Research

On the importance of intraspecific variability for the quantification of functional diversity

Cécile H. Albert, Francesco de Bello, Isabelle Boulangeat, Gilles Pellet, Sandra Lavorel and Wilfried Thuiller

C. H. Albert (cecile.albert@m4x.org), I. Boulangeat, S. Lavorel, W. Thuiller, Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Univ. Joseph Fourier, BP 53, FR-38041 Grenoble Cedex 9, France. – F. de Bello, Inst. of Botany, Czech Academy of Sciences, Dukelská 135, CZ-379 82 Třeboň, Czech Republic. – G. Pellet, Station Alpine Joseph Fourier, CNRS UMS 2925, Univ. Joseph Fourier, BP 53, FR-38041 Grenoble Cedex 9, France.

Functional diversity (FD) is a key facet of biodiversity used to address central questions in ecology. Despite recent methodological advances, FD remains a complex concept and no consensus has been reached either on how to quantify it, or on how it influences ecological processes. Here we define FD as the distribution of trait values within a community. When and how to account for intraspecific trait variability (ITV) when measuring FD remains one of the main current debates. It remains however unclear to what extent accounting for population-level ITV would modify FD quantification and associated conclusions. In this paper, we address two critical questions: (1) How sensitive are different components of FD to the inclusion of population-level ITV? (2) Does the omission of ITV obscure the understanding of ecological patterns? Using a mixture of empirical data and simulation experiments, we conducted a sensitivity analysis of four commonly used FD indices (community weighted mean traits, functional richness, Rao's quadratic entropy, Petchey and Gaston's FD index) and their relationships with environmental gradients and species richness, by varying both the extent (plasticity or not) and the structure (contingency to environmental gradient due to local adaptation) of population-level ITV. Our results suggest that ITV may strongly alter the quantification of FD and the detection of ecological patterns. Our analysis highlights that 1) species trait values distributions within communities are crucial to the sensitivity to ITV, 2) ITV structure plays a major role in this sensitivity and 3) different indices are not evenly sensitive to ITV, the single-trait FD from Petchey and Gaston being the most sensitive among the four metrics tested. We conclude that the effects of intraspecific variability in trait values should be more systematically tested before drawing central conclusions on FD, and suggest the use of simulation studies for such sensitivity analyses.

Understanding how biodiversity is assembled within communities (Hubbel 2001, Gravel et al. 2006) and how it influences ecosystem functioning (Naeem and Wright 2003, Hooper et al. 2005) are central questions in ecology. Answering these questions usually requires going beyond the taxonomical identity of species by focusing on species' biological characteristics, i.e. functional traits. Functional traits, any measurable feature at the individual level affecting fitness directly or indirectly (Violle et al. 2007), describe organisms' biological activity and directly influence ecological processes (e.g. habitat filtering, competition, mutualism, predation) and properties of communities and ecosystems (Naeem and Wright 2003, Hooper et al. 2005). Functional traits effects are mediated by the kind, range and relative abundance of functional attributes in a given community. These dimensions are collectively known as 'functional diversity' (hereafter FD, Diaz and Cabido 2001, Mason et al. 2005, Diaz et al. 2007) and can be quantified by different components and associated metrics (Pavoine and Bonsall 2011). A community's FD can be generally described by its dominant trait values, usually estimated by a community weighted mean

trait value (i.e. mean of trait values weighted by species abundances, Garnier et al. 2004) and different facets of trait dissimilarity between coexisting species: richness, evenness and divergence (Mouillot et al. 2005). The last decade has been extremely productive from a methodological point of view, with the establishment of multi-level criteria to select the more efficient FD indices for different research questions (Petchey and Gaston 2002, Botta-Dukat 2005, Lepš et al. 2006, Villéger et al. 2008). FD is also increasingly used for a variety of objectives. It is used for instance as a tool to disentangle the effects of abiotic conditions (expected to decrease FD through environmental filtering) and biotic interactions (expected to increase FD through limiting similarity) on community assembly (Petchey and Gaston 2007, Cornwell and Ackerly 2009,). FD remains however a complex concept and no consensus has been reached either on the most efficient way to measure it (Petchey and Gaston 2006, Podani and Schmera 2006), or on how it affects community assembly and ecosystem functioning. One of the main discrepancies regards when and how to account for intraspecific trait variability (ITV) when assessing FD

(Mason et al. 2005, Lepš et al. 2006, Cianciaruso et al. 2009, de Bello et al. 2010, Schleuter et al. 2010).

A change in FD values from one community to another or within a community can be caused by a change in species composition (turnover) and species relative abundance, by ITV or by combinations of these factors (Lepš et al. 2011). Individuals within a species may indeed have highly variable trait values (high ITV, Albert et al. 2010, Hulshof and Swenson 2010, Messier et al. 2010) due to local adaptation or to phenotypic plasticity of traits either along environmental gradients (e.g. bioclimatic, disturbance, Sandquist and Ehleringer 1997), or because individuals may alter their trait values in response to biotic interactions (e.g. competition, Gross et al. 2009). This ITV can lead to significant differences between populations' trait values (population-level ITV), that are more or less driven by environmental gradients and species' niches (Sugiyama 2003, Albert et al. 2010, Thuiller et al. 2010). This contingency is hereafter called 'structure'. Consequently, two communities with the same species composition could potentially have different trait distributions and thus different FD. Although this sounds relatively logical, most studies published so far have ignored ITV in estimating FD, making the implicit assumption that changes in species composition override the effect of population-level ITV (but see Lavorel et al. 2008). The lack of consensus highlighted above could emerge from this omission of population-level ITV.

There is a theoretical understanding that ITV should affect FD (Lepš et al. 2006, Schleuter et al. 2010), and it is becoming obvious that neglecting ITV in the computation of FD might result in a misinterpretation of the real dynamics and functioning of communities (de Bello et al. 2010, Jung et al. 2010). It remains however unclear to what extent accounting for population-level ITV will modify FD quantification and associated conclusions on community structure along environmental gradients. In this paper, we address two critical questions: 1) how sensitive are different components of FD to the inclusion of population-level ITV? 2) Does the omission of ITV obscure the understanding of ecological patterns? To answer these two questions we used field data combined with a simulation study and explored the sensitivity of four common FD indices to the extent and structure of ITV. This original combination of theoretical simulations of ITV with empirical data allowed the investigation of realistic ranges of ITV. Moreover we used ecological niche modeling (Guisan and Thuiller 2005) to simulate the environmental structure of ITV, following Thuiller et al. (2010). We finally tested the robustness, to the inclusion of ITV, of two commonly investigated relationships in ecology: those between FD indices and environmental gradients and those between FD indices and species richness.

Methods

We built a simulation study using abundance data from 46 observed communities (Fig. 1) and species mean trait values from field observations. In order to incorporate ITV we varied trait values within each species artificially around the known mean. We simulated various levels and structures of ITV (random vs environmentally-driven) based on previous

observations in this study area (Albert et al. 2010, Thuiller et al. 2010). As a result, community composition (presence and abundance of species) was observed from the field, while trait values for a given species varied artificially among communities in order to represent population-level ITV. From the simulated data we assessed FD using four different indices and made subsequent analyses of these calculated FD metrics (Fig. 1).

Field data

Study area and vegetation sampling
Plant communities were surveyed across a French alpine valley, called 'Guisane', that encompasses strong climatic gradients (44°5′ to 45°4′N; 6°21′ to 6°40′E, 1200 – 2600 m a.s.l., Albert 2010). The landscape is a mosaic of deciduous and evergreen coniferous forests dominated by *Larix decidua*, *Pinus uncinata* and *Pinus sylvestris*; heaths dominated by *Rhododendron ferrugineum*, *Vaccinium uliginosum* or *V. myrtillus*; subalpine grasslands dominated by *Bromus erectus*, *Trisetum flavesecens*, *Dactylis glomerata* or *Festuca paniculata*; and alpine meadows dominated by *Helictotrichon sedenense*, *Carex curvula*, *Kobresia myosuroides* or *Nardus stricta* (Cadel and Gilot 1963).

Communities were defined as 10 × 10 m plots with homogeneous vegetation. They were selected using a stratification defined by two uncorrelated direct (Austin 1987) gradients: mean minimal temperature in winter (hereafter temperature) and solar radiation in August (hereafter radiation), in order to encompass the broadest possible environmental heterogeneity (Albert et al. 2010). Exhaustive Braun-Blanquet relative abundance relevés (Braun-Blanquet 1946) were made within each community during the summer 2007 using a six-level cover scale, converted to relative abundance using median values of percentage cover classes (1%, 5%, 17%, 37%, 62% and 87%, Braun-Blanquet 1946). Among all the observed communities (174 communities, 442 species), we selected the ones that did not contain tree species (grasslands and shrub heaths) and for which we had trait data for at least 80% of the total cover (Pakeman and Quested 2007). The 46 communities contained a total of 91 'structuring' species representing 80 % of the total cover in each community (280 occurring species). Selected communities contained between 22 and 51 species each (hereafter NSp) and only 4 to 19 structuring species (hereafter NSpSt) and occurred between 1500 and 2700 m a.s.l. Each of the 91 structuring species occurred in 1 to 29 communities. The herbaceous communities considered are managed with low disturbance regimes including summer grazing by sheep, cattle or horses and mowing (see S1 in Albert et al. 2010 and Quétier et al. 2007 for more about current and past land use in the area).

Trait data
For the 91 structuring species, mean trait values were extracted from the alpine plants trait database ANDROSACE (Thuiller et al. unpublished) which compiles trait values from field measurements in the study area (including Choler 2005, Quétier et al. 2007, Albert et al. 2010, Lavorel et al. 2011) and from several other databases (VISTA, Garnier et al. 2007, LEDA, Knevel et al. 2003). We used two different traits that capture important axes of plant functional differentiation
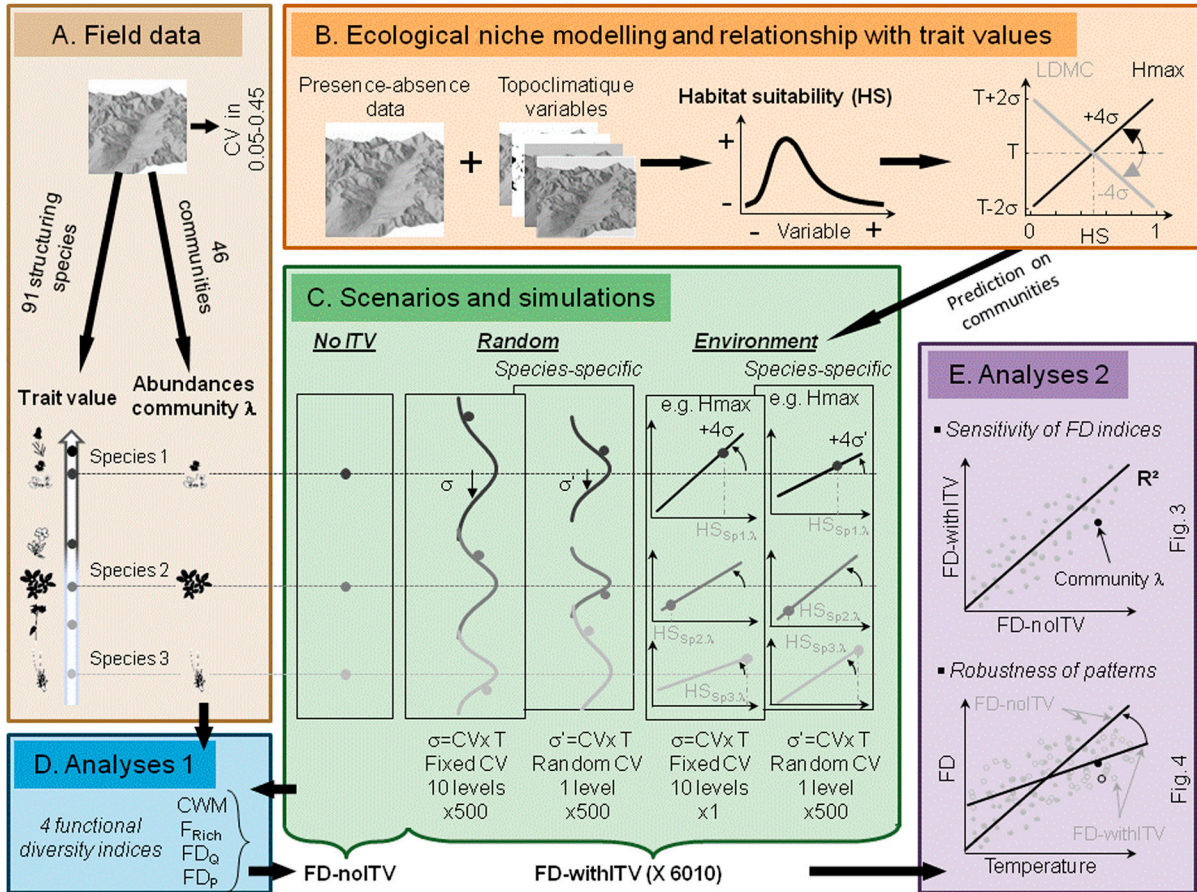
Figure 1. General framework for varying trait values artificially. (A) Field data. We used abundance data from 46 observed communities composed by 91 structuring species and species mean trait values from field observations (T = species mean trait values). Coefficients of variation (CV) orders of magnitude were based on previous observations in this study area (Albert et al. 2010). (B) Ecological niche modelling and relationship with trait values. We used presence/absence data as well as topo-climatic variables to derive a habitat suitability index for each structuring species within the study site by calibrating ecological niche models (ENM). For each species we prescribed a relationship between trait values and habitat suitability following the relationships revealed by Thuiller et al. (2010). These relationships were used to simulate environmentally-driven ITV in the 'Environment' scenarios. (C) Scenarios and simulations. In order to incorporate ITV we varied trait values within each species artificially around the known mean (σ = standard deviation). We simulated various levels (CV from 0.05 to 0.45) and structures of ITV following four main scenarios. The scenarios considered two different types of ITV structure, i.e. random versus environmentally-driven ITV, in order to depict trait's contingency to environmental gradients due to local adaptation or plasticity. Within each we considered two scenarios where ITV was similar for all species or variable across species ('species-specific') that is more ecologically realistic (Albert et al. 2010). As a result, community composition (presence and abundance of species) was observed from the field (A), while trait values for a given species varied artificially among communities in order to represent population-level ITV (B and C). (D) Analyses 1. From the simulated data we assessed FD using four different indices: CWM = community weighted mean. FRich = functional richness. FD_Q = Rao's quadratic entropy, FD_P = FD index from Petchey and Gaston. All four were calculated with (FD-withITV) and without ITV (FD-noITV). (E) Analyses 2. To assess the sensitivity of FD indices, we compared the FD-withITV and FD-noITV. The R^2 of the linear regression between both gives the percentage of information included in FD-noITV compared to FD-withITV. To assess the robustness of ecological patterns, we tested the effect of ITV on the relationships between the CWMs and the temperature gradient and between FD_P and specific richness. For each simulation, i.e. each scenario and each level of ITV, we fitted a linear regression between CWM (resp. FD_P, or log transformed for linearity) and temperature (resp. species richness) and extracted the slope and the corresponding p-value.

(Westoby et al. 2002). 1) Leaf dry matter content (LDMC, expressed in mg g^{-1}), i.e. the oven-dried mass of a leaf divided by its water-saturated fresh mass (Cornelissen et al. 2003), depicts the leaf economics spectrum (Wright et al. 2004) as it is related to the average density of leaf tissue. 2) Maximum vegetative height (Hmax, in cm), i.e. the distance between the top of photosynthetic tissue and the ground, is associated with plant competitive vigour and tends to be allometrically correlated with above-ground biomass (Cornelissen et al. 2003). We did not study the third recommended leading axis of plant functional differentiation, seed mass (Westoby

1998), for which our database was too incomplete, and because this trait is generally considered as invariable within species (but see Violle et al. 2009). This selection allowed us to compare two traits 1) with different levels of ITV (Albert et al. 2010), as Hmax is expected to be much more variable than LDMC within species (but see de Bello et al. 2010), and 2) that varied strongly across species and across communities (Fig. 2).

Simulation parameters were based on observed values. In the same study area, Albert et al. (2010) found on average (16 species) coefficients of variation around 0.26 for Hmax
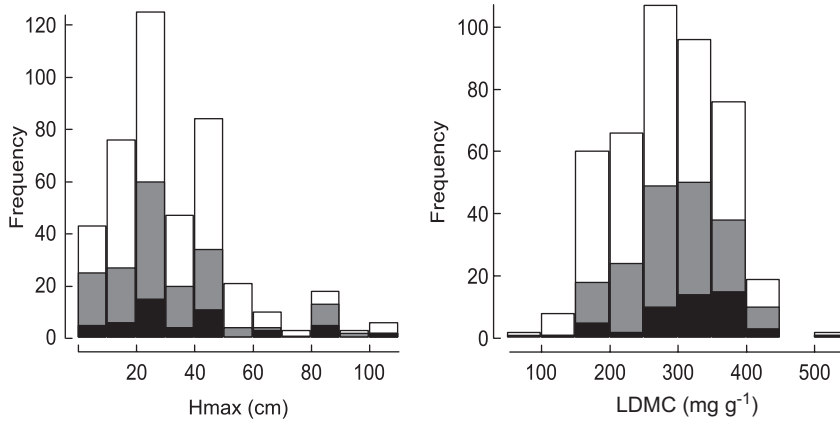
Figure 2. Trait values distribution within the regional pool. Distribution of species mean trait values within the regional pool (46 communities with a total of 436 occurrences): overall distribution (white), distribution for species with a relative abundance above 0.1 (50% of occurrences, light grey), distribution for species with a relative abundance above 0.18 (10% of occurrences, dark grey). For Hmax (left) and LDMC (right).

(approximately a normal distribution with SD = 0.07) and 0.09 for LDMC (approximately a normal distribution with SD = 0.03). These orders of magnitude also fit with the ones suggested by Cornelissen et al. (2003), which are 17–35% for Hmax and 4–10% for LDMC.

Ecological niche modelling and relationship with trait values

In order to simulate the influence of environment on population-level ITV, we used ecological niche models to estimate habitat suitability (HS) for each species and prescribed a relationship between functional trait values and habitats suitability following the relationships revealed by Thuiller et al. (2010).

Ecological niche models

For each of the 91 structuring species, we built ecological niche models (ENM) to extract the species' habitat suitability in the Guisane valley (Guisan and Thuiller 2005). Presence-absence data were extracted from the database of the Conservatoire Botanique National Alpin (577 community plots in the valley, independent of the 46 community plots detailed above). ENMs were calibrated using generalised additive models in an information theory approach (complete protocol in Thuiller et al. 2010) using three topoclimatic variables at 50 m resolution: annual degree days above 5°C, annual mean of daily solar radiation (kJ m^{-2} day^{-1}) and topographic wetness index. We then calculated species' habitat suitability index by standardizing the probabilities of occurrence from ENMs to have comparable values between species (Albert and Thuiller 2008). We reduced values to their 0.15–0.85 quantiles to avoid extreme values, centered them to 0.5 and rescaled them between 0 and 1. We finally extracted the predicted habitat suitability of each species in each of the 46 studied communities.

Environmentally-driven ITV

For each species we prescribed a relationship between trait values and habitat suitability following the relationships revealed by Thuiller et al. (2010). The relationships were as follow:

$$Hmax_{ic} = Hmax_i - 2\,\sigma_{Hmaxi} + 4\,\sigma_{Hmaxi}\,HS_{ic} \qquad (1)$$

$$LDMC_{ic} = LDMC_i + 2\,\sigma_{LDMCi} - 4\,\sigma_{LDMCi}\,HS_{ic} \qquad (2)$$

with $Hmax_{ic}$ (or $LDMC_{ic}$) and HS_{is} the trait value and the habitat suitability index of species i in community c, $Hmax_i$ (or $LDMC_i$) the mean trait value for species i in the ANDROSACE database, and σ_{Hmaxi} (or σ_{LDMCi}) the standard deviation of the trait from its mean for species i according to the different simulation scenarios defined in the next section. These equations represent within-species trait variation along the environmental gradients in a manner that is consistent with species environmental requirements (Albert et al. 2010, Wright et al. 2006). Trait values change differently for each species depending on how far the species is from its optimum. When the species grows in suitable areas, its individuals tend to have higher Hmax than average (respectively lower LDMC) and when it grows in unsuitable areas the contrary is expected, the mean trait value being obtained for HS = 0.5 (average suitability across the entire valley). The choice of an interval of variation [mean –2σ; mean + 2σ] allowed us to compare the 'Environment' and 'Random' scenarios, as 95% of random values following a normal distribution fall within this interval. We chose linear relationships, the simplest monotonic function, as the true relationship shape was unknown.

Scenarios and simulations: varying trait values artificially

In order to test the impact of the extent and structure of ITV on FD measurements, we developed four simulation scenarios. The scenarios considered two different types of ITV structure, i.e. random versus environmentally driven ITV, in order to depict trait's contingency to environmental gradients due to local adaptation or plasticity. Within each we considered two scenarios where ITV was similar for all species or variable across species ('species-specific') that is more ecologically realistic (Albert et al. 2010).

'Random': trait values were generated randomly for each species within each community. The trait value for the species i in the community c ($Trait_{ic}$) was a realization of the

normal law with a mean *Trait$_i$* (mean trait value of species *i*) and a standard deviation CV × *Trait$_i$*. CV varied between 0.05 and 0.45 (observed range, Albert et al. 2010) by steps of 0.05. This scenario corresponded to the hypothesis that population-level ITV was not structured by environmental gradients, all species being evenly variable.

'Random species-specific': as well as for the 'Random' scenario, trait values were generated randomly. The trait value for species *i* in the community *c* (*Trait$_{ic}$*) was a realization of the normal law with a mean *Trait$_i$* (mean trait value of species *i*) and a standard deviation CV$_i$ × *Trait$_i$*. CV$_i$ was randomly attributed to species following for Hmax a normal distribution with a mean 0.26 and a standard deviation (SD) 0.07 (as observed in the field, thus approximately between 0.12 and 0.40) and for LDMC a normal law with a mean 0.09 and a SD 0.03 (as observed in the field, thus approximately between 0.03 and 0.15). This scenario corresponded to the hypothesis that population-level ITV was not structured by environmental gradients, all species having a different variability of random magnitude.

'Environment': trait values were constrained by the habitat suitability index following Eq. 1 and 2 for Hmax and LDMC respectively. CV varied between 0.05 and 0.45 by steps of 0.05. This scenario corresponded to the hypothesis that population-level ITV was structured by species environmental requirements, all species being evenly variable.

'Environment species-specific': as for the 'Environment' scenario, trait values are constrained by HS following Eq. 1 and 2, but CVs are specific to each species: CV$_i$. CV$_i$ was randomly attributed to each species following for Hmax a normal distribution with a mean 0.26 and a standard deviation 0.07 (thus approximately between 0.12 and 0.40) and for LDMC a normal law with a mean 0.09 and a SD 0.03 (thus approximately between 0.03 and 0.15). This scenario corresponded to the hypothesis that population-level ITV was structured by species environmental requirements, all species being evenly variable.

The scenario 'Random' was run 500 times for each level of variability. The scenario 'Environment' was run once for each level of variability. The scenarios 'Random species-specific' and 'Environment species-specific' were run 500 times. An example of results for one run is given in supplementary material.

Analyses

Functional diversity indices
For each simulated dataset described above we evaluated FD within the communities using four commonly used indices:

1) Community weighted mean traits (hereafter CWM) represent the expected functional trait value of a random community sample and indicate the mean functional character of the community. CWM can be calculated for each trait as the mean of the trait values (*T$_{ic}$* for species *i*) in the community *c* containing S species, weighted by the relative abundance of the species (p$_{ic}$ for species i, Violle et al. 2007):

$$CWM_c = \sum_{i=1}^{S} p_{ic} T_{ic} \qquad (3)$$

2) Functional richness (hereafter FRich) corresponds to the volume of the functional space (range of trait values for single traits) occupied by the community. The larger the space, the richer the community. We chose here to evaluate relative functional richness for each community as the ratio between the community range and the regional range (volume of the functional space for the 46 studied communities) in order to allow comparisons between the different simulations.

3) Functional divergence characterizes the dispersion of trait values and associated abundances around the CWM. We evaluated this component of FD using Rao's quadratic entropy index (hereafter FD$_Q$) given by

$$FD_{Qc} = \sum_{i=1}^{S-1} \sum_{j=i+1}^{S} d_{ijc}\, p_{ic}\, p_{jc} \qquad (4)$$

where *dijc* is the dissimilarity between the *i*-th and *j*-th species (*dij* = *dji* and *dii* = 0) in community *c*. FD$_Q$ expresses the average dissimilarity between two randomly selected individuals with replacements (Rao 1982, Botta-Dukat 2005). The higher the abundances of species with extreme trait values, the higher FD$_Q$. Since we used only single quantitative continuous traits, we chose the Euclidian distance as a metric of dissimilarity.

4) The FD index from Petchey and Gaston (2002, hereafter FD$_P$), measures the extent of functional differences among species' trait values by estimating the dispersion of species in the trait space (Pavoine and Bonsall 2011). FD$_P$ measures the total branch length of the functional dendrogram based on species trait dissimilarities. It was calculated using the Xtree R function (J. Schumacher 2003) with Euclidian distance as the metric of dissimilarity.

We calculated these four indices on each single trait because traits are not evenly variable (Albert et al. 2010) and because this is a more rational approach when assessing the effects of ITV on FD (de Bello et al. 2010). Consequently, trait values were not standardized as in multi-trait calculations, especially given that standardizing values is not straightforward when the range of trait values is variable across simulations.

Statistical analyses
To estimate the possible bias in calculating FD indices without considering ITV, we compared the indices calculated with ITV (based on the four scenarios, FD-withITV) and without ITV (using only mean species' trait values, FD-noITV, Fig. 1 and Supplementary material Appendix A1). The R² of the linear regression between FD-noITV and FD-withITV gives the percentage of information included in FD-noITV compared to FD-withITV. The lower is the R², the less reliable is the estimation of FD using species' mean trait values as compared with including ITV into the calculation. In other words, low R² would give little support to any conclusion on community structure or ecosystem properties while ignoring ITV. Conversely, high R² meant that conclusions about response of communities to environmental gradients using

species' mean trait values are reliable and were explained by a change in community composition (species identity and abundance) rather than ITV.

FD measures are often related to other abiotic and biotic variables in order to test hypotheses on FD drivers. We aimed at testing the robustness of such relationships by including, or not, the effects of ITV. Among all possible relationships between FD indices and environmental gradients or species richness, we chose to analyse two combinations that have already been much investigated. First, we investigated the relationships between the CWMs and the temperature gradient (or altitude, Lavorel et al. 2011). CWMs are expected to respond to environmental gradients as mean properties of the community (Garnier et al. 2004) and quantifying their response is of interest for ecologists as it could lead to more predictive functional ecology (Sonnier et al. 2010). Second, we studied the relationship between FD_P and specific richness that is broadly described in the literature (Diaz and Cabido 2001, Petchey and Gaston 2002, 2006, Cianciaruso et al. 2009). Recent literature has questioned the independence between functional divergence and species richness and identifying causes and conditions for their correlation or independence is an essential question (Diaz and Cabido 2001). For each simulation, i.e. each scenario and each level of ITV, we fitted a linear regression between CWM (resp. FD_P or log transformed for linearity) and temperature (resp. species richness) and extracted the slope and the corresponding p-value (Fig. 1).

Results

Observed trait and abundance distributions

Among the communities, CWM-noITV for Hmax increased significantly with temperature ($p < 0.001$, $R^2 = 0.30$), with an average increase of 8 cm per °C. CWM-noITV for LDMC decreased marginally with temperature ($p = 0.09$, $R^2 = 0.06$), with an average decrease of 8 mg g^{-1} per °C (Fig. 4). The two traits had contrasted distributions. Species with small Hmax were more represented (positive skew), while the LDMC distribution was nearly centered (Fig. 2). Dominant species also tended to have slightly higher Hmax values than the mean and rather higher LDMC values than the mean (Fig. 2).

Sensitivity of different FD indices to ITV

In general, and as expected, increasing ITV led to decreased reliability of estimated FD-noITV indices (Fig. 3). The four FD indices were sensitive to the inclusion of ITV in the calculation, but not equally. The informative character of FD-noITV indices was however better conserved for CWM, and particularly less conserved for FD_P than for the FD_Q and FRich (Fig. 3). For instance with a CV of 0.2 for Hmax (low ITV), R^2 were above 0.8 for CWM, around 0.8 for FD_Q and FRich and around 0.65 for FD_P. This meant that the error made in estimating FD_P for Hmax while ignoring ITV was around 35% if the actual ITV is CV = 0.2. Similarly the error made in estimating FD_P for LDMC while ignoring ITV was more than 50% if the actual ITV is CV = 0.1.

The influence of ITV on the calculation of FD indices (difference between FD-noITV and FD-withITV) varied across the two investigated traits (Fig. 3). This influence was stronger and quicker for LDMC than for Hmax. While R^2 fell down to 0.4 for indices calculated with Hmax, they reached nearly 0 for indices calculated with LDMC. This meant that FD estimations became less reliable for LDMC than for Hmax as soon as a small amount of population-level ITV existed in the data.

The fundamental structure of population-level ITV, whether it was structured along the environmental gradients or random and whether it was evenly distributed across species or not (species-specific), had very contrasting consequences for FD calculations. While there were no significant differences between the scenarios 'Random' (with CVs of 0.25 for Hmax and 0.1 for LDMC) and 'Random species-specific', and between 'Environment' (with CVs of 0.25 for Hmax and 0.1 for LDMC) and 'Environment species-specific', strong differences emerged between the 'Random' structures and the 'Environment' scenarios. For the indices calculated with LDMC, the R^2 obtained in the 'Environment' scenario were far lower than the ones obtained in the 'Random' scenario. For the indices calculated with Hmax, the differences were less marked, R^2 obtained under the 'Environment' scenario being either lower (CWM) or higher (FD_Q and FRich) than the ones obtained under the 'Random' scenario.

Robustness of ecological patterns

The observed relationships of FD with environmental factors or species richness were altered when accounting for ITV (Fig. 4, 5, 6, 7). As expected from the previous results, relationships between FD metrics and environmental gradients and between FD metrics and species richness were more affected by the inclusion of ITV for LDMC than for Hmax. The positive relationship between CWM-Hmax and temperature was not altered when including ITV. The slope always remained positive and significant (Fig. 5). In contrast, the relationship between CWM-LDMC and temperature was strongly altered when including ITV (Fig. 5), the slope being negative, positive or non-significant depending on ITV's quantity and structure, and on simulation run. In particular the slope became positive with increasing ITV in the scenario 'Environment' although it was negative when using mean trait values, meaning that this relationship was not robust to ITV (Fig. 5). The relationships between FD_P (for both Hmax and LDMC) and species richness (NSp) were also altered when including ITV. The intercept increased with increasing ITV and slopes were either positive or non-significant (Fig. 6, 7).

Discussion

Sensitivity of functional diversity patterns to intraspecific variability

We conducted a sensitivity analysis of FD indices and their relationships with environmental gradients and species richness to population-level intraspecific trait variability. Our results suggest that ITV may strongly alter the quantification
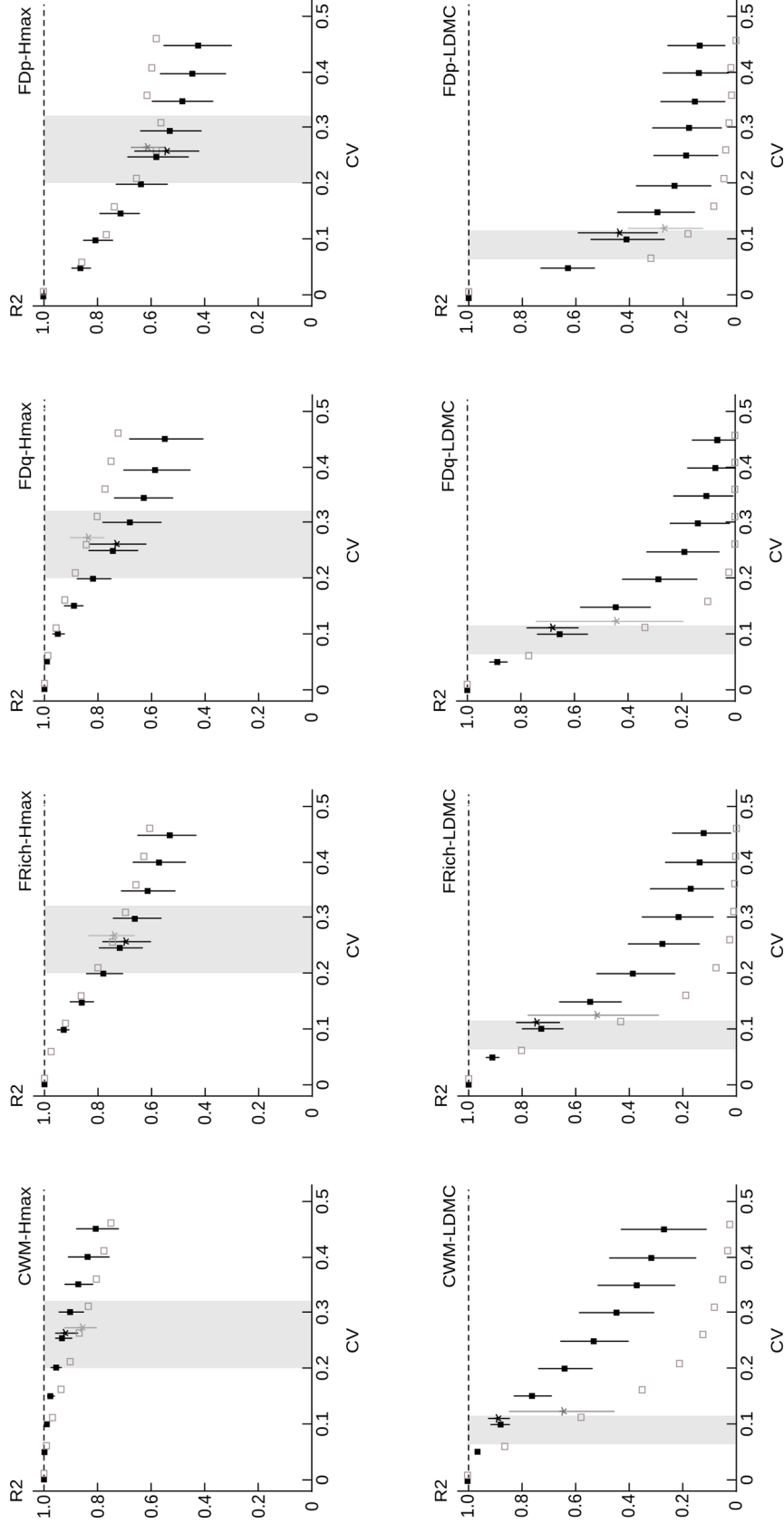
Figure 3. Sensitivity of different FD indices to ITV. R^2 of the relationships between FD-noITV and FD-withITV (Fig. 1) are presented as a function of increasing ITV (CV = coefficients of variation). Each panel represents a different index, from left to right: CWM = community weighted mean, FRich = functional richness, FD_q = Rao's quadratic entropy, FD_p = FD index from Petchey and Gaston (2002). The first row presents indices calculated for Hmax and the second for LDMC. The grey areas represent the 10–90% quantiles of the ITV observed in the field for each trait (Albert et al. 2010). Black squares represent the scenario 'random', grey open squares the scenario 'random specific', black stars the scenario 'environment' and grey stars the scenario 'environment specific'. The vertical segments represent the 10–90% quantiles of the results obtained from 500 runs. Symbols are slightly shifted to avoid they overlap.
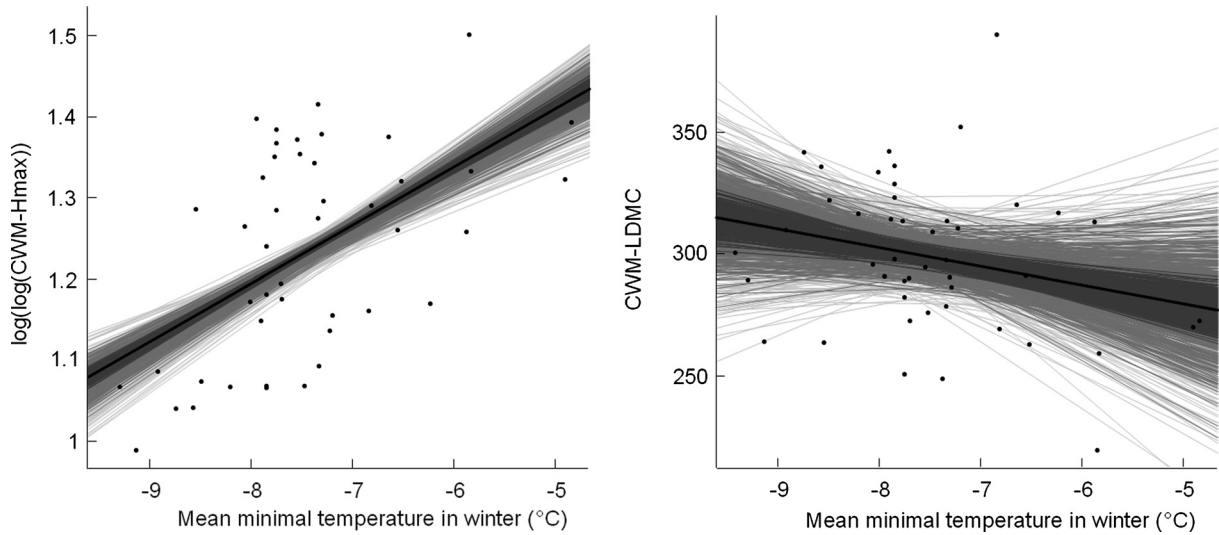
7

Figure 4. Robustness of the relationship between community weighted mean traits and temperature: resulting curves for the scenario 'Random'. Regression lines for the scenario 'random' and CV = 0.45 (light grey, 500 runs), CV = 0.3 (medium grey, 500 runs), CV = 0.1 (dark grey, 500 runs), CWM-noITV (black). Black dots illustrate the relationship CWM-noITV versus temperature. CWM calculated for Hmax (left) and LDMC (right).

of FD and the detection of patterns of ecological variation. Extending the scope of Cianciaruso et al. (2009), we tested several indices and revealed that they were not all as sensitive to ITV, and specifically that the single-trait FD from Petchey and Gaston (2002) was, generally, among the most sensitive of all. Our analysis also highlights three new findings concerning the inclusion of ITV in FD estimations.

First, we obtained contrasted results for single-trait indices calculated with Hmax and LDMC, both in terms of reliability of functional indices and robustness of ecological patterns. This is probably due to the fact that these traits were very differently distributed within the study communities and with respect to species abundances. Trait value distributions were

not normally distributed, and in particular trait value distributions for species with the highest abundances were even more skewed than values across all species (Fig. 2). Moreover, most abundant species had higher LDMC values on average and then more variable trait values when ITV was included (as SD = CV × mean), likely leading to higher sensibility of indices calculated with LDMC. Combining both traits for multi-trait indices calculations, we observed slower and lower decrease in R^2 (for FD_p, R^2 above 0.5 for random ITV, Supplementary material Appendix A1). Traits being differently distributed and differently variable in the field (Garnier et al. 2001, Cornelissen et al. 2003, Albert et al. 2010, de Bello et al. 2010), calculating multi-traits indices may 1) minimize
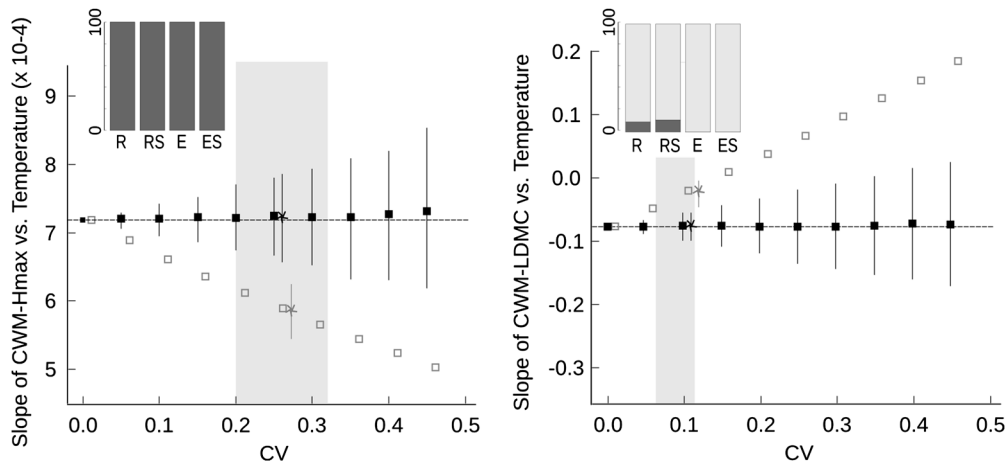

Figure 5. Robustness of the relationship between community weighted mean traits and temperature: synthesis of the results. Slopes of the relationship CWM-withITV versus temperature as a function of increasing ITV (CV = coefficients of variation). The dot line gives the slope of the relationship CWM-noITV versus temperature. The grey areas represent the 10–90% quantiles of the ITV observed in the field for each trait (Albert et al. 2010). Black squares represent the scenario 'random' (R), black stars the scenario 'random specific' (RS), grey open squares the scenario 'environment' (E) and grey stars the scenario 'environment specific' (ES). The vertical segments represent the 10–90% quantiles of the results obtained from 500 runs. The histograms give the percentage of significant slopes (at 0.05 in dark grey) vs. non significant (light grey). CWM calculated for Hmax (left) and for LDMC (right). Symbols are slightly shifted to avoid they overlap.
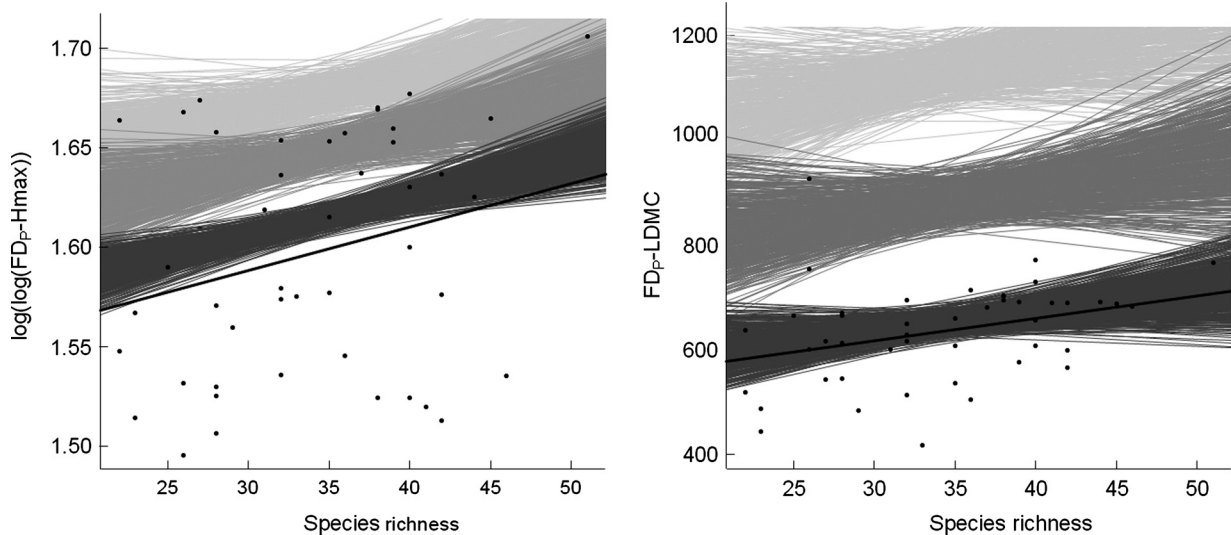
Figure 6. Robustness of the relationship between FD_P and species richness: resulting curves for the scenario 'Random'. Regression lines for the scenario 'random' and CV = 0.45 (light grey, 500 runs), CV = 0.3 (medium grey, 500 runs), CV = 0.1 (dark grey, 500 runs), FD_P-noITV (black). Black dots illustrate the relationship FD_P-noITV versus species richness. FD_P calculated for Hmax (left) and LDMC (right).

effects of ITV (as it is the case in our case study) or 2) lead to over-simplistic tests of the robustness of FD indices to ITV inclusion, in which all traits are treated as evenly variable.

Second, our analysis revealed a new and important feature: the way population-level ITV is structured within the environment (random vs environmental contingency), and not only the intensity of this variability, matters for the assessment of FD. The structuring of ITV by species environmental requirements ('Environment' and 'Environment specific' scenarios) enhanced ITV's effects on FD quantification and on the detection of ecological patterns in comparison with a random structuring. In this case there were thus synergistic effects on FD indices of 1) species presence, 2) their abundance, and 3) variation in their trait values among

communities (as they all depend differently on species environmental requirements).

Third, we found that all indices are not evenly sensitive to ITV. Our study is – to the best of our knowledge – the first systematic comparison of the robustness of different FD indices to ITV inclusion. Consistently with Lavorel et al. (2008), we found a low sensitivity of CMW to ITV. In contrast, we found that FD_P was highly sensitive to population-level ITV. This is not fully consistent with Cianciaruso et al. (2009) who found a low sensitivity of FD_P to ITV, but they considered individual-level and not population-level ITV and used a FD_P index that was based on multiple and not single traits. This high sensitivity of FD_P to population-level ITV also led to a high sensitivity of the relationships between
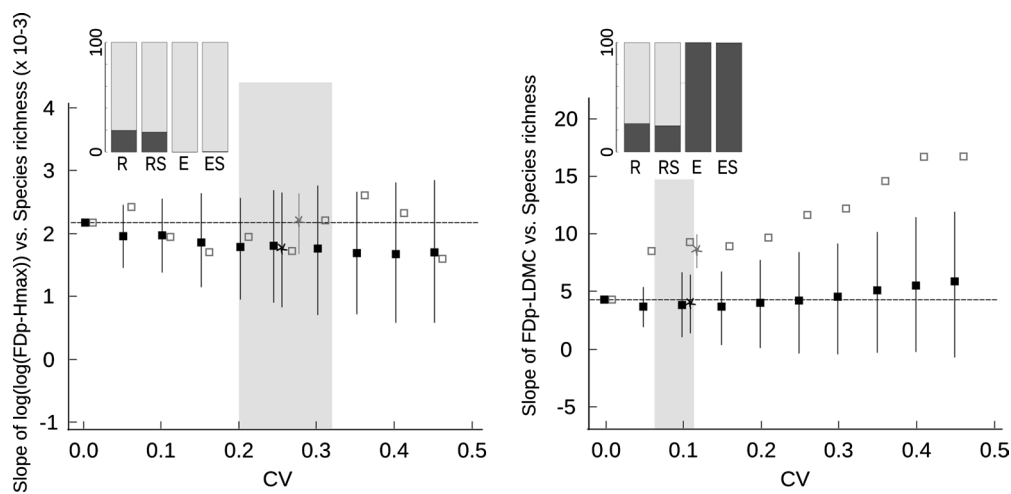

Figure 7. Robustness of the relationship between FD_P and species richness: synthesis of the results. Slopes of the relationship FD_P-withITV vs species richness as a function of increasing ITV (CV = coefficients of variation). The dot line gives the slope of the relationship FD_P-noITV vs. species richness. The grey areas represent the 10–90% quantiles of the ITV observed in the field for each trait (Albert et al. 2010). Black squares represent the scenario 'random' (R), black stars the scenario 'random specific' (RS), grey open squares the scenario 'environment' (E) and grey stars the scenario 'environment specific' (ES). The vertical segments represent the 10–90% quantiles of the results obtained from 500 runs. The histograms give the percentage of significant slopes (at 0.05 in dark grey) versus non significant (light grey). FD_P calculated for Hmax (left) and LDMC (right). Symbols are slightly shifted to avoid they overlap.

9

FD_P (for both Hmax and LDMC) and species richness. This interesting result should be interpreted carefully as our study communities contained between 22 and 51 species belonging to the local pool of 280 species, which was really far from the range of species richness studied in theoretical papers (2 to 25, Petchey and Gaston 2002, Cianciaruso et al. 2009).

Perspectives

Until recently, ITV was considered as negligible in trait-based approaches because it was assumed to be much lower than interspecific variability (1999, Garnier et al. 2001, Roche et al. 2004,). Recent studies have however shown that this assumption is not always met (de Bello et al. 2010, Messier et al. 2010) and even when met, ITV could still play a crucial role in determining FD. For instance, by using locally measured trait values, Jung et al. (2010) revealed habitat filtering mechanisms that were not detectable with mean trait values even with an ITV much lower than interspecific trait variation (at least three-fold lower). Consequently, omitting (vs accounting for) ITV in trait-based studies, in particular in FD assessments, requires first evaluating its effect on subsequent analyses (e.g. calculation with and without including it) and then deciding whether the results obtained without ITV are satisfactory or not.

Such evaluation can be done, as illustrated in our case study, by using virtual simulations (Zurell et al. 2010). We recommend the use of semi-virtual simulations, i.e. at least partially based on observed data, rather than 'fully' virtual experiments. From our results it seems indeed crucial 1) to use realistic trait distributions (skewed) and not necessarily normal distributions; 2) to keep realistic links between trait values and abundances (more abundant species with lower or higher trait values) because they might lead to synergistic effects on FD indices; 3) to use realistic orders of magnitude for ITV (e.g. CV) for different traits (Cornelissen et al. 2003, Albert et al. 2010); 4) to be sure that combining several traits with different levels of ITV is possible (traits are not evenly variable). Omitting these points might lead to erroneous conclusions or to results difficult to interpret. Simulation results can be used to decide whether to neglect ITV or not, depending on the expected reliability. This requires setting acceptable error levels as done in statistical tests (e.g. 5%). If the tolerated error level is exceeded, one will have to use local trait values and not means, such as those often obtained from data bases.

Several important issues emerged from the patterns of the simulated data. Based on Albert et al. (2010), we believed that the more realistic scenarios in our case study would be the ones that implied species-specific CVs. However, as these scenarios did not lead to significantly different results from the others, we could consider that scenarios without species-specific variability were sufficient in our case to conclude about the need to account for ITV or not. Furthermore, because environmental gradients explained only a small part of the variability in our study case (Albert et al. 2010), we believe that the 'Random' scenario is already a good approach for a test in this case. In addition, the CVs observed in the field (0.26 for Hmax and 0.09 for LDMC, Albert et al. 2010) include all levels of variability (inter-population, within-population, within-individual) and thus overestimate population-level variability alone (which

was simulated here). Within the ranges of ITV observed in the field, R^2 obtained for indices calculated with Hmax or LMDC (excepted for FD_P) were around or above 0.75 (for CWM under the random scenario, R^2 around 0.9). If we accept a 10% error as satisfactory, we should conclude that only CWMs calculated without including ITV (within the observed ranges of CVs) can be satisfactory.

Conclusions

Conclusions about FD and its role for community assembly and ecosystem functioning are being drawn from calculated indices and from relationships such as FD vs. environmental gradients (Sonnier et al. 2010). Such analyses have also been considered as the foundations for more predictive quantitative approaches to community assembly (Shipley et al. 2006, Sonnier et al. 2010, Violle et al. 2010). Our results suggest that the inclusion of ITV would improve such novel approaches that rely on FD indices. However, different FD indices and indices calculated with different traits are contrastingly sensitive to both extent and structure of ITV, making it difficult to delineate systematic rules to decide when to account for ITV. Testing a priori the robustness of the quantification of FD to ITV through simulation studies, as illustrated in this paper, appears consequently useful before engaging in the extensive field work required to quantify ITV for a given system. We believe that a more systematic evaluation of the effects of intraspecific trait variability on FD indices and their relationships to environmental variables or other community properties is warranted as part of the development of functional ecology.

References

Albert, C. H. and Thuiller, W. 2008. Favourability functions versus probability of presence: advantages and misuses. – Ecography 31: 417–422.

Albert, C. H. et al. 2010. Intraspecific functional variability: extent, structure and sources of variation. – J. Ecol. 98: 604–613.

Austin, M. P. 1987. Models for the analysis of species' response to environment gradients. – Vegetatio 69: 35–45.

Botta-Dukat, Z. 2005. Rao's quadratic entropy as a measure of functional diversity based on multiple traits. – J. Veg. Sci. 16: 533–540.

Braun-Blanquet, J. 1946. Über den Deckungswert der Arten in den Pflanzengesellschaften der Ordnung Vaccinio-Piceetalia. – Jahresber. Naturforsch. Ges. Graubündens 130: 115–119.

Cadel, G. and Gilot, J.-C. 1963. Documents pour la carte de la végétation des Alpes. – Ed. Fac. Sci. Grenoble I: 91–140

Choler, P. 2005. Consistent shifts in alpine plant traits along a mesotopographical gradient. – Arct. Antarct. Alp. Res. 37: 444–453.

Cianciaruso, M. V. et al. 2009. Including intraspecific variability in functional diversity. – Ecology 90: 81–89.

Cornelissen, J. H. C. et al. 2003. Handbook of protocols for standardised and easy measurement of plant functional traits worldwide. – Aust. J. Bot. 51: 335–308.

Cornwell, W. K. and Ackerly, D. D. 2009. Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. – Ecol. Monogr. 79: 109–126.

de Bello, F. et al. 2010. Quantifying the relevance of intraspecific trait variability for functional diversity. – Meth. Ecol. Evol. 2: 163–174

Diaz, S. and Cabido, M. 2001. Vive la difference: plant functional diversity matters to ecosystem processes. – Trends Ecol. Evol. 16: 646–655.

Diaz, S. et al. 2007. Incorporating plant functional diversity effects in ecosystem service assessments. – Proc. Natl Acad. Sci. USA 104: 20684–20689.

Garnier, E. et al. 2001. Consistency of species ranking based on functional leaf traits. – New Phytol. 152: 69–83.

Garnier, E. et al. 2004. Plant functional markers capture ecosystem properties during secondary succession. – Ecology 85: 2630–2637.

Garnier, E. et al. 2007. Assessing the effects of land-use change on plant traits, communities and ecosystem functioning in grasslands: a standardized methodology and lessons from an application to 11 European sites. – Ann. Bot. 99: 967–985.

Gravel, D. et al. 2006. Reconciling niche and neutrality: the continuum hypothesis. – Ecol. Lett. 9: 399–409.

Gross, N. et al. 2009. Linking individual response to biotic interactions with community structure: a trait based approach. – Funct. Ecol. 23: 668–679.

Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.

Hooper, D. U. et al. 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. – Ecol. Monogr. 75: 3–35.

Hubbel, S. P. 2001. The unified neutral theory of biodiversity and biogeography. – Princeton Univ. Press.

Hulshof, C. M. and Swenson, N. G. 2010. Variation in leaf functional trait values within and across individuals and species: an example from a Costa Rican dry forest. – Funct. Ecol. 24: 217–223.

Jung, V. et al. 2010. Intraspecific variability and trait-based community assembly. – J. Ecol. 98: 1134–1140.

Knevel, I. C. et al. 2003. Life-history traits of the northwest European flora: the LEDA database. – J. Veg. Sci. 14: 611–614.

Lavorel, S. et al. 2008. Assessing functional diversity in the field - methodology matters! – Funct. Ecol. 22: 134–147.

Lavorel, S. et al. 2011. Using plant functional traits to understand the landscape-scale distribution of multiple ecosystem services. – J. Ecol. 99: 135–137.

Lepš, J. et al. 2006. Quantifying and interpreting functional diversity of natural communities: practical considerations matter. – Preslia 78: 481–501.

Lepš, J. et al. 2011. Community trait response to environment: disentangling species turnover vs. intraspecific trait variability effects. – Ecography doi: 10.1111/j.1600-0587.2010.06904.x

Mason, N. W. H. et al. 2005. Functional richness, functional evenness and functional divergence: the primary components of functional diversity. – Oikos 111: 112–118.

Messier, J. et al. 2010. How do traits vary across ecological scales? A case for trait-based ecology. – Ecol. Lett. 13: 838–848.

Mouillot, D. et al. 2005. Functional regularity: a neglected aspect of functional diversity. – Oecologia 142: 353–359.

Naeem, S. and Wright, J. P. 2003. Disentangling biodiversity effects on ecosystem functioning: deriving solutions to a seemingly insurmountable problem. – Ecol. Lett. 6: 567–579.

Pakeman, R. and Quested, H. 2007. Sampling plant functional traits: what proportion of the species need to be measured? – Appl. Veg. Sci. 10: 91–96.

Pavoine, S. and Bonsall, M. B. 2011. Measuring biodiversity to explain community assembly: a unified approach. – Biol. Rev. doi:10.1111/j.1469-185X.2010.00171.x

Petchey, O. L. and Gaston, K. J. 2002. Functional diversity (FD), species richness and community composition. – Ecol. Lett. 5: 402–411.

Petchey, O. L. and Gaston, K. J. 2006. Functional diversity: back to basics and looking forward. – Ecol. Lett. 9: 741–758.

Petchey, O. L. and Gaston, K. J. 2007. Dendrograms and measuring functional diversity. – Oikos 116: 1422–1426.

Podani, J. and Schmera, D. 2006. On dendrogram-based measures of functional diversity. – Oikos 115: 179–185.

Quétier, F. et al. 2007. Linking vegetation and ecosystem response to complex past and present land use changes using plant traits and a multiple stable state framework. – Ecol. Monogr. 77: 33–52.

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

Roche, P. et al. 2004. Congruency analysis of species ranking based on leaf traits: which traits are the more reliable? – Plant Ecol. 174: 37–48.

Sandquist, D. R. and Ehleringer, J. R. 1997. Intraspecific variation of leaf pubescence and drought response in *Encelia farinosa* associated with contrasting desert environments. – New Phytol. 135: 635–644.

Schleuter, D. et al. 2010. A user's guide to functional diversity indices. – Ecol. Monogr. 80: 469–484.

Shipley, B. et al. 2006. From plant traits to plant communities: s statistical mechanistic approach to biodiversity. – Science 314: 812–814.

Sonnier, G. et al. 2010. Plant traits, species pools and the prediction of relative abundance in plant communities: a maximum entropy approach. – J. Veg. Sci. 21: 318–331.

Sugiyama, S. 2003. Geographical distribution and phenotypic differentiation in populations of *Dactylis glomerata* L. in Japan. – Plant Ecol. 169: 295–305.

Thuiller, W. et al. 2010. Variation in habitat suitability models does not always relate to variation in species' plant functional traits. – Biol. Lett. 6: 120–123.

Villéger, S. et al. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. – Ecology 89: 2290–2301.

Violle, C. et al. 2007. Let the concept of trait be functional! – Oikos 116: 882–892.

Violle, C. et al. 2009. Intraspecific seed trait variations and competition: passive or adaptive response? – Funct. Ecol. 23: 612–620.

Westoby, M. 1998. A leaf-heigh-seed (LHS) plant ecology strategy scheme. – Plant 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.

Wright, I. J. et al. 2004. The worldwide leaf economics spectrum. – Nature 428: 821–827.

Wright, J. W. et al. 2006. Experimental verification of ecological niche modeling in a heterogeneous environment. – Ecology 87: 2433–2439.

Zurell, D. et al. 2010. The virtual ecologist approach: simulating data and observers. – Oikos 119: 622–635.