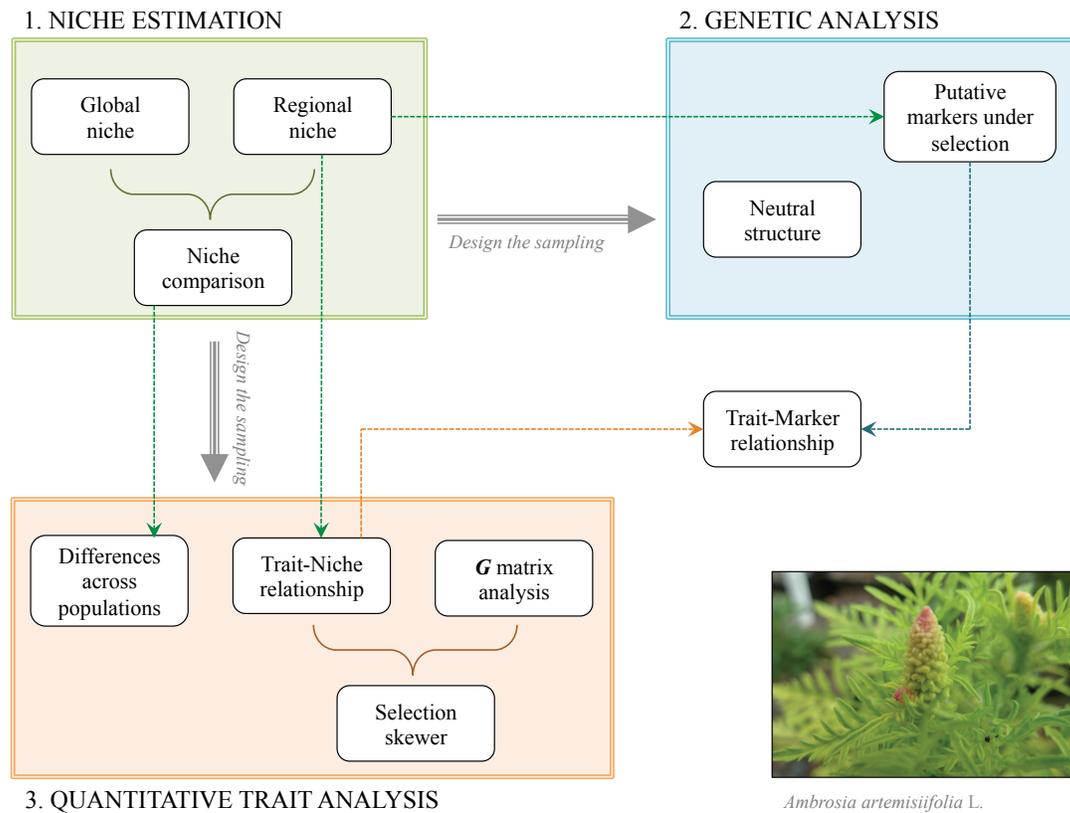
Vos, P., Hogers, R., Bleeker, M., Reijmans, M., Vandeele, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP - A new technique for DNA-fingerprinting. *Nucleic Acids Research*, 23, 4407-4414.

Walther, G.-R., Roques, A., Hulme, P.E., Sykes, M.T., Pysek, P., Kühn, I., Zobel, M., Bacher, S., Botta-Dukat, Z., Bugmann, H., Czucz, B., Dauber, J., Hickler, T., Jarosik, V., Kenis, M., Klotz, S., Minchin, D., Moora, M., Nentwig, W., Ott, J., Panov, V.E., Reineking, B., Robinet, C., Semchenko, V., Solarz, W., Thuiller, W., Vila, M., Vohland, K. & Settele, J. (2009). Alien species in a warmer world – risks and opportunities. *Trends in Ecology & Evolution*, 686-693.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, 15, 323-354.

## Appendix 1. More detailed description of the methods

### 1. Overview of the methods



Schematic representation of methods used in this study, and the links between them. In general three datasets were used to perform the different analysis: niche (green box), genetic (blue box), and quantitative traits (orange box). However, some analysis involved the combination of different datasets, here represented by dotted arrows. On the bottom-right corner is a picture of *Ambrosia artemisiifolia* from our common garden experiment.

### 2. Amplified Fragment Length Polymorphism (AFLP) protocol

Digestion of total DNA was performed together with ligation of double stranded adaptors for 4h at 37 °C in a 11µl mix using 1U of MseI, 5U of EcoRI (New England Biolabs), 1U of T4 DNA ligase (Roche and its associated buffer), 50mM NaCl and 50ng/µl BSA. Products were diluted 1:10 and the preselective amplification (120s at 72°C; 25 cycles 30s at 94°C, 30s at 56°C, 120s at

72°C; with a final elongation 10min at 72°C) was carried out in a 25µl volume containing 5pmol of the *EcoRI*+A primer, 5pmol of the *MseI*+C primer, 200µM of each dNTP, 1.5 mM MgCl<sub>2</sub>, 1X PCR Buffer II pH 8.3, 0.5U *AmpliTag* Gold DNA polymerase (Applied Biosystems) and 3µL of diluted digestion-ligation reaction. After a 1:20 dilution of preselective PCR products, selective amplifications were conducted with three primer combinations: *EcoRI*+ACG / *MseI*+CAA, *EcoRI*+ACT / *MseI*+CTG and *EcoRI*+AGC / *MseI*+CTG. *EcoRI*+3 primers were labelled with 6-carboxy-fluorescein (6-FAM). Each selective amplification reaction contained 2.5pmol of each primer, 200µM of each dNTP, 2.5mM MgCl<sub>2</sub>, 1× Taq polymerase buffer, 8ng/µL BSA, 0.5U *AmpliTag* Gold DNA polymerase (Applied Biosystems) and 2.5µL of diluted preselective amplification reaction, in a final volume of 12.5µL. We used the following PCR conditions: 95°C for 10min; 13 cycles of 30s at 94°C, 60s at 65-55°C, and 60s at 72°C; 23 cycles of 30s at 94°C, 60s at 56°C, 60s at 72°C; and a final elongation of 10min at 72°C. PCR products were purified using columns of half to half 5% Sephadex G50 and Sephacryl S200. Finally, 1.5µl of the diluted FAM labelled products were mixed with 10µl of HiDi formamide and 0.1µl Genescan ROX 500 size standard (Applied Biosystems), and electrophoresed on an ABI PRISM XL 3130 capillary sequencer.

### 3. Description of the neutral genetic structure

#### *3.1. Fis, Fst and He calculations*

To calculate the inbreeding coefficient  $F_{IS}$  and the degree of population differentiation  $F_{ST}$ , which are not directly available for dominant markers such as AFLPs, we used Bayescan v.2.1 (Foll *et al.* 2010). The method infers the proportion of heterozygote individuals based AFLP band intensities. BayeScan assumes that all sampled populations stem from an ancestral population according to the F-model (Falush *et al.* 2003).  $F_{IS}$  and  $F_{ST}$  statistical significance were tested using the randomization procedure with 20 pilot runs, 5000 iterations, 50,000 burn in and 50,000 of thinning interval (Foll *et al.* 2010; Fischer *et al.* 2011). To compare with recent studies, a global  $F_{ST}$  for all populations was also generated with AFLP-surv v.1.0 (with 9,999 randomization, choosing the non-uniform prior distribution and an average  $F_{IS}$  estimated from Bayescan analysis). For each population the genetic diversity  $H_e$  (Nei 1987), was also calculated with AFLP-surv v.1.0 (with 9,999 randomization, choosing the non-uniform prior distribution and an average  $F_{IS}$  estimated from Bayescan analysis; Vekemans *et al.* 2002).

### *3.2. Genetic structure with Structure*

To investigate the genetic structure of the populations we used the Bayesian clustering program Structure 2.3 (Prichard *et al.* 2000) from which we extracted the optimal number clusters. This method determines the more likely number of genetic clusters from the dataset and assigns to each individual a probability of belonging to these clusters. Knowing the high gene flow of this anemophilous invader we used the admixture model, using the LocPrior option (Hubisz *et al.* 2000) and assuming correlated allele frequencies across populations (Pritchard *et al.* 2000; Falush *et al.* 2007). We run Structure for 1,000,000 iterations after a burn-in of 500,000 iterations, and repeated the procedure 5 times for each K value from  $K=1$  to  $K=20$  ( $K$  being the number of clusters). The most parsimonious number of cluster was identified using the log-likelihood curve, and we chose the one at the beginning of the log-likelihood plateau. For the most likely number of clusters ( $K=4$ ) we performed five more runs, averaged the ten runs using CLUMPP v.1.1.2 (Jakobsson & Rosenberg 2007), and displayed it using the software distruct v.1.1 (Rosenberg 2007).

### *4. Common garden protocol*

The dormancy of all seeds was broken at the same time (12<sup>st</sup> April 2011) and germination condition were of 20/10°C 12h/12h light/dark conditions per day (daily permutation of seed family position in the germination room to avoid position effect; following Willemsen 1975a recommendations). The germinated seeds were then sown in 10x10x20 cm<sup>3</sup> pot size containing a mixture of 1/3 sand, 1/3 compost soil and 1/3 row soil. These pots were placed for the first 15 days under a canvas for solar radiation protection in order to protect the individuals that came from sites with lower levels of solar radiations, and then randomly placed into 3 rows of 10 blocks containing each about 100 individuals. Throughout the experiment the plants were watered until soil saturation every two days.

## Appendix 2. Testing for experimental design effects.

The common garden was composed of three “main blocks” (called *lines*), sub-divided in ten “smaller blocks” (called *block*). Four each trait measured at the end of the experiment, these two block effects were tested to detect whether they have a significant influence on the measurements. The results are presented in the table below.

	Biomass	Height	Shoot/Root	LDMC
Line/Block effect	p-val<0.01	p-val=0.014	p-val<0.01	p-val<0.01
Line effect only	p-val<0.01	p-val<0.01	p-val<0.01	p-val=0.386
Block effect only	p-val<0.01	p-val<0.01	p-val<0.01	p-val<0.01

## Appendix 3. Coefficients estimated for the relationship between trait and niche gradients.

	Biomass	Height	ShootRoot	LDMC
Temperature	13.57	-0.102	2.157	-0.017
Temperature <sup>2</sup>	-0.247	–	-0.057	–
Radiation	–	–	-1.409	-0.027
Radiation <sup>2</sup>	–	-0.013	–	–
Temperature:Radiation	–	–	0.052	0.001
$\Delta AIC$	13.7 (***)	2.5 (*)	6.8 (**)	3.5 (*)

ANOVA signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1

## Appendix 4. Results of the tests of isolation by distance

	IBD geographic		IBD environmental	
All markers	–	p-val: 0.259	positive	p-val: 0.008
Only neutral markers	–	p-val: 0.787	–	p-val: 0.893



# *Synthesis*



## 1. ECOLOGICAL AND METHODOLOGICAL FINDINGS

Once an invasive species is introduced, three key drivers influence its success: environmental filtering, competitive interactions and rapid evolution. Because these drivers are usually studied independently, we still know little about how they interact. Over the course of my PhD I have attempted to better understand these three drivers independently and to use this knowledge to develop a better understanding of their interactions. In the following three sections, the major findings concerning each driver are summarized, and within each section the link between the focal driver and the two other drivers are discussed.

### 1.1. ENVIRONMENTAL FILTERING

#### *(1) Identification of population invasion stages*

Combining the literature review and synthesis of Chapter 1 with the empirical analyses of Chapter 2, I have demonstrated that the apparent methodological limitations of statistical niche models (i.e. strong assumptions) can be turned into an advantage (i.e. improve process understanding) when interpreting the results carefully. My strategy has been to compare the invader's environmental niches estimated in the invaded region of interest (the *regional niche*) and at a global scale (the *global niche*). The regional niche is not necessarily estimated in an equilibrium situation in which the invader's occurrences match its favourable environmental conditions. On the contrary, the global niche is assumed to be stable, as it is constituted by a high number of occurrences distributed worldwide which can buffer local dis-equilibrium situations (Gallien *et al.* 2010). Based on the comparison of both niches in the studied region, I have developed a framework that allows us to learn more about the stage of invasion of different populations. Using this framework, I have identified four

different stages for invading populations: equilibrium, colonisation, adaptation and maladaptation (sink populations). Because we usually do not or cannot *a priori* know the stage of invasion of all invading populations (without long-term time series data), such information could bring complementary knowledge to other fields of investigation in invasion research, such as community ecology or population genetics. Furthermore, this screening could be of great interest to environmental managers who have to decide in which populations to invest most effort in order to prevent the invasion of pest species.

### *(2) Linking invasion stages to community ecology*

In the context of community ecology, the relative importance of different assembly processes is known to vary among invasion stages (Richardson *et al.* 2000, Theoharides & Dukes 2007). During the introduction phase one could expect that neither the environmental nor the biotic filters influence the species presence, and that only propagule pressure is important (many populations can be considered as sink populations). During the establishment phase, both the environmental filtering and the niche opportunity could become more important, as the invader has to maintain its population size without immigration. Finally, the importance of competitive exclusion should increase during the invasion and landscape colonisation stages, as the invader can spread widely over various kinds of native communities. The four stages that I proposed in Chapter 2 (i.e. equilibrium, colonisation, adaptation and maladaptation) do not perfectly match the different temporal phases of invasion that are usually considered (i.e. introduction, establishment and invasion). However, some parallels between these two classifications can be drawn and our proposed framework could therefore help to identify different temporal phases of the invasion process. For instance, maladapted populations can be considered as just introduced (as they are sink populations), while colonising and equilibrium populations can be either establishing or invading. Furthermore, one may wonder whether adapting populations show mechanisms of assembly different from other population stages. As an example, one could expect a decrease in the importance of competitive exclusion compared to colonizing populations due to functional

trade-offs between physiological traits involved in the adaptation to new environments and the ones involved in competitive interactions. Such a study could play an important role in enhancing understanding of species range dynamics.

### *(3) Linking invasion stage and population genetics*

In Chapter 5 I have shown how the information on the invasion stage can be used to answer research questions in the field of population genetics. My case study species has been the invasive weed *Ambrosia artemisiifolia* L. In this study, I successfully verified that populations detected as being in the “adaptation stage” by the framework developed in chapter 2 presented unique adaptive genetic characteristics, detected via an independent approach of population genetics. In a similar way, it could be possible to take advantage of the populations identified as “colonising populations” to gain an in-depth understanding of the influence of founder events and genetic drift on population functional characteristics. For example, it would be possible to follow the  $\mathbf{G}$  matrix evolution (e.g. its shape, size and direction, Kirkpatrick 2009, Boucher *et al.* 2013) during colonisation. This would enable to identify whether and how genetic drift can increase  $\mathbf{G}$ 's volume, or whether and how founder events change  $\mathbf{G}$ 's shape and direction. This avenue of research could bring new insights into how the genetic component of the populations' functional structure can reveal parts of their colonisation history, and could help to answer questions such as “How much of the observed phenotypic diversity is due to the colonisation history?”

## 1.2. BIOTIC FILTERING

Focusing on the role of community assembly rules in driving biological invasions, we provided new insights into known methodological limitations: (i) the influence of the spatial scale in the detection of assembly processes, and (ii) the comparison of different metrics and null model performance at detecting the rules. My most important conclusion from this work is that a clear understanding of a methodology's limitations before it is applied is always beneficial. This is the only means of avoiding seemingly contradictory results, as they are widespread in the literature on drivers of community assembly and invasion dynamics (Shea & Chesson 2002, Proches *et al.* 2008, Thuiller *et al.* 2010).

### *(1) Addressing the scale issue with a field data analysis*

The choice of the spatial scale in invasion ecology, but also in community ecology in general, can significantly affect the detection of the drivers of species assemblages. In Chapter 3.2 we have shown, in agreement with theoretical expectations, that broad scale species records favour the detection of environmental filtering, while finer sampling resolutions also enable competitive interactions to be detected (Vamosi *et al.* 2009).

In the same line of thought, our results in Chapter 4 suggest that the assembly processes we are able to detect using field data also depend on the choice of the co-existing species that compose the community. In this study, considering the case of the invasion by *Solidago canadensis*, we showed that indices using either all species recorded in the community (i.e. herbaceous and tree species) or only species of the same growth form as the invader (i.e. only herbaceous species for an herbaceous invader) produced different patterns. In the former case, competitive interactions seemed to be prevailing (over-dispersion pattern), while in the latter case environmental filtering (clustering pattern) was detected. Considering both indices together, we identify that *S. canadensis* was more likely to invade communities containing trees species (trees being functionally

very different, which explains the over-dispersion pattern), because the trees modify the local abiotic conditions by providing for instance more shade (which explains the clustering pattern of shade tolerant herbaceous species). As a conclusion, it seems wise to a priori define the subset of species that is most likely to interact with regard to the study question, use them for the analysis, and potentially test the sensitivity of the results to the subset choice (e.g. Münkemüller *et al.* submitted, Appendix 2).

*(2) Addressing the metric and null model issue with a virtual ecologist approach*

In Chapter 3.3, we also showed the usefulness of a virtual ecologist approach when estimating the methodological efficiency and power in detecting underlying processes via simulated experiments (Zurell *et al.* 2010). Simulated experiments are an important tool in the field of ecology and evolution, as they enable to assess whether existing methods and subsequent analyses are actually fit for purpose. In other words, they ask whether, with perfect data (i.e. where we know exactly what is driving the community structure), the methods or analyses can provide a perfectly clear and correct answer to the question under study. For instance, in the context of community ecology, it is usually assumed that even though the existing indices do not perfectly describe the reality (because they are a simplification of it), they should still, on average, provide reliable answers. Similarly, it is usual to assume that even though the field-datasets are not perfect they should on average provide reliable information on the studied processes.

As a preliminary result of Chapter 3.3, I could see (with a rule-based community assembly model) that among four different indices proposed to investigate mechanisms of invasion, some perform better at detecting particular processes than others. For instance, the mean distance of an invader to every native species (MDNS) more efficiently detects competitive interactions than the distance to the most abundant species (DMAS), while the distance to the nearest native species (DNNS) weakly captures interactions between

competition and environmental filtering processes. Moreover, these differences are strongly influenced by the environmental heterogeneity of the pool of communities considered. Indeed, for all indices, the competition component is over-estimated when the environment is homogeneous across communities, and the interaction between competitive interaction and environmental filtering difficultly captured. These results highlight once again that data sampling is a critical action that should be tailored to the question that is addressed.

Further developments of the simulation model include the implementation of the process of competitive exclusion via fitness differences (asymmetric competition, Chesson 2000). The implementation of fitness difference in the simulation model will be made feasible by using different traits to simulate the environmental niche and the asymmetric competition. Indeed, if one would use the same traits to define species niche and competitive superiority, the patterns observed under competitive exclusion would always be the same as those obtained under environmental filtering. Moreover, with the simulation of two traits, it is possible to vary species fitness according to the environmental conditions of the community. For example, species fitness could be lower at the niche edge than at the niche core. Including species exclusion due to asymmetric competition in the simulation model would be worthwhile in order to develop and test appropriate indices that could efficiently reveal competitive exclusion and distinguish it from environmental filtering.

### *(3) Competitive interactions & the environmental niche*

Learning from the different competitive interactions and how to measure them, in Chapter 4 I have incorporated competitive interaction indices into statistical niche models in order to identify the drivers of alpine plant invasions in the French Alps. In this study, I have shown that none of the competitive interaction strategies affect invasive plant presence in the French Alps, at least at the spatial grain of the analysis ( $\sim 100\text{m}^2$ ). Instead, the results indicate that these indices provide information about fine scale environmental filters (such as disturbance regime or soil nutrient content). However, with no further information on these

fine scale filters we reach the limits of our potential to disentangle biotic and the abiotic filters with statistical models. This is partly because species presence does not only depend on the environment, but can also influence the environment itself (Tilman 1982, HilleRisLambers *et al.* 2012). It should be noted that invader abundance is possibly more affected by competitive interactions than its presence (as suggested for native species in Boulangeat *et al.* 2012a). Unfortunately, the effect of competition on invader abundance was difficult to test with our dataset, as it contains only relative species abundances estimations.

A complementary analysis to what we have done in the French Alps could be to apply the same tests on similar datasets within the invaders' native ranges, as well as other invaded ranges. The comparison of the results across native and adventive ranges could foster understanding about why some invaders are particularly virulent in some regions of the world. Indeed, one possibility would be that the invaders more easily win competitive exclusion interactions in some adventive regions than in the native one because the different competitors in different regions do not have the same characteristics. Alternatively, environmental filtering could be more important in the native range than in the invaded regions, explaining why some invaders that have a very narrow native range can be so successful abroad (e.g. Argentinean ant, Roura-Pascual *et al.* 2009).

### 1.3. EVOLUTIONARY FILTERING

#### *(1) Identification of adaptive niche expansion*

In Chapter 5, with the case study of *A. artemisiifolia*, we showed that the ecological niche of a species could evolve quickly, leading to rapid adaptation towards colder climates and thus niche expansion. This work has two particularly interesting aspects. First, it was made possible with the *a priori* estimation of the two main climatic niche axes of the species in the French Alps (i.e. temperature and solar radiation), which we used to adequately design the sampling in every region of its niche (Albert *et al.* 2010). Second, instead of considering the evolution of single traits independently, we instead looked at the evolution of variances and covariances of multiple traits, which capture more adequately the evolutionary potential of natural populations subjected to directional selection. Overall, this result shows the first example of rapid climatic niche expansion for an invasive species over extremely short time scales (at most 150 years in this case study). If these results are generalizable, the observed rapidity of the adaptation process would be a severe issue for the identification of current and future hotspots of invasion because ignoring the possibility of rapid niche expansion could lead to an underestimation of the threat caused by invaders, i.e. the speed and range of their future expansion. However, it should be noted that this work also suggests that further niche expansion toward colder conditions will be limited. The results, nonetheless, emphasize the current call for the evolutionary mechanisms that influence species distributions to be integrated into semi-mechanistic hybrid models (Gallien *et al.* 2010, Thuiller *et al.* 2013).

*(2) Does niche evolution affect competitive interactions?*

As a perspective and to go further than studying the evolution of the environmental niche alone, it could be interesting to simultaneously identify whether biotic interactions between the invader and the native species evolve as well (see also Johansson 2008, Johansson & Jonzen 2012). Indeed, species competitive ability and stress resistance are often thought to involve opposing physiological characteristics. For instance, resistance to cold conditions usually involves small plant size and slow growth rates, while high competitive capacities often rely on high plant height and rapid growth rates (Grime 1974). Therefore, on the one hand, one could expect that niche expansion towards stressful environmental conditions will induce a decrease in species competitive abilities, due to functional trade-offs between traits involved in both. On the other hand, the competitive superiority of the invader over the native species may not be affected if the functional trade-off is weaker than the one of the native species. In both cases, the study of the assembly rules in populations that are adapting to new climatic conditions could bring fresh insights into the drivers of species range limits.

## **2. GENERAL LIMITATIONS & PERSPECTIVES**

Among the theoretical and methodological findings brought by my PhD, I realised that a variety of common limitations remain unclear and that clarifying them would enhance the general understanding of invasion dynamics. Here is a brief outline of some future development areas: (1) the influence of historical contingencies on local and regional species assemblages, (2) the consequences of ignoring dispersal mechanisms, and (3) the underestimation of the impacts of invasion on the natural communities.

## 2.1. HISTORICAL CONTINGENCIES

### *Is co-existence facilitated by co-evolution?*

Invasive species are often thought to be successful as they did not co-evolve with the resident native species (Mitchell *et al.* 2006, Wilsey *et al.* 2009). This assumption is often used as an argument for the use of phylogenies in invasion community ecology (e.g. Strauss *et al.* 2006), but may not be valid in every situation. In the context of invasions, phylogenetic distances between the invader and the native species are expected to help detect successful invaders based on the duration of co-evolution with the native species. Co-evolution can be seen as facilitating co-existence via the processes of niche differentiation (e.g. Losos *et al.* 2003). On the contrary, one can also argue that when an invader did not evolve in the same region as the native species it can be resistant to the native predators, which do not have the necessary adaptation to feed on it (e.g. Mitchell *et al.* 2006). Although seemingly logical, these statements consider that the observed native species assemblages are fixed sets of species that have always co-existed together. Indeed, native assemblages largely result from the dynamics of extinction and colonisation, for example due to past climatic cycles during which species assemblages have been continuously re-organised. In other words, the observed co-existence (at one point in time) does not necessarily imply co-evolution (Stuart & Losos 2013). Therefore, the assumption of coexistence facilitated by co-evolution can in some cases be an inappropriate concept at the local scale of the community. Of course the “co-evolution advantage” does exist in some situations, as clear examples such as the introduction of top predators into insular systems have proven the opposite (i.e. feeding on species un-adapted to predation, see example in Box 2). However, I believe that this is not necessarily the case, as for instance in the particular situation of invasion within a given continent (e.g. Mediterranean species moving northwards in Europe). A global scale analysis could make it possible to test whether the phylogenetic distance from the invader to the native assemblage

is a good indicator of invasion success, while taking into account long-term regional climatic stability.

## 2.2. THE MISSING 4<sup>TH</sup> MUSKETEER: DISPERSAL

During my PhD, I did not explicitly study dispersal even though it is the fourth key component of biological invasions. The dispersal capacities of an invader are important as they allow it to colonise quickly fragmented landscapes (Pysek & Richardson 2007). Indeed, their incorporation into hybrid models of invasion has already proved to be insightful, for instance in identifying invasion debt (Dullinger *et al.* 2012). Furthermore, a better inclusion of dispersal into the tests of community assembly rules (e.g. including dispersal distances in the randomisation algorithm of the null model tests, Chalmandrier *et al.* 2013) and understanding of population genetic dynamics (discussed below) could also be promising.

### *Dispersal, population genetics & invasion management*

Identifying the main routes of gene flow between populations is crucial for estimating and forecasting the speed and direction of adaptive evolution. Understanding the spatial structure of gene flow and its impact upon rapid adaptation may also be useful in order to identify target zones of invasive species management. Spatial structure of gene flow can be studied thanks to landscape genetics tools. In a given landscape, gene flow between two populations is partly limited by geographic distances, but also by the “resistance” of the landscape, which is the difficulty and cost for a given species (e.g. pollen, seed) of crossing a particular environment. For plant invaders, landscape resistance can be affected by a variety of factors: distance to roads, presence of rivers, forests or steep slopes. One possible (data driven) strategy to identify the most important resistance components is simple. As a first step, one has to create resistance maps based on a priori hypothesis of the putative resistance

components (Spear *et al.* 2010). For example, if one thinks that rivers are impassable barriers and road sides are facilitating gene flow, it is possible to create three maps: one with the presence/absence of the river, one with a gradient of cost increasing with distance to roads, and one combining both previous maps. As a second step, one has to calculate the minimum cost of crossing the landscape (also called the *least cost path*) for each pair of populations and every resistant map. As a final step, the selected resistance components are the ones for which population pair wise genetic distances are best explained by their least cost path. From the best resistance map, one can then estimate easiest gene flow corridors by calculating, for instance, the 100 best least cost paths linking all pairs of populations. Finally, once the preferential roads of gene flow have been identified it would be possible to identify whether and how gene flow is promoting adaptive evolution. For example, it can be tested whether all adapting populations are highly interconnected, or whether they suffer gene flow from maladapted populations.

### 2.3. IMPACTS OF INVASIONS

#### *(1) Abiotic & biotic modifications*

In my work I neglected the impacts that the invaders could have on both the local abiotic and biotic conditions. In general invasive plants can have two types of impacts (see Levine *et al.* 2003 and Mitchell *et al.* 2006 for reviews). First, they can modify the local abiotic conditions by changing for example the habitat type, physicochemical components or disturbance regimes. Second, they can change community interaction networks by outcompeting native species, attracting new predators or transmitting diseases. However, the impacts of invasions are difficult to estimate, especially if they take a long time to show an effect. If at global and regional scales there is little evidence of extinctions caused by plant invaders (Davis 2003, Sax & Gaines 2003, 2008), at local scales extirpations of native plants are likely to be a more common outcome of invasion (Hejda *et al.* 2009, Vila *et al.* 2011). In the French Alps, invasions took

place before detailed ecological records were collected, so it is difficult to know how native diversity has changed following invasions. Not accounting for the impact of invaders could have influenced our results if the invaders have (i) created an invasion meltdown (Simberloff & Von Holle 1999, Hobbs *et al.* 2006), (ii) replaced a keystone species, or (iii) were ecosystem engineers (Jones *et al.* 1994, Levine *et al.* 2003). However, as most of the studied invaders of the French Alps do not seem to belong to any of these three categories and are not superior competitors (see Chapter 4), their impacts on the native community are likely to be limited.

## *(2) Invasion and community integration*

During this PhD, invasion and biotic interactions were essentially studied from the community ecology perspective. A complementary approach could shift the viewpoint and use the perspective of species network approaches. In this context, I think that using the concept of *integrated community* (coined by Lortie *et al.* 2004) can be helpful. Community integration, as a parallel to phenotypic integration (used in Chapter 5), can be seen as a measure of the level of inter-dependency between the species present in the community, and could theoretically be used for describing intra- or inter-trophic levels. Communities can thus range from highly individualistic communities (i.e. little integration) to highly interdependent ones (i.e. high integration). For plant communities, with a large number of plant community records, it could be possible to measure the level of species interdependency by constructing a plant species network at the regional level (see also Kéfi *et al.* 2012). In such a network all species of the region could be represented as a node and each pair of species would be connected or not by an edge that can be weighted according to a co-occurrence index (index type as in Boulangéat *et al.* 2012a). Once the regional network has been estimated, local networks could be drawn for each community by simply subsampling the regional one according to the species observed locally. Finally, it would be possible to test whether the presence/absence of an invader (or its abundance) is related to community network characteristics (e.g. total number of links, clustering, nestedness, Ings

*et al.* 2009). For example, one could ask whether the invader is attracted or repulsed by highly integrated communities. From there, it may become possible to draw useful conclusions to inform invasive species management based on the current state of local species networks. Such an approach, I believe, has a great potential to provide a rapid and simple way of building a large number of ecological networks, otherwise impossible to obtain over large spatial scales.

### **3. PERSONAL CONCLUSIONS**

To conclude, by combining niche modelling, community ecology, simulation models and population genetic approaches, I have shown that along the niche gradient different processes interact to hamper (e.g. native resistance) or favour biological invasions (e.g. rapid evolution). These complex dynamic processes make it difficult to accurately forecast invaders' potential distribution, but provide interesting theoretical insights into the important mechanisms to consider and material for new research questions.

## *References*



- Albert, C.H., Yoccoz, N.G., Edwards, T.C., Graham, C.H., Zimmermann, N.E. & Thuiller, W. (2010). Sampling in ecology and evolution - bridging the gap between theory and practice. *Ecography*, 33, 1028-1037.
- Araujo, M.B. & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22, 42-47.
- Baker, H.G. & Stebbins, G.L. (1965). *The Genetics of Colonizing Species*. Academic Press, New York.
- Beerling, D.J., Huntley, B. & Bailey, J.P. (1995). Climate and the distribution of *Fallopia japonica*: use of an introduced species to test the predictive capacity of response surface. *Journal of Vegetation Science*, 6, 269-282.
- Blumenthal, D.M. (2006). Interactions between resource availability and enemy release in plant invasion. *Ecology Letters*, 9, 887-895.
- Boucher, F.C., Thuiller, W., Arnoldi, C., Albert, C.H. & Lavergne, S. (2013). Unravelling the architecture of functional variability in wild populations of *Polygonum viviparum* L. *Functional Ecology*, 27, 382-391.
- Boulangeat, I., Gravel, D. & Thuiller, W. (2012a). Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters*, 15, 584-593.
- Boulangeat, I., Lavergne, S., Van Es, J., Garraud, L. & Thuiller, W. (2012b). Niche breadth, rarity and ecological characteristics within a regional flora spanning large environmental gradients. *Journal of Biogeography*, 39, 204-214.
- Brandes, D. & Nitzsche, J. (2007). Ecology, distribution and phytosociology of *Ambrosia artemisiifolia* L. in Central Europe. *Tuexenia*, 27, 167-194.
- Bridle, J.R. & Vines, T.H. (2007). Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution*, 22, 140-147.
- Broennimann, O. & Guisan, A. (2008). Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, 4, 585-589.
- Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007). Evidence of climatic niche shift during biological invasion. *Ecology Letters*, 10, 701-709.
- Buffon, G.L. (1788). *Histoire naturelle, générale et particulière*.
- Callaway, R. & Maron, J.L. (2006). What have exotic plant invasions taught us over the past 20 years? *Trends in Ecology and Evolution*, 21, 369-374.
- Callaway, R.M., Brooker, R.W., Choler, P., Kikvidze, Z., Lortie, C.J., Michalet, R., Paolini, L., Pugnaire, F.I., Newingham, B., Aschhoug, E.T., Armas, C., Kikodze, D. & Cook, B.J. (2002). Positive interactions among alpine plants increase with stress. *Nature*, 417, 844-848.
- Callaway, R.M. & Walker, L.R. (1997). Competition and facilitation: A synthetic approach to interactions in plant communities. *Ecology*, 78, 1958-1965.

- Calsbeek, B. & Goodnight, C.J. (2009). Empirical comparison of **G** matrix test statistics: finding biologically relevant change. *Evolution*, 63, 2627-2635.
- Carboni, M., Münkemüller, T., **Gallien, L.**, Lavergne, S., Acosta, A. & Thuiller, W. (2012). Darwin's naturalization hypothesis: scale matters in coastal plant communities. *Ecography*, 35, 001-009.
- Cavender-Bares, J., Keen, A. & Miles, B. (2006). Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. *Ecology*, 87, 109-122.
- Chalmandrier, L., Münkemüller, T., **Gallien, L.**, de Bello, F., Mazel, F., Lavergne, S. & Thuiller, W. (2013). A family of null models to disentangle habitat filtering vs. biotic interactions from functional diversity patterns. *Journal of Vegetation Science*, in press.
- Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31, 343-366.
- Cheverud, J.M. (1996). Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. *Journal of Evolutionary Biology*, 9, 5-42.
- Chew, M.K. & Hamilton, A.L. (2010). The Rise and Fall of Biotic Nativeness: A Historical Perspective. In: *Fifty Years of Invasion Ecology: The Legacy of Charles Elton*. Wiley-Blackwell Oxford, UK.
- Clavero, M. & Garcia-Berthou, E. (2005). Invasive species are a leading cause of animal extinctions. *Trends in Ecology & Evolution*, 20, 110-110.
- Clements, F.E. (1916). *Plant succession; an analysis of the development of vegetation*. Carnegie Institution of Washington, Washington.
- Dantonio, C.M. & Vitousek, P.M. (1992). Biological invasions by exotic grasses, the grass fire cycle, and global change. *Annual Review of Ecology and Systematics*, 23, 63-87.
- Darwin, C.R. (1859). *The Origin of Species*. John Murray, London.
- Davis, M.A. (2003). Biotic globalization: does competition from introduced species threaten biodiversity? *Bioscience*, 481-489.
- Davis, M.A., Grime, J.P. & Thompson, K. (2000). Fluctuating resources in plant communities: a general theory of invasibility. *Journal of Ecology*, 88, 528-534.
- de Candolle, A.P. (1839). *Introduction to the Natural System of the Plant Kingdom*.
- de Candolle, A.P. (1855). *Géographie botanique raisonnée, ou exposition des faits principaux et des lois concernant la distribution géographique des plantes de l'époque actuelle*. Masson, Paris.
- De Roy, K., Massimo, M., Negroni, A., Thas, O., Balloi, A., Fava, F., Verstraete, W., Daffonchio, D. & Boon, N. (2013). Environmental conditions and community evenness determine the outcome of biological invasion. *Nature communication*, doi:10.1038/ncomms2392.

- Diamond, J. & Bellwood, P. (2003). Farmers and their languages: The first expansions. *Science*, 300, 597-603.
- Diez, J.M., Sullivan, J.J., Hulme, P.E., Edwards, G. & Duncan, R.P. (2008). Darwin's naturalization conundrum: dissecting taxonomic patterns of species invasions. *Ecology Letters*, 11, 674-681.
- Dlugosch, K.M. & Parker, I.M. (2008). Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology Letters*, 11, 701-709.
- Dullinger, S., Gattringer, A., Thuiller, W., Moser, D., Zimmermann, N.E., Guisan, A., Willner, W., Plutzer, C., Leitner, M., Mang, T., Caccianiga, M., Dirnboeck, T., Ertl, S., Fischer, A., Lenoir, J., Svenning, J.-C., Psomas, A., Schmatz, D.R., Silc, U., Vittoz, P. & Huelber, K. (2012). Extinction debt of high-mountain plants under twenty-first-century climate change. *Nature Climate Change*, 2, 619-622.
- E.E.A. The European Environmental Agency (2012). The impacts of invasive alien species in Europe. (ed. Publication Office of the European Union).
- Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129-151.
- Elton, C.S. (1958). *The ecology of invasions by animals and plants*, London, Methuen.
- Engler, R., Randin, C.F., Thuiller, W., Dullinger, S., Zimmermann, N.E., Araujo, M.B., Pearman, P.B., Le Lay, G., Piedallu, C., Albert, C.H., Choler, P., Coldea, G., De Lamo, X., Dirnboeck, T., Gegout, J.-C., Gomez-Garcia, D., Grytnes, J.-A., Heegaard, E., Hoistad, F., Nogues-Bravo, D., Normand, S., Puscas, M., Sebastia, M.-T., Stanisci, A., Theurillat, J.-P., Trivedi, M.R., Vittoz, P. & Guisan, A. (2011). 21st century climate change threatens mountain flora unequally across Europe. *Global Change Biology*, 17, 2330-2341.
- Excoffier, L., Foll, M. & Petit, R.J. (2009). Genetic Consequences of Range Expansions. In: *Annual Review of Ecology Evolution and Systematics*, pp. 481-501.
- Facon, B., Genton, B.J., Shykoff, J., Jarne, P., Estoup, A. & David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution*, 21, 130-135.
- Fargione, J., Brown, C.S. & Tilman, D. (2003). Community assembly and invasion: An experimental test of neutral versus niche processes. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 8916-8920.
- Fitch, A. (1861). Sixth report on the noxious and other insects of the state of New York. *New York State Agric Soc Trans*, 745-868.
- Forbes, S.A. (1883). *Insects affecting corn* (ed. Office ISE).

- Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010). Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology*, 98, 790-799.
- Gallien, L.**, Douzet, R., Pratte, S., Zimmermann, N.E. & Thuiller, W. (2012). Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography*, 21, 1126-1136.
- Gallien, L.**, Münkemüller, T., Albert, C.H., Boulangéat, I. & Thuiller, W. (2010). Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, 16, 331-342.
- Gonzalez-Moreno, P., Pino, J., Carreras, D., Basnou, C., Fernandez-Rebollar, I. & Vila, M. (2013). Quantifying the landscape influence on plant invasions in Mediterranean coastal habitats. *Landscape Ecology*.
- Graham, C.H., Parra, J.L., Rahbek, C. & McGuire, J.A. (2009). Phylogenetic structure in tropical hummingbird communities. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 19673-19678.
- Grime, J.P. (1974). Vegetation classification by reference to strategies. *Nature*, 250, 26-31.
- Grinnell, J. (1917). The niche relationship of the California Thrasher. *Auk*, 34, 427-233.
- Guisan, A. & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993-1009.
- Halleri, D.A. (1742). An account of a treatise. 'Archiatrī regīī elect. medicin. anatomiae. botan. praelect. c. enumeratio methodica stirpium helvetiae indigenarum. qua omnium brevis descriptio & synonymia. compendium virium medicarum. dubiarum declaratio. navarum & rariorum uberior historia & icones continentur. Gottingiae' Extracted and translated from the Latin by W. Watson. *Phil. Trans. R. Soc. Lond. B*, 369-380.
- Hejda, M., Pyšek, P., Pergl, J., Sadlo, J., Chytrý, M. & Jarosik, V. (2009). Invasion success of alien plants: do habitat affinities in the native distribution range matter? *Global Ecology and Biogeography*, 18, 372-382.
- Hierro, J.L., Maron, J.L. & Callaway, R.M. (2005). A biogeographical approach to plant invasions: the importance of studying exotics in their introduced and native range. *Journal of Ecology*, 93, 5-15.
- HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M. & Mayfield, M.M. (2012). Rethinking Community Assembly through the Lens of Coexistence Theory. *Annu. Rev. Ecol. Evol. Syst.*, 43, 227-248.
- Hobbs, R.J., Arico, A., Aronson, J., Baron, J.S., Bridgewater, P., Cramer, V.K., Epstein, P.R., Ewel, J.J., Klink, C.A., Lugo, A.E., Norton, D., Ojima, D., Richardson, D.M., Sanderson, E.W., Valladares, F., Vilà, M., Zamora, R. & Zobel, M. (2006). Emerging ecosystems: theoretical and management aspects of the new ecological world order. *Global Ecology and Biogeography*, 15, 1-7.

- Holt, R.D., Barfield, M. & Gomulkiewicz, R. (2005). Theories of niche conservatism and evolution: could exotic species be potential tests? In: *Species Invasions: Insights into Ecology, Evolution, and Biogeography* (eds. Sax D, Stachowicz J & Gaines SD). Sinauer Associates Sunderland, MA., pp. 259–290.
- Hooker, J.D. (1867). *Handbook of the New Zealand flora*.
- Howard, L.O. (1898). *The principal insects affecting the tobacco plant*.
- Huston, M.A. (1999). Local processes and regional patterns: appropriate scales for understanding variation in the diversity of plants and animals. *Oikos*, 86, 393-401.
- Hutchinson, G.E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 145-159.
- Huxley, T.H. (1870). *A Liberal Education, Lay Sermons, Addresses, and Reviews*.
- Ibanez, I., Silander, J.A., Wilson, A.M., Lafleur, N., Tanaka, N. & Tsuyama, I. (2009). Multivariate forecasts of potential distributions of invasive plant species. *Ecological applications*, 19, 359-375.
- Inderjit, Cadotte, M.W. & Colautti, R.I. (2005). The ecology of biological invasions: past, present and future. In: *Invasive Plants: Ecological and Agricultural Aspects* (ed. Inderjit). Birkhäuser Verlag Basel, pp. 19–44.
- Ings, T.C., Montoya, J.M., Bascompte, J., Bluethgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H. & Woodward, G. (2009). Ecological networks - beyond food webs. *Journal of Animal Ecology*, 78, 253-269.
- Johansson, J. (2008). Evolutionary responses to environmental changes: How does competition affect adaptation? *Evolution*, 62, 421-435.
- Johansson, J. & Jonzen, N. (2012). Game theory sheds new light on ecological responses to current climate change when phenology is historically mismatched. *Ecology Letters*, 15, 881-888.
- Jones, C.G., Lawton, J.H. & Shachak, M. (1994). Organisms as ecosystem engineers. *Oikos*, 69, 373-386.
- Kawecki, T.J. (2008). Adaptation to Marginal Habitats. *Annual Review of Ecology Evolution and Systematics*, 39, 321-342.
- Kéfi, S., Berlow, E., Wieters, E., Navarrete, S., Petchey, O., Wood, S., Boit, A., Joppa, L., Lafferty, K., Williams, R., Martinez, N., Menge, B., Blanchette, C., Iles, A. & U., B. (2012). More than a meal... Integrating non-feeding interactions into food webs. *Ecology Letters*, 15, 291-300.
- Kettunen, M., Genovesi, P., Gollasch, S., Pagad, S., Starfinger, U., ten Brink, P. & Shine, C. (2009). Technical support to EU strategy on invasive species (IAS) — Assessment of the impacts of IAS in Europe and the EU. In: *Final report for the European Commission, Institute for European Environmental Policy (IEEP), Brussels, Belgium*.

- Kikvidze, Z., Suzuki, M. & Brooker, R. (2011). Importance versus intensity of ecological effects: why context matters. *Trends in Ecology & Evolution*, 26, 383-388.
- Kirkpatrick, M. (2009). Patterns of quantitative genetic variation in multiple dimensions. *Genetica*, 136, 271-284.
- Kolbe, J.J., Glor, R.E., Schettino, L.R.G., Lara, A.C., Larson, A. & Losos, J.B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, 431, 177-181.
- Körner, C. (1999). *Alpine Plant Life*. Springer-Verlag, Berlin.
- Kunstler, G., Lavergne, S., Courbaud, B., Thuiller, W., Vieilledent, G., Zimmermann, N.E., Kattge, J. & Coomes, D.A. (2012). Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic or functional similarity: implications for forest community assembly. *Ecology Letters*, 15, 831-840.
- Lachmuth, S., Durka, W. & Schurr, F.M. (2010). The making of a rapid plant invader: genetic diversity and differentiation in the native and invaded range of *Senecio inaequidens*. *Molecular Ecology*, 19, 3952-3967.
- Lavergne, S. & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 3883-3888.
- Lavergne, S., Mouquet, N., Thuiller, W. & Ronce, O. (2010). Biodiversity and Climate Change: Integrating Evolutionary and Ecological Responses of Species and Communities. In: *Annual Review of Ecology, Evolution, and Systematics*, Vol 41 (eds. Futuyma DJ, Shafer HB & Simberloff D). Annual Reviews Palo Alto, pp. 321-350.
- Levine, J.M. (2000). Species Diversity and Biological Invasions: Relating Local Process to Community Pattern. *Science*, 288, 852-854.
- Levine, J.M., Adler, P.B. & Yelenik, S.G. (2004). A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters*, 7, 975-989.
- Levine, J.M. & D'Antonio, C.M. (1999). Elton revisited: a review of evidence linking diversity and invasibility. *Oikos*, 87, 15-26.
- Levine, J.M., Vila, M., D'Antonio, C.M., Dukes, J.S., Grigulis, K. & Lavorel, S. (2003). Mechanisms underlying the impacts of exotic plant invasions. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 270, 775-781.
- Lindroth, C.H. (1957). *The faunal connections between Europe and North America*. Wiley, New York, USA.
- Lodge, D.M. (1993). Biological Invasions: Lessons for ecology. *Trends in Ecology and Evolution* 8, 133-137.
- Lonsdale, W.M. (1999). Global patterns of plant invasions and the concept of invasibility. *Ecology*, 80, 1522-1536.

- Lortie, C.J., Brooker, R.W., Callaway, R.M., Choler, P., Kikvidze, Z., Michalet, R. & Pugnaire, F.I. (2004). Rethinking plant community theory. *Oikos*, 107, 433-438.
- Losos, J.B., Leal, M., Glor, R.E., de Queiroz, K., Hertz, P.E., Schettino, L.R., Lara, A.C., Jackman, T.R. & Larson, A. (2003). Niche lability in the evolution of a Caribbean lizard community. *Nature*, 424, 542-545.
- Lowe, S., Browne, M., Boudjelas, S. & DE Poorter, M. (2000). 100 of the World's Worst Invasive Alien Species A selection from the Global Invasive Species Database. (ed. The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN)).
- Lyell, C. (1832). *Principles of geology*.
- MacArthur, R.H. (1970). Species-packing and competitive equilibrium for many species. *Theoretical Population Biology*, 1-11.
- MacDougall, A.S., Gilbert, B. & Levine, J.M. (2009). Plant invasions and the niche. *Journal of Ecology*, 97, 609-615.
- Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M. & Bazzaz, F.A. (2000). Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications*, 10, 689-710.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, 15, 59-69.
- Maron, J.L., Vilà, M., Bommarco, R., Elmendorf, S. & Beardsley, P. (2004). Rapid evolution of an invasive plant. *Ecological Monographs*, 74, 261-280.
- Martin, G., Chapuis, E. & Goudet, J. (2008). Multivariate Q(st)-F-st Comparisons: A Neutrality Test for the Evolution of the G Matrix in Structured Populations. *Genetics*, 180, 2135-2149.
- Medina, F.M., Bonnaud, E., Vidal, E., Tershy, B.R., Zavaleta, E.S., Donlan, C.J., Keitt, B.S., Le Corre, M., Horwath, S.V. & Nogales, M. (2011). A global review of the impacts of invasive cats on island endangered vertebrates. *Global Change Biology*, 17, 3503-3510.
- Mitchell, C.E., Agrawal, A.A., Bever, J.D., Gilbert, G.S., Hufbauer, R.A., Klironomos, J.N., Maron, J.L., Morris, W.F., Parker, I.M., Power, A.G., Seabloom, E.W., Torchin, M.E. & Vazquez, D.P. (2006). Biotic interactions and plant invasions. *Ecology Letters*, 9, 726-740.
- Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.-F., Chave, J., Coueron, P., Dalecky, A., Fontaine, C., Gravel, D., Hardy, O.J., Jabot, F., Lavergne, S., Leibold, M., Mouillot, D., Muenkemüller, T., Pavoine, S., Prinzing, A., Rodrigues, A.S.L., Rohr, R.P., Thebault, E. & Thuiller, W. (2012). Ecophylogenetics: advances and perspectives. *Biological Reviews*, 87, 769-785.
- Muenkemüller, T., de Bello, F., Meynard, C.N., Gravel, D., Lavergne, S., Mouillot, D., Mouquet, N. & Thuiller, W. (2012). From diversity indices to community assembly processes: a test with simulated data. *Ecography*, 35, 468-480.

- Olden, J.D., Poff, N.L., Douglas, M.R., Douglas, M.E. & Fausch, K.D. (2004). Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution*, 19, 18-24.
- Panetta, F.D. & Mitchell, N.D. (1991). Homoclimate analysis and the prediction of weediness. *Weed Research*, 31, 273-284.
- Peterson, A.T., Papes, M. & Kluza, D.A. (2003). Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science*, 51, 863-868.
- Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012). Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders. *Science*, 335, 1344-1348.
- Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'Connell, C., Wong, E., Zern, J., Aquino, T. & Tsornondo, T. (2001). Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems and Environment*, 84, 1-20.
- Pimentel, D., Zuniga, R. & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, 52, 273-288.
- Prentis, P.J., Wilson, J.R.U., Dormontt, E.E., Richardson, D.M. & Lowe, A.J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science*, 13, 288-294.
- Proches, S., Wilson, J.R.U., Richardson, D.M. & Rejmánek, M. (2008). Searching for phylogenetic pattern in biological invasions. *Global Ecology and Biogeography*, 17, 5-10.
- Pulliam, H.R. (2000). On the relationship between niche and distribution. *Ecology Letters*, 3, 349-361.
- Pyšek, P. & Richardson, D.M. (2007). Traits associated with invasiveness in alien plants: where do we stand? *Biological Invasions*, 97-122.
- Qian, H. & Kissling, W.D. (2010). Spatial scale and cross-taxon congruence of terrestrial vertebrate and vascular plant species richness in China. *Ecology*, 91, 1172-1183.
- Rejmánek, M., Richardson, D.M., Higgins, S.I., Pitcairn, M.J. & Grotkopp, E. (2005). Ecology of invasive plants: state of the art. In: *Invasive alien species: a new synthesis* (eds. Mooney HA et al.) Island Press, Washington, D.C., pp. 104-161.
- Rey, O., Estoup, A., Vonshak, M., Loiseau, A., Blanchet, S., Calcaterra, L., Chifflet, L., Rossi, J.-P., Kergoat, G.J., Foucaud, J., Orivel, J., Leponce, M., Schultz, T. & Facon, B. (2012). Where do adaptive shifts occur during invasion? A multidisciplinary approach to unravelling cold adaptation in a tropical ant species invading the Mediterranean area. *Ecology Letters*, 15, 1266-1275.
- Richardson, D.M. & Bond, W.J. (1991). Determinants of plant distribution evidence from pine invasions. *The American Naturalist*, 157, 639-668.
- Richardson, D.M. & Pyšek, P. (2012). Naturalization of introduced plants: ecological drivers of biogeographical patterns. *New Phytologist*, 196, 383-396.

- Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000). Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, 6, 93-107.
- Richardson, D.M., Rouget, M., Ralston, S.J., Cowling, R.M., van Rensburg, B.J. & Thuiller, W. (2005). Species richness of alien plants in South Africa: environmental correlates and the relationship with native plant species richness. *Ecoscience*, 12, 391-402.
- Rikli, M. (1903). Die Anthropochoren und der Formenkreis des *Nasturtium palustre* (Leys.). DC. *Berichte der Zürcherischen Botanischen Gesellschaft*, 71–82.
- Roura-Pascual, N., Bas, J.P., Thuiller, W., Hui, C., Krug, R.M. & Brotons, L. (2009). From introduction to equilibrium: reconstructing the invasive pathways of the Argentine ant in a Mediterranean region. *Global Change Biology*, 15, 2101-2115.
- Sax, D.F. (2002). Native and naturalized plant diversity are positively correlated in scrub communities of California and Chile. *Diversity and Distributions*, 8, 193-210.
- Sax, D.F. & Gaines, S.D. (2003). Species diversity: from global decreases to local increases. *Trends in Ecology and Evolution*, 11, 561-566.
- Sax, D.F. & Gaines, S.D. (2008). Species invasions and extinction: The future of native biodiversity on islands. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 11490-11497.
- Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S., Grosberg, R.K., Hastings, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I. & Rice, W.R. (2007). Ecological and evolutionary insights from species invasions. *Trends in Ecology and Evolution*, 22, 465-471.
- Sax, D.F., Stachowicz, J.J. & Gaines, S. (2005). *Species Invasions: Insights into Ecology, Evolution, and Biogeography*. Sinauer Associates Inc., U.S., Sunderland, United States.
- Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution*, 50, 1766-1774.
- Scott, J.K. & Panetta, F.D. (1993). Predicting the Australian weed status of southern African plants. *Journal of Biogeography*, 20, 87-93.
- Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009). Evolution and Ecology of Species Range Limits. *Annual Review of Ecology Evolution and Systematics*. Annual Reviews Palo Alto, 415-436.
- Sexton, J.P., Strauss, S.Y. & Rice, K.J. (2011). Gene flow increases fitness at the warm edge of a species' range. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 11704-11709.
- Shea, K. & Chesson, P. (2002). Community ecology theory as a framework for biological invasions. *Trends in Ecology and Evolution*, 17, 170-176.
- Simberloff, D. & Von Holle, B. (1999). Positive interactions of nonindigenous species: invasional meltdown? *Biological Invasions*, 21-32.

- Spear, S.F., Balkenhol, N., Fortin, M.-J., McRae, B.H. & Scribner, K. (2010). Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, 19, 3576-3591.
- Strauss, Y.E., Webb, C.O. & Salamin, N. (2006). Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Sciences of the USA*, 103, 5841-5845.
- Stuart, Y.E. & Losos, J.B. (2013). Ecological character displacement: glass half full or half empty? *Trends in Ecology & Evolution*, in press.
- Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. & Zimmerman, J.K. (2006). The problem and promise of scale dependency in community phylogenetics. *Ecology*, 87, 2418-2424.
- Thellung, A. (1912). La flore adventice de Montpellier. *Mémoires de la Société Nationale des Sciences Naturelles et Mathématiques de Cherbourg*, 57-728.
- Theoharides, K.A. & Dukes, J.S. (2007). Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. *New Phytologist*, 176, 256-273.
- Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, 10, 2020-2027.
- Thuiller, W., Gallien, L., Boulangéat, I., de Bello, F., Muenkemüller, T., Roquet, C. & Lavergne, S. (2010). Resolving Darwin's naturalization conundrum: a quest for evidence. *Diversity and Distributions*, 16, 461-475.
- Thuiller, W., Lavorel, S., Araujo, M.B., Sykes, M.T. & Prentice, I.C. (2005). Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 8245-8250.
- Thuiller, W., Münkemüller, T., Lavergne, S., Mouillot, D., Mouquet, N., Schifffers, K. & Gravel, D. (2013). A road map for integrating eco-evolutionary processes into biodiversity models. *Ecology Letters*, in press.
- Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 10854-10861.
- Treier, U.A., Broennimann, O., Normand, S., Guisan, A., Schaffner, U., Steinger, T. & Mueller-Schaerer, H. (2009). Shift in cytotype frequency and niche space in the invasive plant *Centaurea maculosa*. *Ecology*, 90, 1366-1377.
- Urban, M.C., Phillips, B.L., Skelly, D.K. & Shine, R. (2008). A toad more traveled: The heterogeneous invasion dynamics of cane toads in Australia. *American Naturalist*, 171, 134-148.
- Vamosi, S.M., Heard, S.B., Vamosi, J.C. & Webb, C.O. (2009). Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, 18, 572-592.

- Vila, M., Espinar, J.L., Hejda, M., Hulme, P.E., Jarosik, V., Maron, J.L., Pergl, J., Schaffner, U., Sun, Y. & Pyšek, P. (2011). Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters*, 14, 702-708.
- Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Rejmánek, M. & Westbrooks, R. (1997). Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, 21, 1-16.
- von Humboldt, A. (1807). *Voyage aux régions équinoxiales du Nouveau Continent* (avec Aimé Bonpland).
- Wallace, A.R. (1870). *Contributions to the Theory of Natural Selection*.
- Wallace, A.R. (1876). *The Geographical Distribution of Animals*.
- Walther, G.-R., Roques, A., Hulme, P.E., Sykes, M.T., Pyšek, P., Kühn, I., Zobel, M., Bacher, S., Botta-Dukat, Z., Bugmann, H., Czucz, B., Dauber, J., Hickler, T., Jarosik, V., Kenis, M., Klotz, S., Minchin, D., Moora, M., Nentwig, W., Ott, J., Panov, V.E., Reineking, B., Robinet, C., Semchenko, V., Solarz, W., Thuiller, W., Vila, M., Vohland, K. & Settele, J. (2009). Alien species in a warmer world – risks and opportunities. *Trends in Ecology & Evolution*, 686-693.
- Warming, E. & Vahl, M. (1909). *Oecology of Plants - an introduction to the study of plant communities*.
- Watson, H.C. (1847). *Cybele Britannica; or British plants and their geographical relations*. Longman, London.
- Welden, C.W. & Slauson, W.L. (1986). The intensity of competition versus its importance - an overlooked distinction and some implications. *Quarterly Review of Biology*, 61, 23-44.
- Wilkinson, D.M. (2002). Ecology before ecology: biogeography and ecology in Lyell's 'Principles' - A critical re-assessment of the significance of Charles Lyell (1832) *Principles of Geology*, vol. II. John Murray, London. *Journal of Biogeography*, 29, 1109-1115.
- Willis, K.J. & Whittaker, R.J. (2002). Species diversity-scale matters. *Science*, 295, 1245-1248.
- Wilsey, B.J., Teaschner, T.B., Daneshgar, P.P., Isbell, F.I. & Polley, H.W. (2009). Biodiversity maintenance mechanisms differ between native and novel exotic-dominated communities. *Ecology Letters*, 12, 432-442.
- Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dormann, C.F., Forchhammer, M.C., Grytnes, J.-A., Guisan, A., Heikkinen, R.K., Høye, T.T., Kuhn, I., Luoto, M., Maiorano, L., Nilsson, M.-C., Normand, S., Ockinger, E., Schmidt, N.M., Termansen, M., Timmermann, A., Wardle, D.A., Aastrup, P. & Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological reviews of the Cambridge Philosophical Society*, 88, 15-30.

- Ziska, L., Knowlton, K., Rogers, C., Dalan, D., Tierney, N., Elder, M.A., Filley, W., Shropshire, J., Ford, L.B., Hedberg, C., Fleetwood, P., Hovanky, K.T., Kavanaugh, T., Fulford, G., Vrtis, R.F., Patz, J.A., Portnoy, J., Coates, F., Bielory, L. & Frenz, D. (2011). Recent warming by latitude associated with increased length of ragweed pollen season in central North America. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 4248-4251.
- Zurell, D., Berger, U., Cabral, J.S., Jeltsch, F., Meynard, C.N., Muenkemüeller, T., Nehrbass, N., Pagel, J., Reineking, B., Schroeder, B. & Grimm, V. (2010). The virtual ecologist approach: simulating data and observers. *Oikos*, 119, 622-635.

# *Appendix 1*





## **A family of null models to distinguish between environmental filtering and biotic interactions in functional diversity patterns**

L. Chalmandrier, T. Münkemüller, L. Gallien, F. de Bello, F. Mazel, S. Lavergne & W. Thuiller

### **Keywords**

Assembly rules; Biotic and abiotic filtering; Limiting similarity; Null models; Simulated communities

### **Nomenclature**

✉ ~~XXXXXXXXXXXXXXXXXXXX~~

Received 3 April 2012

Accepted 5 November 2012

Co-ordinating Editor: Rasmus Ejrnæs

**Chalmandrier, L.** (corresponding author, loic.chalmandrier@ens.fr), **Münkemüller, T.** (✉), **Gallien, L.** (✉), **Mazel, F.** (✉), **Lavergne, S.** (✉) & **Thuiller, W.** (✉): Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Université Joseph Fourier, BP 53, Grenoble Cedex 9, 38041, France

**de Bello, F.** (✉): Institute of Botany, Czech Academy of Sciences, Třeboň, CZ-379 82,

✉ Czech Republic

### **Abstract**

**Questions:** Traditional null models used to reveal assembly processes from functional diversity patterns are not tailored for comparing different spatial and evolutionary scales. In this study, we present and explore a family of null models that can help disentangling assembly processes at their appropriate scales and thereby elucidate the ecological drivers of community assembly.

**Location:** French Alps.

**Methods:** Our approach gradually constrains null models by: (1) filtering out species not able to survive in the regional conditions in order to reduce the spatial scale, and (2) shuffling species only within lineages of different ages to reduce the evolutionary scale of the analysis. We first tested and validated this approach using simulated communities. We then applied it to study the functional diversity patterns of the leaf–height–seed strategy of plant communities in the French Alps.

**Results:** Using simulations, we found that reducing the spatial scale correctly detected a signature of competition (functional divergence) even when environmental filtering produced an overlaying signal of functional convergence. However, constraining the evolutionary scale did not change the identified functional diversity patterns. In the case study of alpine plant communities, investigating scale effects revealed that environmental filtering had a strong influence at larger spatial and evolutionary scales and that neutral processes were more important at smaller scales. In contrast to the simulation study results, decreasing the evolutionary scale tended to increase patterns of functional divergence.

**Conclusion:** We argue that the traditional null model approach can only identify a single main process at a time and suggest to rather use a family of null models to disentangle intertwined assembly processes acting across spatial and evolutionary scales.



## *Appendix 2*



## Scale decisions can reverse conclusions on community assembly processes

Münkemüller, T.<sup>1\*</sup>, Gallien, L.<sup>1</sup>, Lavergne, S.<sup>1</sup>, Renaud, J.<sup>1</sup>, Roquet, C.<sup>1</sup>, Abdulayak, S.<sup>2</sup>, Dullinger, S.<sup>3</sup>, Garraud, L.<sup>2</sup>, Guisan, A.<sup>4</sup>, Lenoir, J.<sup>5</sup>, Svenning, J.-C.<sup>6</sup>, Van Es, J.<sup>2</sup>, Vittoz, P.<sup>4</sup>, Willner, W.<sup>7</sup>, Wohlgemuth, T.<sup>8</sup>, Zimmermann, N.<sup>8</sup> and Thuiller, W.<sup>1</sup>

<sup>1</sup> Laboratoire d'Ecologie Alpine, UMR CNRS 5553, University Joseph Fourier, Grenoble 1, BP 53, 38041 Grenoble Cedex 9, France

<sup>2</sup> Domaine de Charance, Conservatoire Botanique National Alpin, Gap, 05000, France

<sup>3</sup> Dep. of Conservation Biology, Vegetation- and Landscape Ecology, Faculty Centre of Biodiversity, Rennweg 14, 1030 Vienna

<sup>4</sup> Department of Ecology & Evolution, University of Lausanne, 1015 Lausanne; Institute of Earth Sciences, University of Lausanne, 1015 Lausanne, Switzerland

<sup>5</sup> Ecologie et Dynamique des Systèmes Anthropisés (EA 4698), Plant Biodiversity Lab, Jules Verne University of Picardie, 1 rue des Louvels, FR-80037 Amiens Cedex 1, France

<sup>6</sup> Ecoinformatics & Biodiversity Group, Department of Bioscience, Aarhus University, Ny Munkegade 114, DK-8000 Aarhus C, Denmark

<sup>7</sup> Vienna Institute for Nature Conservation and Analyses, Giessergasse 6/7, 1090 Vienna, Austria

<sup>8</sup> Landscape Dynamics, Swiss Federal Research Institute WSL, CH-8903 Birmensdorf, Switzerland

- Corresponding author, email [tamara.muenkemueller@ujf-grenoble.fr](mailto:tamara.muenkemueller@ujf-grenoble.fr), phone +33 (0)476 635733, fax: +33 (0)4 76 51 42 79

### ABSTRACT

#### Aim

Community ecology studies increasingly use information on phylogenetic diversity patterns to infer assembly processes. Here, we quantify how these patterns are influenced by scale choices in terms of spatial and environmental extent and organismic scales.

#### Location

European Alps.

#### Methods

We applied 42 sampling strategies differing in their combination of focal scales. For each resulting sub-dataset, we estimated the phylogenetic diversity of the species pool, phylogenetic  $\alpha$ -diversities of local communities, and assembly process signals in the composition of local communities by comparing observed values with null models (i.e. phylogenetic clustering vs. over-dispersion). Finally, we studied the effects of scale choices on these measures using regression and partial regression analyses.

#### Results

Scale choices were decisive for revealing signals in diversity patterns and thus for conclusions concerning assembly processes. Notably, changes in focal scales sometimes reversed a pattern of over-dispersion into clustering. Organismic scale had a stronger effect than spatial and environmental extent. However, we did not find general rules for the direction of change from over-dispersion to clustering with changing scales. Importantly, these scale issues had only a weak influence when focusing on regional diversity patterns that change along abiotic gradients





## Abstract

Biological invasions, the second major threat to biodiversity, pose significant challenges to conservation management and eco-evolutionary research. Even though invasion processes have been studied for more than 150 years, our capacity to predict their presence today and in the future is still rudimentary. This deficiency stems mainly from the difficulty involved in reliably assessing the ecological niche of an invader, i.e. those environmental and biotic conditions that allow the species to maintain viable populations. In particular, disentangling the abiotic and biotic components of the ecological niche and accounting for their changing over space and time due to evolutionary dynamics is difficult, albeit crucial for the quality of predictions.

The main objective of my PhD has been to address these challenges by improving methodological approaches of niche estimation, advancing our understanding of the role of biotic interactions for invasion processes and studying in greater detail how evolution may affect spatio-temporal niche dynamics. More precisely, (1) with a comprehensive literature review, I started by describing the limits of the different modelling approaches usually applied to predict invasive species distributions. (2) Then, I provided a modelling framework for improving regional environmental niche estimations. (3) Thirdly, I focused on the identification of biotic interactions, and the methods commonly used to identify patterns of symmetric competition in ecological communities. I also implemented a simulation model of community assembly to test the efficiency of these methods. (4) In a fourth part, I studied invaded alpine plant communities and showed that characteristics of the biotic environment in these communities (e.g. symmetric vs. asymmetric competition) were good predictors of invaders' presence. (5) Finally, I provided a first example of a genetic-based, climatic niche expansion of the invasive weed *Ambrosia artemisiifolia* L. in the French Alps by combining information on its environmental niche, genetic structure and functional traits. Taken together, the results of these studies highlight how tightly the different facets of invasion ecology and evolution are interrelated and open the way to an integrated modelling approach that would advance both eco-evolutionary research on invasion dynamics and applied tools for biodiversity protection.

## Résumé

L'invasion biologique, deuxième menace majeure de la biodiversité, pose d'importants défis pour la conservation de la biodiversité et la recherche en éco-évolution. Bien que les espèces invasives aient été étudiées depuis plus de 150 ans, nos capacités à prédire leurs présences aujourd'hui et dans le futur restent rudimentaires. Ce problème est principalement dû à la difficulté d'estimer à la fois les composantes biotiques et abiotiques de la niche des espèces invasives, ainsi que leur évolution dans le temps et l'espace.

L'objectif de ma thèse a été de travailler sur ces défis en améliorant les méthodes d'estimation de niche, en enrichissant notre compréhension du rôle des interactions biotiques dans le processus d'invasion, et en étudiant en détail comment les processus évolutifs peuvent affecter la dynamique spatio-temporelle des niches. Plus précisément, (1) à l'aide d'une revue de la littérature, j'ai commencé par décrire les limites des différentes approches de modélisation utilisées pour prédire la distribution des espèces invasives. (2) Ensuite, j'ai proposé un cadre de modélisation permettant d'améliorer l'estimation des niches abiotiques régionales. (3) Puis, je me suis intéressée à la caractérisation des interactions biotiques, et aux méthodes communément utilisées pour identifier les patrons de compétition symétrique en écologie des communautés. J'ai également implémenté un modèle de simulation d'assemblage de communautés pour tester la performance de ces méthodes. (4) Ces premières études m'ont permis d'étudier à la fois les composantes biotiques et abiotiques des communautés de plantes envahies dans les Alpes. (5) Finalement, j'ai étudié l'évolution de la niche environnementale chez une espèce invasive des Alpes françaises *Ambrosia artemisiifolia* L., à travers une approche reliant *niche environnementale, trait fonctionnels et structure génétique*. Dans l'ensemble, les résultats de ces études montrent à quel point les différentes facettes de l'écologie et l'évolution en invasion sont fortement intriquées. De plus, ils soulignent la nécessité d'une modélisation intégrant les processus écologiques et évolutifs pour ainsi comprendre la dynamique des invasions et proposer des outils de protection de la biodiversité efficaces.