Botta-Dukát Z; Czúcz B (2016) **Testing the ability of functional diversity indices to detect trait convergence and divergence using individual-based simulation**. - METHODS IN ECOLOGY AND EVOLUTION 7: (1) pp. 114-126. DOI: 10.1111/2041-210X.12450

Testing the ability of functional diversity indices to detect trait convergence and divergence using individual-based simulation

Running title: Detecting trait convergence and divergence

Author: Zoltán Botta-Dukát, Bálint Czúcz

Manuscript length: 8798 words

Address: MTA Centre for Ecological Research, Institute of Ecology and Botany, Alkotmány 2-4, Vácrátót, H-2163, Hungary.

E-mail: botta-dukat.zoltan@okologia.mta.hu

Summary

1. The quest for ‘assembly rules’, i.e. the processes shaping the species composition of communities, is a central issue in community ecology. Nevertheless, so far there is no general agreement on a framework to detect assembly rules in real life data: several key elements are still missing or heavily disputed, including the choice of the appropriate test statistic (e.g. functional diversity index) and randomization strategy for each major assembly process.
2. Simulation studies based on artificial communities can help to explore the usefulness of different approaches in detecting assembly rules. Nevertheless, the currently dominant approach to simulate artificial communities (i.e. selecting species from a pool based solely on trait values) oversimplifies the complex processes involved in community assembly and thus fails to produce realistic patterns. Consequently, its value for testing methodologies is seriously limited.
3. In this study we implemented a flexible, individual-based algorithm simulating real-life community processes (individuals are born, survive, compete for resources, reproduce and die), to generate artificial species composition data. With the help of this algorithm, we estimated the type I error rates and the statistical power of five different diversity indices (FRic, Rao’s quadratic entropy, FEve, the variance of functional distances, and the variance of nearest neighbor distances) in combination with three randomization strategies (randomization of trait values in the whole dataset, within plots and within the range of trait values occurring in each plot) for detecting two underlying assembly processes (habitat filtering and limiting similarity). We also tested the influence of all adjustable simulation parameters on the simulation results in a sensitivity analysis framework.
4. The results of the sensitivity analysis show that the individual-based simulation framework proposed here can be used for creating artificial community data with realistic pattern of

trait values. Based on the results, Rao’s quadratic entropy performed best for detecting both habitat filtering (trait convergence) and limiting similarity (trait divergence). Functional richness may also be suitable for detect traiting convergence. Functional evenness and variance of nearest neighbor distances, however, should not be used for finding assembly rules.

Keywords: assembly rules, Type I error rate, statistical power, functional richness, functional evenness, Rao’s quadratic entropy

Introduction

Understanding the rules of community assembly from a regional species pool is a central issue in community ecology (Keddy 1992). Assembly rules are constraints on species coexistence (Weiher *et al.* 2011; Götzenberger *et al.* 2012) that predict species presence and abundance in the local community (Keddy 1992). These constraints can be organized into a hierarchy of filters (Belyea & Lancaster 1999; Götzenberger *et al.* 2012). In community assembly studies focusing at a small area with negligible dispersal limitation, two such constraints are frequently considered. On one hand, individuals have to survive and reproduce under the given environmental conditions; this filter is often referred to as environmental (or habitat) filtering. On the other hand, species may be absent from suitable habitats due to interspecific competition, which forms the second filter. Limiting similarity theory (MacArthur & Levins 1967; Meszéna *et al.* 2006) predicts that species can only coexist if they are regulated differently (e.g. use different resources). The theory was originally developed in the context of resource competition, but there are several other potential stabilizing mechanisms (Chesson 2000; Wilson 2011) that can be based on differences in species attributes.

Different approaches proposed to detect these two filters (i.e. habitat filtering and limiting similarity) have been reviewed by Götzenberg et al. (2012). In the last few years, the trait-based approach became dominant in this field. Habitat filtering and limiting similarity influence the distribution of trait values in opposite ways (Mouillot *et al.* 2007; Cornwell & Ackerly 2009; Götzenberger *et al.* 2012). Habitat filtering leads to lower variation in trait values than random selection from the species pool (i.e. trait convergence) by excluding trait values not adapted to the local conditions (Figure 1). On the other hand, if coexisting species use different resources, they should differ markedly in the related traits, thus exhibiting more variance in trait values than would be expected for a random assembly (trait divergence).

Although trait convergence and divergence are two opposite patterns, habitat filtering and limiting similarity may act simultaneously (Weiher *et al.* 2011). When using a single test statistic with just one null-model, trait convergence and divergence are mutually exclusive outcomes (de Bello *et al.* 2012), and the lack of significant departure from the null-model may even indicate a balance between these two processes (Mason *et al.* 2008). Simultaneous effects of habitat filtering and limiting similarity can be detected only using more than one null model (e.g. Bernard-Verdier *et al.* 2012) or test statistic (Joner *et al.* 2012).

Although there is considerable evidence both for trait convergence and divergence, most of the tests (72% in the meta-analysis by Götzenberg et al. 2012) report no significant departure from the null-model. Possible reasons are that (i) the tested traits are neutral, (ii) the studied dataset is too small to detect departures from randomness or (iii) the applied test statics and/or null-models were inappropriate. Indeed, there is no consensus on which test statistic and null-model should be used, and the methods applied show large variation (see Appendix S1 for illustration).

Previous attempts to check the ability of functional diversity indices to distinguish between trait convergence and trait divergence were based on algorithms that select species from the species pool following trait-based rules (Mouchet *et al.* 2010; de Bello *et al.* 2012; Aiba *et al.* 2013; Mason *et al.* 2013). The disadvantage of this approach is that it does not try to simulate the real processes, but only aims at reproducing the expected pattern (trait convergence or divergence). There is only one study so far (Münkemüller et al. (2012) which applied a spatially explicit, individual-based modeling strategy to simulate the real underlying community processes: i.e. individuals are born, survive, compete for resources, reproduce and die. However, even this study applied a cellular-automaton simulation model with only one individual in each cell, which means that one of the key processes (limiting similarity) could not be tested.

The aim of this study is two-fold. We developed an individual-based simulation framework capable of testing methods and hypotheses regarding assembly rules and tested the general applicability of this framework. Secondly we evaluated the ability of functional diversity indices to detect habitat filtering and limiting similarity, using artificial data from the simulations.

Methods

Individual based simulation

We simulate the species composition of a set of locations along an environmental gradient using individual-based simulation (Black & McKane 2012). The simulation operates on an ecological time-scale; the regional species pool defined in the beginning does not change during the simulation. Each species is characterized by the values of three numerical traits: trait A is related to habitat matching, trait B regulates resource acquisition, while trait C is neutral. There is no within-species variation in trait values. Individuals compete for space and resources. Competition for space is strict: for each local community, the total number of individuals is limited, and a new individual can enter only after another resident has died. Competition between individuals depends on their similarity in traits related to resource acquisition; thus the competition is symmetric and it is strongest between conspecifics,. The slope of the “difference in trait B” vs.” strength of competition” curve depends on parameter σ_B (width of the competition kernel) with lower values resulting in a steeper slope, and if

σ_B=0, only individuals with the same trait values compete (in practice, this means that there is no interspecific competition). In addition to competition, the vital rates of a species are also determined by its how well it is adapted to the local conditions. The adaptedness is set to be different for different environmental conditions, depending on the difference between the position of local community along the environmental gradient and the value of trait A for the species. Thus each location along the environmental gradient favors a different value in trait A, and the survival probability of seedlings decreases with increasing difference between the actual and the locally optimal trait value. The speed of this decrease depends on the parameter σ_A (the strength of filtering) with lower values resulting in a steeper decrease. If $\sigma_A=\infty$, survival rates become constant. Local communities are not isolated, thus there is a continuous propagule inflow from the other local communities with a low rate. The model is not spatiality explicit, so the position of the individuals within the community does not influence competition, and the position of the communities does not influence propagule exchange. Nevertheless, as we consider each local community to be represented by one "plot" (a sample from a specific location at a specific position of the gradient) that contains only the entire local community, we use the terms "plot" and "local community" as synonyms in this study.

The simulation consists of a community initialization followed by an iterative simulation of a "disturbance-regeneration" cycle. The main steps of the simulation are illustrated in Figure 2, and explained in detail in Appendix S2. The parameters regulating the simulation algorithm and their values are shown in Table 1.

Functional diversity indices

Functional diversity is a complex concept, which is composed of three primary components: functional richness, functional divergence and functional evenness (Mason *et al.* 2005; Villéger *et al.* 2008). It is expected that habitat filtering and limiting similarity influence different components of functional diversity (Raevel *et al.* 2012): habitat filtering decreases the functional richness (by excluding non-adapted species) and functional divergence (due to higher abundance of optimally adapted species), while limiting similarity increases functional divergence and functional evenness by increasing the difference between dominant species (Figure 1). These expectations fit well into the theoretical framework developed by Boulangeat et al. (2012): species are first filtered by the abiotic environment which may be followed by further exclusions due to competition, and the abundance of occurring species are determined by competition and environmental filtering acting together.

We selected five indices for testing based on the comprehensive review of Pavoine and Bonsall (2011) so as to represent all three components of functional diversity: richness, divergence and evenness. We focused on indices calculated from distance matrices, as these are also applicable in phylogenetic studies, and offer a relatively straightforward way for incorporating intraspecific trait variation (de Bello *et al.* 2013a). Nevertheless, as there is no widely accepted distance-based measure of functional richness, and because it is known to be a good indicator of habitat filtering (e.g. Cornwell *et al.* 2006), we also examined the convex hull volume – which is thus the single metric in this study that is calculated directly from trait values. Information on the selected indices is summarized in Table 2.

Although distance-based indices can be used to quantify a multivariate functional diversity based on several traits at the same time, we still tested each trait separately. There are several reasons for this decision. Including all traits into a single analysis may hide existing patterns: opposite departures from randomness (i.e. convergence and divergence) in different traits may cancel each other out (Spasojevic & Suding 2012), while including neutral traits weakens the statistical tests (Butterfield & Suding 2013).

Null-models

Many different randomization algorithms have been used in trait-based assembly rules studies. Since the scope of the present study was testing functional diversity indices, not randomization algorithms, only three such algorithms were applied (Table 3). The first one involves reshuffling trait values among the species in the whole dataset containing all of the local communities (hereafter called "between-plots" randomization strategy). Note that this reshuffling is essentially equivalent to randomly drawing species from the pool of observed species (e.g. Cornwell *et al.* 2006; Montaña *et al.* 2014). The between-plot randomization strategy assumes a null-model that any species can occur in any local community with any abundance. This means that both habitat filtering and limiting similarity can cause departures from the random pattern.

In the second algorithm species traits are shuffled among the species occurring at each plot separately (hereafter called "within-plot" randomization strategy). This strategy relies on the inherent assumption that the occurrence of species is already determined by environmental filtering, but their abundance can still be shaped by interspecific competition. This strategy is furthermore equivalent to randomizing abundances among the species present, a null-model applied for example by Mason et al. (2008, 2013) and Pakeman et al. (2011). In the within plot randomization strategy the null hypothesis is that any of the occurring species could be dominant; functional divergence (i.e. larger than expected dissimilarities between the most abundant species) is a consequence of limiting similarity.

The third randomization strategy is reshuffling abundances within the environmentally filtered pool (hereafter called "restricted" randomization) that is, among species whose trait values fall within the range of observed trait values in a community (Cornwell & Ackerly 2009). This null-model hypothesizes that environmental filtering excludes species with trait values outside from this range, but the occurrence and abundance of the non-excluded species are independent from their trait values.

The distribution of the functional diversity indices under the null models were characterized by their values in 999 independent randomizations. One-sided tests were applied, which means that p-values (probability of type I error) were calculated separately for trait convergence (in the between-plot randomization strategy only) and divergence hypotheses (in all randomization strategies).

Simulation experiments

To test the effect of the different parameters on the simulated communities, we applied a sensitivity analysis consisting of 243 simulation experiments. These experiments were following a 3^{11-6} fractional factorial design constructed with the help of the R package *planor* (Kobilinsky *et al.* 2014). We used the parameter values shown in Table 1, which generally involved taking a baseline value, reducing it by 33% and increasing it by 50% thus forming a geometric series of three values for the sensitivity analysis. In the case of σ_A and σ_B, we only used the first three values in Table 1. We quantified all of the functional diversity indices for all of the null models, interpreting the type Ierror rates and the power of the tests (see below) as goodness metrics revealing the appropriateness of the simulation setup (parameter values) for use in evaluating functional diversity indices.

As an additional exercise to explore the consequences of no habitat filtering and/or competition in the simulation, we performed a further 4x4x3x3 full factorial experiment using all values of σ_A, σ_B, the number of species in the regional species pool (S) and the number of individuals in a local community (J), and the baseline values (i.e. first value in Table 1) of all other parameters from Table 1. The impact of the key parameters σ_A (the strength of environmental filtering) and σ_B (width of the competition kernel) on the simulation is further explored in Appendix S2.

Evaluation of the results of randomization tests

The randomization tests resulted in one p-value for each local community. The proportion of p-values lower than 5% was calculated for each meta-community. These proportions (interpreted as type I error or power depending on the parameters) were used in the subsequent analysis. We first checked whether the rate of type I errors (the proportion of significant results if there is no effect) is equal to the predefined significance level in two different ways: (1) by switching off filtering and/or competition by setting $\sigma_A = \infty$ or $\sigma_B = 0$ respectively, or (2) by testing the neutral trait the value of which did not influence the community assembly. Then we estimated the power of the test, i.e. the proportion of significant results, when there is a known direct effect. The effect of the simulation parameters on the type I error rate and power were explored by fitting classification tree models to the results from the fractional factorial experiment. We fitted conditional, inference-based classification trees for this purpose using the ctree algorithm implemented in the “party” package in R (Hothorn *et al.* 2006). The big advantage of ctree over most of the traditional classification and regression tree (CART) algorithms (e.g. Breiman *et al.* 1984) is that ctree offers unbiased variable selection and a statistically sound stopping rule (Hothorn *et al.* 2006), thus the fitted models offer an easily interpretable yet statistically sound ‘decision key’ for selecting the optimal algorithm in different study contexts. To offer insight into the selection of appropriate functional diversity indices, we also included the type of the test statistic as an explanatory variable into the ctree models. As the dependent variable, we used the proportion of power above 0.8, an arbitrary, but widely used threshold for power (like 0.05 for significance levels). Separate analyses were done for the three randomization methods. Similar ctree models were also fit to compare alternative methods for detecting the same process in order to guide future field-based studies. We used the difference in power as the dependent variable and parameters which can be determined from field data, including alpha- and beta-diversity, as predictors in these ctree models.

Results

General overview and Type I error rates

The communities constructed with the simulation algorithm exhibited general community structures very similar to real life plant communities: the mean species richness of local communities (alpha diversity) ranged from 4.7 to 57.7 (mean: 21.2), and the total number of species in the final simulation results (gamma diversity) was between 7 and 300 (mean: 89.6), which lead to beta-diversity values (i.e. the ratio of gamma- and alpha-diversity) ranging from 1.1 to 20.3 (mean: 4.635). The distributions of the traits in the simulation results differed from their distribution in the initial regional species pool in the expected way (Appendix S3).

For the in between- and within-plot randomizations, the type I error rates did not differ significantly from the predefined 5% threshold the vast majority of the test cases (Figure 3), while they often greatly exceeded the predefined threshold in restricted randomization. The cumulative distributions of type I error rates estimated in the two ways (i.e. by switching off both effects or by using neutral traits) did not differ considerably (Figure S4.1).

Power for detecting trait convergence due to habitat filtering

Only FRic and RaoQ had acceptable power in some (but not all) parameter combinations (Figure 4). The power of the other indices was generally low (i.e. below 0.8) for all combinations of the simulation parameters. Not surprisingly, the strength of habitat filtering in the simulated communities strongly influenced the power of the tests (Figure S4.2). Setting up the ctree model for *FRic* and *RaoQm* we found that the only simulation parameters which significantly influenced the detectability of trait convergence in the parameter space explored were the strength of filtering (σ_A), the width of the competition kernel (σ_B), and the correlation between traits A and B (c) (Figure 5). RaoQ performed much better than FRic, if habitat filtering is weak, while FRic has slightly higher power if the habitat filtering is strong. Using diversity values to predict differences between the power of the two functions, we found that RaoQ is preferable if local richness is high, while beta diversity is relatively low. Any correlation between the traits related to habitat filtering and competition decreases the detectability of habitat filtering, irrespective of its sign.

Power for detecting trait divergence due to competition

For trait divergence, we observed a generally low power of detection for all indices and randomization strategies. There was a low power of detection using the between-plot randomization strategy for trait divergence, and only *RaoQ* and the variance of differences (*Vd*) exceeded the threshold value in some cases (Figure 4.b). Not surprisingly, the parameters fine-tuning the competition process (i.e. σ_B and K) significantly influenced the detection rate, but the probability of reaching a high level of power remained low, even when competition was strong (Figure 5.b). *RaoQ* performed better than *Vd* if beta diversity was low (i.e. in a less heterogeneous environment where habitat filtering plays a minor role, Figure S4.4). Furthermore, the lack of habitat filtering significantly improved the power of *RaoQ* but not the other indices (Figure S4.5).

We also experienced low power levels for the within-plot randomization strategy with *RaoQ* being the only index which gave some above threshold cases (Figure 4.c). RaoQ was better at detecting intraspecific competition when the local communities were larger (high J, Figure 5.c).

In restricted randomization, *FRic* and *RaoQ* gave a relatively high power (Figure 4.d), but this came at the price of a high type I error rate, compromising the practical applicability of this randomization strategy.

Within plot randomization performed the best of the two randomization strategies, with acceptable type I error rates for *RaoQ.* It exhibited significantly higher overall power than the between-plot strategy (median difference is 0.06; $p < 0.001\%$, with a Wilcoxon paired rank sum test). Nevertheless, there were also several cases when between-plot randomization performed better (Figure S4.6), and the difference between the two methods could not be predicted by the diversity values.

Discussion

Functional richness and quadratic entropy indicate habitat filtering

Habitat filtering is expected to restrict the range of trait values and thus decrease the functional richness (Weiher *et al.* 1998; Cornwell *et al.* 2006). For this reason, habitat filtering is often considered to be indicated by lower than expected *FRic* values (e.g. Cornwell *et al.* 2006; Bernard-Verdier *et al.* 2012; Raevel *et al.* 2012). Our results pointed out that this relationship strongly depends on the environmental heterogeneity of the datasets studied (Willis *et al.* 2010; de Bello 2012): the power of the tests becomes low if the gradient sampled is short relative to the tolerance width of the species (Figure 5). With low environmental heterogeneity, trait values get filtered almost the same way in all plots. Thus there is no considerable difference to be expected in their plot level minimum and maximum values, and consequently in the actual and expected values of *FRic*.

As Rao's quadratic entropy is expected to combine functional richness and functional divergence (Mouchet *et al.* 2010), its use for detecting a decrease in functional richness due to habitat filtering is justified. According to Raevel et al. (2012), environmental filtering may influence not only the range of trait values, but also the position of dominant species in the trait space (functional divergence sensu Villéger *et al.* 2008). Rao's quadratic entropy is influenced by both effects, which may be advantageous when the aim is the detection of environmental filtering. And yet, even though *RaoQ* may also be influenced by limiting similarity, we found it highly appropriate for detecting habitat filtering, too. One possible explanation for the superior performance of *RaoQ* is that it is less sensitive to extreme trait values than *FRic* (Cornwell & Ackerly 2009). In a multi-trait analysis, the usage of *RaoQ* can also help to avoid the problems associated with convex hull volume discussed by Podani (2009).

Detecting limiting similarity remained an unresolved problem

None of the indices that we tested proved unequivocally appropriate for detecting limiting similarity. Even the best performing, Rao's quadratic entropy (*RaoQ*) and variance of distances (*Vd*) indices, had relatively low statistical power using any randomizations with reasonable type I error rates. The theoretical minimum value of *Vd* is attained if the species are placed equidistantly in the trait space, thus relatively low values can emerge due to interspecific competition which is supposed to homogenize the size of gaps among species in the trait space. The low power of this test statistic may accordingly be explained by the fact that *Vd* does not use the abundances, and even a small number of non-complying rare species may break the even spacing of trait values. Another possible explanation could be that the range of *Vd* is highly sensitive to the range of trait values (that is *FRic*), which is, on the other hand, highly influenced by habitat filtering (Appendix S5). Nevertheless, the fact that the power of *Vd* is not sensitive to the strength of habitat filtering, and RaoQ outperforms Vd at low beta diversity (i.e. in homogeneous environments where habitat filtering plays a minor role) contradicts the latter explanation.

As we have already discussed, *RaoQ* combines aspects of functional richness (i.e. the range of trait values) and functional divergence (i.e. the position of dominant species relative to the center of trait range). It seems highly improbable that limiting similarity would influence *RaoQ* values through altering functional richness (Cornwell & Ackerly 2009; Bernard-Verdier *et al.* 2012). Under what conditions can limiting similarity lead to high functional divergence? Functional divergence is high if abundant species are situated near the border of the occupied trait space. Limiting similarity predicts that dominant species are situated as far as possible from one another in the trait space. When considering only one trait, functional divergence is maximal if dominant species are situated at the two opposite ends of the trait gradient. This is in agreement with the pattern expected due to limiting similarity for two dominant species but contradicts the expected equidistant spacing of dominants if there are more than two dominant species. Note that increasing the dimensionality of the trait space can increase the number of dominant species that can be placed near the border and far from each other at the same time. Thus functional divergence measures (including *RaoQ*) probably perform better in multi-trait studies. The dimensionality of the trait space can be higher than the number of traits for nominal (multistate) traits or when the overlap of trait distribution is used as distance measure (de Bello *et al.* 2013a).

Contrary to what its name suggests, *FEve* (“functional evenness index”) was unsuitable for detecting limiting similarity. One possible explanation is that *FEve* considers only neighbors in trait space (true neighbors in the unidimensional space of our tests, which is generalized as a minimum spanning tree in the multidimensional case), while limiting similarity predicts that dominant species highly differ in traits (**Hiba! A hivatkozási forrás nem található.**Figure 6).

The variance (or standard deviation) of nearest neighbor distances (*Vnnd*) is another metric frequently used to detect limiting similarity in studies with real data, but which did not work well for this purpose in our simulations. This index is often successfully used to detect even spacing of species along a trait axis, which is interpreted as a result of limiting similarity (e.g. Cornwell & Ackerly 2009; Katabuchi *et al.* 2012). Nevertheless, *Vnnd* is similar to *FEve* in that only neighboring trait values are considered. Accordingly, many fundamentally different patterns may result in the same *Vnnd* value (Figure 6), which might explain the poor performance of this popular and seemingly well-suited metric. Based on our negative test results, we suggest that neither *Vnnd* nor *FEve* should be used to detect limiting similarity or trait divergence.

A quest for appropriate randomization strategies

All of the available randomization strategies test null hypotheses that slightly differ from the theoretical hypotheses that can be deduced from the processes (Figure 1). When testing for the effects of environmental filtering, the trait distribution in the local community should be compared with that of a community assembled randomly from the regional species pool. The between-plot strategy comes close to this, but it uses the pool of the species “observed” (i.e. set of species occurring in at least one plot) instead of the entire regional pool. If all local communities sampled are assembled from the same regional pool (i.e. the dispersal limitation is negligible), the pool of observed species is a subset of the regional pool. If a lot of local communities are sampled, and they cover the whole range of environmental variability within the region, the difference between the two sets is small. This small difference explains the high power of between-plot randomization to detect environmental filtering, when a wide range is sampled along the environmental gradient.

For testing limiting similarity, the actual trait distribution in the local community should be compared to that of a community randomly assembled from the “environmental” species pool, i.e. the species that can theoretically occur under the given environmental conditions. Between-plot randomization uses the set of species occurring in the whole dataset instead of the environmental species pool. If the dataset consists of local communities from different environments, it can contain a lot of species that do not belong to the environmental species pool. The more environmentally homogeneous the dataset is, the less difference there is between the two sets of species. Within-plot randomization, on the other hand, uses the list of species that actually occur in the tested community instead of environmental species pool. In contrary to between-plot randomization, this set is narrower than it should be (i.e. all species occurring in the local community belong to the environmental pool, but many of species from the environmental pool may be absent, including the species actually filtered out by competition). This is a significant shortcoming when using plot simulations to test for limiting similarity, which can explain why we saw low power with within-plot randomization. The detectability of limiting similarity could be improved by using a randomization strategy with a reference set closer to the environmental pool (Cornwell & Ackerly 2009; de Bello *et al.* 2012).

Restricted randomization aims at defining a more plausible environmental pool using the range of observed trait values as the reference set. However, this definition resulted in high type I error rate for indices influenced by range of traits, i.e. FRic and RaoQ. To understand this shortcoming, one should consider that the range of trait values can never be higher in the random communities created by restricted randomization than in the observed one. If the value of a functional diversity index depends on the trait range, this restriction obviously leads to an artificial “trait divergence” (i.e. higher observed value than mean of random values) even in neutral communities. Therefore we suggest that this way of creating an environmentally filtered pool for randomization should be avoided if possible. Nevertherless, this limitation is not important when the filtered pool is constructed through spatial scaling (Swenson *et al.* 2006) or based on species habitat preference (de Bello *et al.* 2012), or if species are selected with probabilities depending on the suitability of local habitat (Chalmandrier *et al.* 2013). Since the novel concept of dark diversity (Pärtel *et al.* 2011) is defined as the species absent from the local community but belong to the environmentally filtered pool, methods for estimating dark diversity, e.g. Beals smoothing (Ewald 2002; Botta-Dukát 2012), may also be useful for testing limiting similarity.

Detectability of assembly rules

In the meta-analysis of Götzenberg et al. (2012) only 358 of the 1966 tests reviewed (18%) showed significant departures from randomness. One possible reason for this low proportion is that the applied statistical tests have low power. Our simulations confirmed that some of the widely used test statistics have low power. Therefore we suggest that non-significant results found in such studies should be, if possible, re-analyzed using more powerful statistical tests.

64% of the significant tests examined by Götzenberg et al. (2012) had divergence of trait values that seems to contradict our results on the low power of tests dedicated to detecting trait divergence. However, a test level meta-analysis may be biased by case studies from homogeneous environments with a high number of tests, where the effect of environmental filtering is hardly detectable (examples of such studies are Stubbs & Wilson 2004; Mason & Wilson 2006). Aggregating results at the level of studies could show the opposite pattern. Merging the data collected by Emerson & Gillespie (2008: Table 1), Vamosi et al. (2009: Table 1) and HilleRisLambers et al. (2012: Table 1), 33 studies find trait or phylogenetic convergence, 11 studies find divergence, and 22 studies find both. We think these numbers mirror the higher detectability of environmental filtering, which is in line with our results.

A niche for individual-based simulations

There is a broad variety of methods developed for detecting assembly rules. As this study exemplifies, checking the statistical power of the methods is necessary even if they seem to be well-established theoretically. The individual-based simulation framework introduced in this paper can be a useful tool for such studies. The simulation presented in this paper was developed in R, a widely used high-level statistical scripting language, and the source code is available in Appendix S6.

This study focused on a very limited set of popular distance-based functional diversity indices. The broad spectrum of available test statistics is worth a deeper investigation with individual-based simulation models. Other functional or phylogenetic diversity indices (Pavoine & Bonsall 2011) or test statistics unrelated to the concept of functional diversity (e.g. Pillar *et al.* 2009; Shipley *et al.* 2012; de Bello *et al.* 2013b; Shipley 2014) are less widely used, but they could potentially highly outperform the ones studied in this paper. A systematic analysis of all available choices within a well-designed IBM environment would undoubtedly offer major methodological guidance for field studies.

We focusing primarily on the performance of the test statistics, so we applied only three widely used randomization strategies. There are several more possible randomization approaches, an exhaustive testing of which would also be necessary, with special regard for methods appropriate for detecting limiting similarity.

The individual-based simulation model presented in this paper is an attempt at constructing a minimal in silico representation of two major processes shaping the composition of ecological communities: habitat filtering and limiting similarity. Nevertheless, such a model is necessarily incomplete, and the model can be refined for producing a more realistic representation at the price of adding more complexity to the simulation. In the present version, there is no within-species variation in trait values, but this property can be implemented in a relatively simple and straightforward way within this framework. If individuals differ not only in their traits within the species, but this difference is also heritable, the simulation framework can potentially be made suitable for studying evolutionary processes.

Beyond environmental filtering, there may be alternative processes that lead to trait convergence. The most important such processes are asymmetric competition (Mayfield & Levine 2010) and dispersal limitation (Münkemüller *et al.* 2012). Both of these complexities can potentially be incorporated into this simulation framework. Competition can be made asymmetric by replacing Equation 3 in Appendix S2 with the formula of Kisdi (Kisdi 1999), whereas spatially limited dispersal can be generated relatively easily after setting up a spatial neighborhood or distance matrix for the local communities in step 4 (Appendix S2). Nevertheless, the power of methods for detecting the effect of spatial pattern of environment and limited dispersal were exhaustively studied by Münkemüller et al. (2012) using cellular automaton simulation, an approach which might be better suited to studying spatial processes, but which cannot easily handle limiting similarity with only one individual living in each cell.

In addition to checking the applicability of methods for detecting assembly rules, a similar simulation framework could also be useful for exploring the factors influencing the alpha-, beta- and gamma-diversity of artificial communities. Such exploration would give new insight into the relative importance of stochastic and deterministic processes in community organization, which has been a highly debated subject in the last few years (Chase & Myers 2011; Gravel *et al.* 2011; Rosindell *et al.* 2012; Vellend *et al.* 2014).

Conclusions

Of the functional diversity indices studied, Rao's quadratic entropy seems to be most suitable for testing for both trait convergence (due to environmental filtering) and trait divergence (due to limiting similarity). Environmental filtering can be detected relatively reliably using the between-plot randomization strategy, if the data set covers a wide range of environmental conditions. If an environmentally homogeneous area was sampled, however, trait convergence can become impossible to detect without using external information on the regional species pool, because the species "filtered out" by the unsuitable environment are absent from the entire dataset.

None of the combinations of diversity indices and randomization strategies tested can reliably detect trait divergence due to limiting similarity under all conditions. . However, if there is a lack of habitat filtering (i.e. if the dataset is environmentally homogeneous), Rao's quadratic entropy combined with a between-plot randomization strategy can detect this process.

For existing datasets, the ability to detect habitat filtering may be improved by increasing the environmental heterogeneity of the dataset. On the other hand, analyzing data from a narrow range of environmental variables increases the detection probability of limiting similarity.

In the light of the outcomes of our simulations, non-significant results in studies with real-data should be interpreted very carefully. In addition to the lack of the effect tested, negative results can occur in several ways ranging from the masking of limiting similarity by environmental filtering to the inability of the chosen index/test to detect an existing effect. As we have shown, several seemingly very well-suited and commonly used indices might be useless under experimental circumstances with artificial data.

Acknowledgement

This research was supported by OTKA K83595 and K91180 research grants. Bálint Czúcz was supported by the János Bolyai research fellowship of the Hungarian Academy of Sciences. We thank Francesco de Bello, Lars Götzenberg, Imelda Somodi, and three anonymous reviewers for their valuable comments, and Emily Rauschert for improving English language of the article.

Data accessibility

No real data were used in this paper. Simulation algorithm is available in form of R script in the Supporting Information.

Literature

Aiba, M., Katabuchi, M., Takafumi, H., Matsuzaki, S.S., Sasaki, T. & Hiura, T. (2013). Robustness of trait distribution metrics for community assembly studies under the uncertainties of assembly processes. *Ecology*, **94**, 2873–2885.

de Bello, F. (2012). The quest for trait convergence and divergence in community assembly: are null-models the magic wand? *Global Ecology and Biogeography*, **21**, 312–317.

de Bello, F., Carmona, C.P., Mason, N.W.H., Sebastia, M.-T. & Leps, J. (2013a). Which trait dissimilarity for functional diversity: trait means or trait overlap? *Journal of Vegetation Science*, **24**, 807–819.

de Bello, F., Price, J.N., Münkemüller, T., Liira, J., Zobel, M., Thuiller, W., Gerhold, P., Götzenberger, L., Lavergne, S., Lepš, J., Zobel, K. & Pärtel, M. (2012). Functional species pool framework to test for biotic effects on community assembly. *Ecology*, **93**, 2263–2273.

de Bello, F., Vandewalle, M., Reitalu, T., Lepš, J., Prentice, H.C., Lavorel, S. & Sykes, M.T. (2013b). Evidence for scale- and disturbance-dependent trait assembly patterns in dry semi-natural grasslands. *Journal of Ecology*, **101**, 1237–1244.

Belyea, L.R. & Lancaster, J. (1999). Assembly rules within a contingent ecology. *Oikos*, **86**, 402–416.

Bernard-Verdier, M., Navas, M.-L., Vellend, M., Violle, C., Fayolle, A. & Garnier, E. (2012). Community assembly along a soil depth gradient: contrasting patterns of plant trait convergence and divergence in a Mediterranean rangeland. *Journal of Ecology*, **100**, 1422–1433.

Black, A.J. & McKane, A.J. (2012). Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution*, **27**, 337–345.

Botta-Dukát, Z. (2012). Co-occurrence-based measure of species' habitat specialization: robust, unbiased estimation in saturated communities. *Journal of Vegetation Science*, **23**, 201–207.

Botta-Dukát, Z. (2005). Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science*, **16**, 533–540.

Boulangeat, I., Gravel, D. & Thuiller, W. (2012). Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters*, **15**, 584–593.

Breiman, L., Friedman, J., Ohlsen, R. & Stone, C. (1984). *Classification and regression trees*. Chapman Hall/CRC Press, New York.

Butterfield, B.J. & Suding, K.N. (2013). Single-trait functional indices outperform multi-trait indices in linking environmental gradients and ecosystem services in a complex landscape. *Journal of Ecology*, **101**, 9–17.

Chalmandrier, L., Münkemüller, T., Gallien, L., de Bello, F., Mazel, F., Lavergne, S. & Thuiller, W. (2013). A family of null models to distinguish between environmental filtering and biotic interactions in functional diversity patterns. *Journal of Vegetation Science*, **24**, 853–864.

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

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, **31**, 343–366.

Clarke, K. & Warwick, R. (2001). A further biodiversity index applicable to species lists: variation in taxonomic distinctness. *Marine Ecology-Progress Series*, **216**, 265–278.

Cornwell, W.K. & Ackerly, D.D. (2009). Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecological Monographs*, **79**, 109–126.

Cornwell, W.K., Schwilk, D.W. & Ackerly, D.D. (2006). A trait-based test for habitat filtering: convex hull volume. *Ecology*, **87**, 1465–1471.

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

Ewald, J. (2002). A probabilistic approach to estimating species pools from large compositional matrices. *Journal of Vegetation Science*, **13**, 191–198.

Götzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A., Lepš, J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel, K. & Zobel, M. (2012). Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biological Reviews*, **87**, 111–127.

Gravel, D., Guichard, F. & Hochberg, M.E. (2011). Species coexistence in a variable world. *Ecology Letters*, **14**, 828–839.

HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M. & Mayfield, M.M. (2012). Rethinking Community Assembly through the Lens of Coexistence Theory. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 227–248.

Hothorn, T., Hornik, K. & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, **15**, 651–674.

Joner, F., Anand, M. & Pillar, V. (2012). Trait-convergence and divergence assembly patterns in a temperate forest herbaceous layer along the gradient of canopy closure. *Community Ecology*, **13**, 178–184.

Katabuchi, M., Kurokawa, H., Davies, S.J., Tan, S. & Nakashizuka, T. (2012). Soil resource availability shapes community trait structure in a species-rich dipterocarp forest. *Journal of Ecology*, **100**, 643–651.

Keddy, P.A. (1992). Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science*, **3**, 157–164.

Kisdi, É. (1999). Evolutionary Branching under Asymmetric Competition. *Journal of Theoretical Biology*, **197**, 149–162.

Kobilinsky, A., Bouvier, A. & Monod, H. (2014). *PLANOR: an R package for the automatic generation of regular fractional factorial designs. R package*.

Laliberté, E. & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, **91**, 299–305.

MacArthur, R. & Levins, R. (1967). The Limiting Similarity, Convergence, and Divergence of Coexisting Species. *The American Naturalist*, **101**, 377–385.

Mason, N.W.H., de Bello, F., Mouillot, D., Pavoine, S. & Dray, S. (2013). A guide for using functional diversity indices to reveal changes in assembly processes along ecological gradients. *Journal of Vegetation Science*, **24**, 794–806.

Mason, N.W.H., Lanoiselée, C., Mouillot, D., Wilson, J.B. & Argillier, C. (2008). Does niche overlap control relative abundance in French lacustrine fish communities? A new method incorporating functional traits. *Journal of Animal Ecology*, **77**, 661–669.

Mason, N.W.H., Mouillot, D., Lee, W.G. & Wilson, J.B. (2005). Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos*, **111**, 112–118.

Mason, N.W.H. & Wilson, J.B. (2006). Mechanisms of species coexistence in a lawn community: mutual corroboration between two independent assembly rules. *Community Ecology*, **7**, 109–116.

Mayfield, M.M. & Levine, J.M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, **13**, 1085–1093.

Meszéna, G., Gyllenberg, M., Pásztor, L. & Metz, J.A.J. (2006). Competitive exclusion and limiting similarity: A unified theory. *Theoretical Population Biology*, **69**, 68–87.

Montaña, C.G., Winemiller, K.O. & Sutton, A. (2014). Intercontinental comparison of fish ecomorphology: null model tests of community assembly at the patch scale in rivers. *Ecological Monographs*, **84**, 91–107.

Mouchet, M.A., Villéger, S., Mason, N.W.H. & Mouillot, D. (2010). Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. *Functional Ecology*, **24**, 867–876.

Mouillot, D., Dumay, O. & Tomasini, J.A. (2007). Limiting similarity, niche filtering and functional diversity in coastal lagoon fish communities. *Estuarine, Coastal and Shelf Science*, **71**, 443–456.

Münkemüller, 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.

Pakeman, R., Lennon, J. & Brooker, R. (2011). Trait assembly in plant assemblages and its modulation by productivity and disturbance. *Oecologia*, **167**, 209–218.

Pärtel, M., Szava-Kovats, R. & Zobel, M. (2011). Dark diversity: shedding light on absent species. *Trends in Ecology & Evolution*, **26**, 124–128.

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

Pillar, V.D., Duarte, L. da S., Sosinski, E.E. & Joner, F. (2009). Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. *Journal of Vegetation Science*, **20**, 334–348.

Podani, J. (2009). Convex hulls, habitat filtering, and functional diversity: mathematical elegance versus ecological interpretability. *Community Ecology*, **10**, 244–250.

Raevel, V., Violle, C. & Munoz, F. (2012). Mechanisms of ecological succession: insights from plant functional strategies. *Oikos*, **121**, 1761–1770.

Rosindell, J., Hubbell, S.P., He, F., Harmon, L.J. & Etienne, R.S. (2012). The case for ecological neutral theory. *Trends in Ecology & Evolution*, **27**, 203–208.

Shipley, B. (2014). Measuring and interpreting trait-based selection versus meta-community effects during local community assembly. *Journal of Vegetation Science*, **25**, 55–65.

Shipley, B., Paine, C.E.T. & Baraloto, C. (2012). Quantifying the importance of local niche-based and stochastic processes to tropical tree community assembly. *Ecology*, **93**, 760–769.

Spasojevic, M.J. & Suding, K.N. (2012). Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. *Journal of Ecology*, **100**, 652–661.

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.

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.

Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleynhans, E.J., Kraft, N.J.B., Letaw, A.D., Macdonald, A.A.M., Maclean, J.E., Myers-Smith, I.H., Norris, A.R. & Xue, X. (2014). Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos*, **123**, 1420–1430.

Villéger, S., Mason, N.W.H. & Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, **89**, 2290–2301.

Weiher, E., Clarke, G.D.P. & Keddy, P.A. (1998). Community Assembly Rules, Morphological Dispersion, and the Coexistence of Plant Species. *Oikos*, **81**, 309–322.

Weiher, E., Freund, D., Bunton, T., Stefanski, A., Lee, T. & Bentivenga, S. (2011). Advances, challenges and a developing synthesis of ecological community assembly theory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2403–2413.

Willis, C.G., Halina, M., Lehman, C., Reich, P.B., Keen, A., McCarthy, S. & Cavender-Bares, J. (2010). Phylogenetic community structure in Minnesota oak savanna is influenced by spatial extent and environmental variation. *Ecography*, **33**, 565–577.

Wilson, J.B. (2011). The twelve theories of co-existence in plant communities: the doubtful, the important and the unexplored. *Journal of Vegetation Science*, **22**, 184–195.

Table 1: Parameters of the individual based simulation and the values used in the sensitivity analysis

Name	**Description**	**Step**	**Values explored**	**Remark**
S	number of species in the regional species pool	all	200, 133, 300	
N	number of local communities	all	50, 33, 75	
J	number of individuals in a local community	all	300, 200, 450	
d	parameter of the symmetric beta distribution used for generating trait values	1	1, 0.5, 1.5	these values give uniform (d=1), arcsine (0.5), and Wigner semicircle distributions (1.5)
c	parameter for setting the correlation between traits A and B	1	0, -6, 0.6	see equation (1) in Appendix S2, at c = 0 there is no correlation, whereas c = ±.6 means a correlation of ~±.83
R	the fraction of the environmental gradient covered by the local communities	1	0.8, 0.6, 0.9	
σ_A	tolerance width	1, 5	0.05, 0.01, 0.25, ∞	determines habitat filtering, with lower values meaning more specialist species (see equation 2 and figure S2.1 in Appendix S2). At $\sigma_A = \infty$ species are maximally generalist (abiotic conditions do not influence their abundance).
σ_B	scaling parameter of the competition strength	3	0.05, 0.01, 0.25, 0	determines limiting similarity, with higher values meaning stronger competition (see equations 4 and 5 and figure S2.1 in Appendix S2) At $\sigma_B = 0$ there is no interspecific competition (no effect of trait B on competition).
b	probability of reproduction at no competition	3	1	see equation (3) in Appendix S2
K_0	critical level of competitive pressure	3	200, 133, 300	see equation (3) in Appendix S2
m	probability of dispersing seeds to another local community	4	0.1, 0.066, 0.15	the target community is selected randomly (without regard to spatial location or closeness)

sim.len	the length of the simulation	100, 66, 150	the total number of iterations is sim.len*J

Table 2: Overview of the tested functional diversity indices

Name	**Description**	**Measured component**[1]	**Abundances used?**	**References**
FRic	Convex hull volume (or range in the one-dimensional case)	functional richness	No	(Cornwell *et al.* 2006; Villéger *et al.* 2008)
RaoQ	Rao's quadratic entropy (=Mean distance among species weighted by species abundance)	functional richness and functional divergence	Yes	(Botta-Dukát 2005)
FEve	Evenness in the branch lengths of a minimum spanning tree	functional evenness	Yes	(Villéger *et al.* 2008)
Vd	Variance of distances	functional evenness	No	(Clarke & Warwick 2001)
Vnnd	Variance in the nearest-neighbor distances	functional evenness	No	(Weiher *et al.* 1998)

[1] sensu Laliberte & Legendre 2010

Table 3: Overview of the alternative hypotheses applied. For calculating type I error rate both habitat filtering and limiting similarity were switched off by appropriate parameter settings or using neutral trait. For testing power, only the trait expected to converge/diverge was used in the calculations.

Randomization	test statistic	Power for detecting convergence	Power for detecting divergence	Type I error rate
Between-plot	FRic RaoQ FEve	one sided observed<random	one sided observed>random	two sided
	Vd Vnnd	one sided observed>random	one sided observed<random	
Within-plot	RaoQ FEve	not applied	one sided observed>random	one sided observed>random
Restricted	FRic	not applied	one sided observed>random	one sided observed>random
	RaoQ			
	FEve			
	Vd		one sided observed<random	one sided observed<random
	Vnnd			

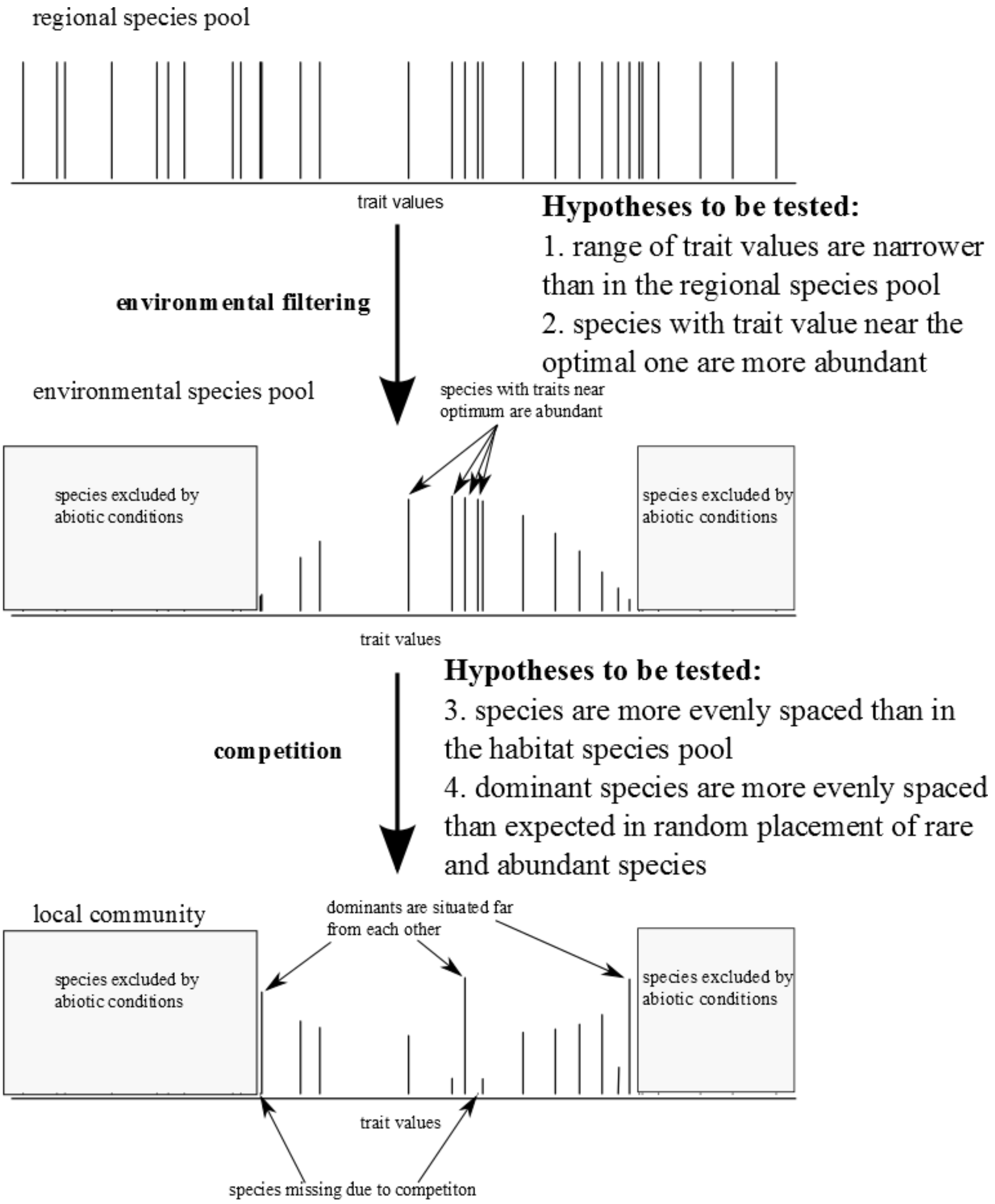


Figure 1: Expected changes in the distribution of trait values due to habitat filtering and competition, and the related hypotheses to be tested. Since the regional and the environmental species pool are often not known, they are substituted by the pool of observed species (in the between-plot randomization strategy) or occurring just in the local plot (in the within-plot strategy).

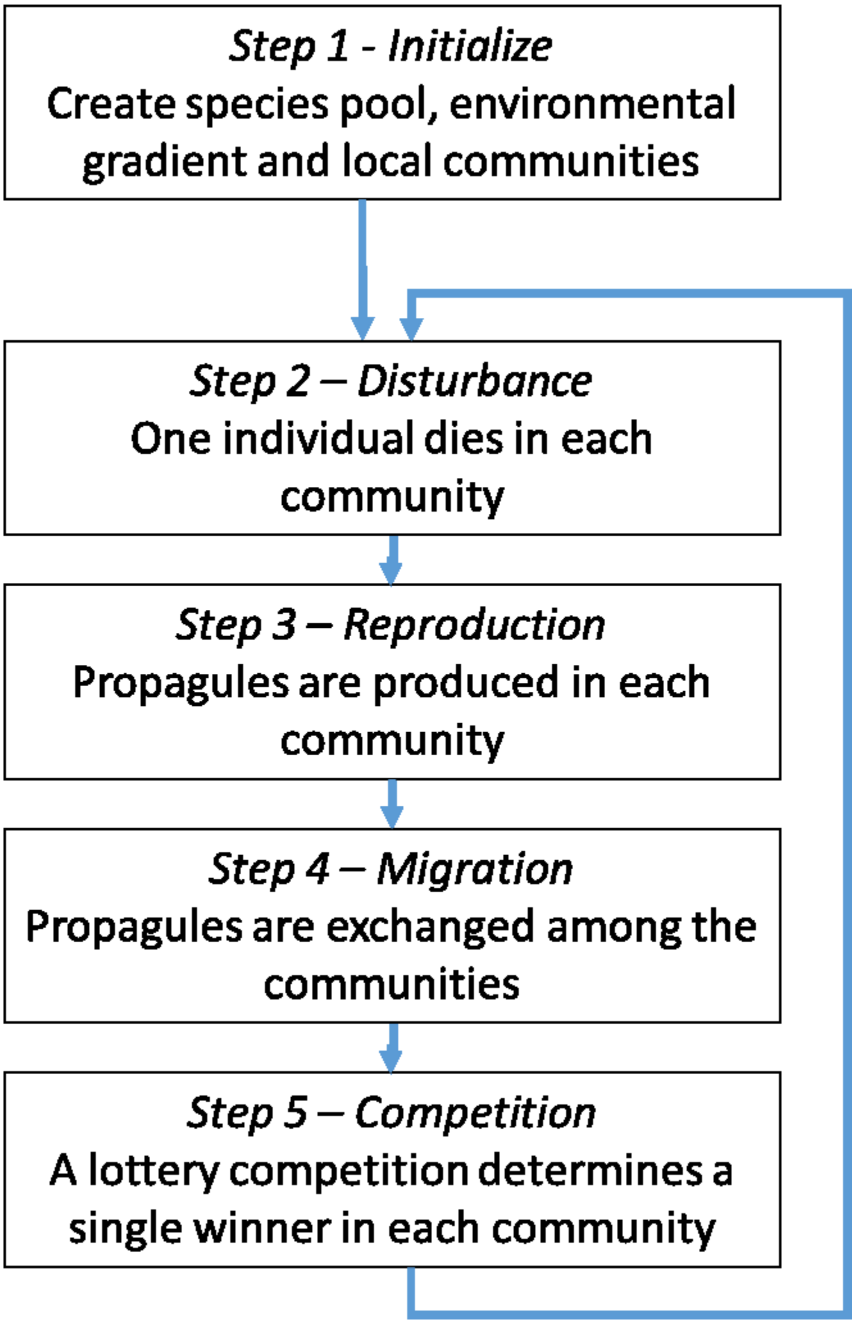
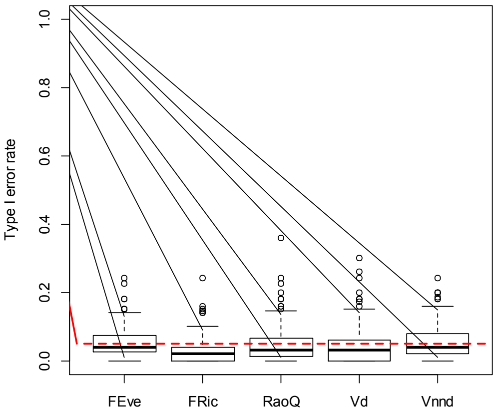
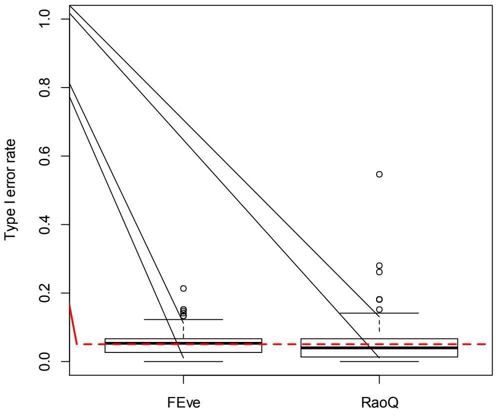


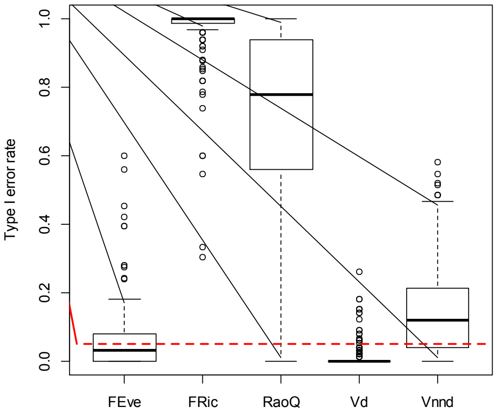
Figure 2: Flow-chart of the individual based simulation.



a)

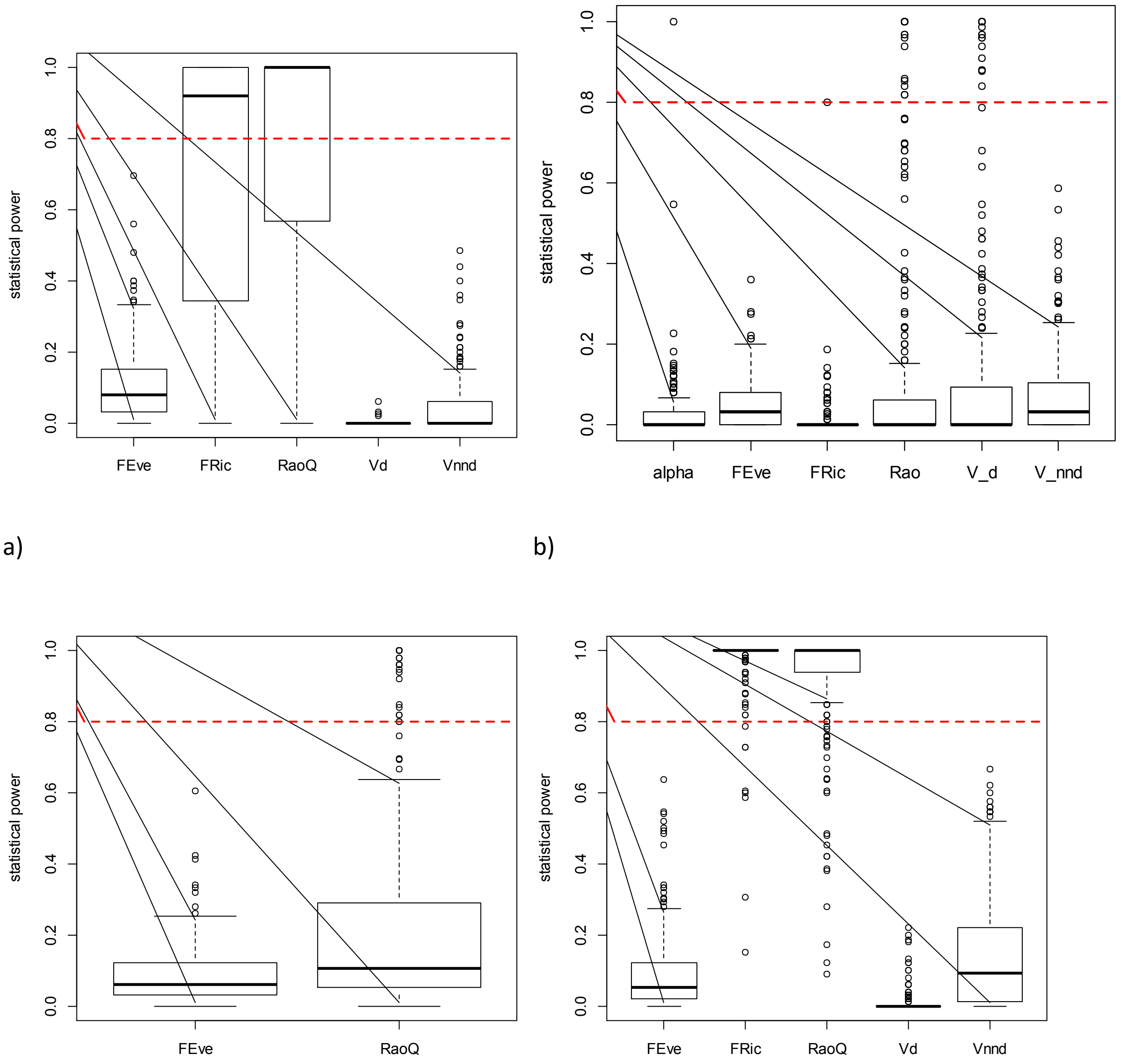


b)



c)

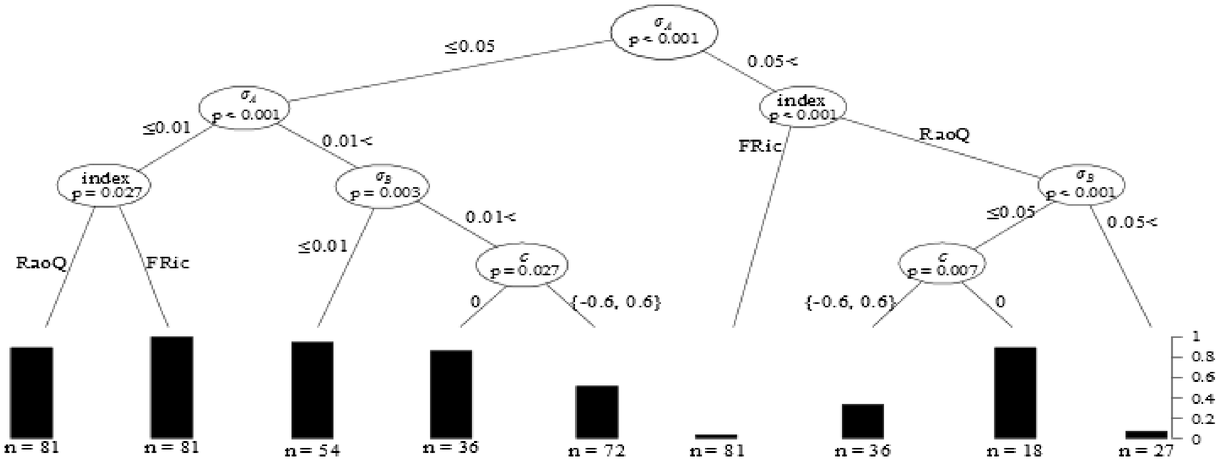
Figure 3: Boxplots of the type I error rates in between-plot (a), within-plot(b) and restricted randomization (c) estimated by using neutral traits . Dashed red line indicates the pre-defined significance level (i.e. 5%).



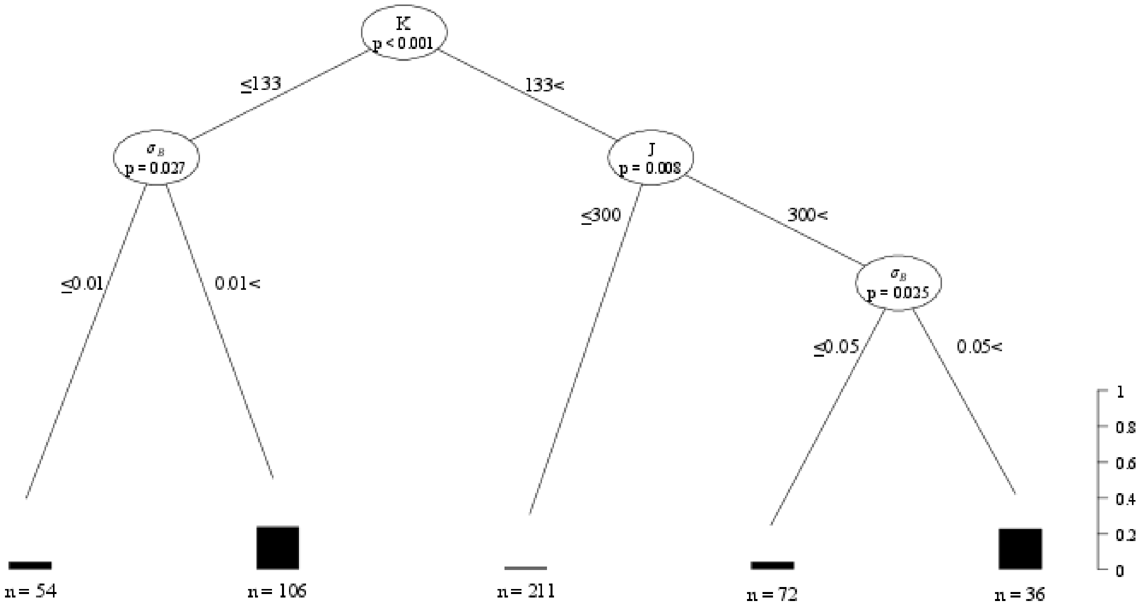
a) b)

c) d)

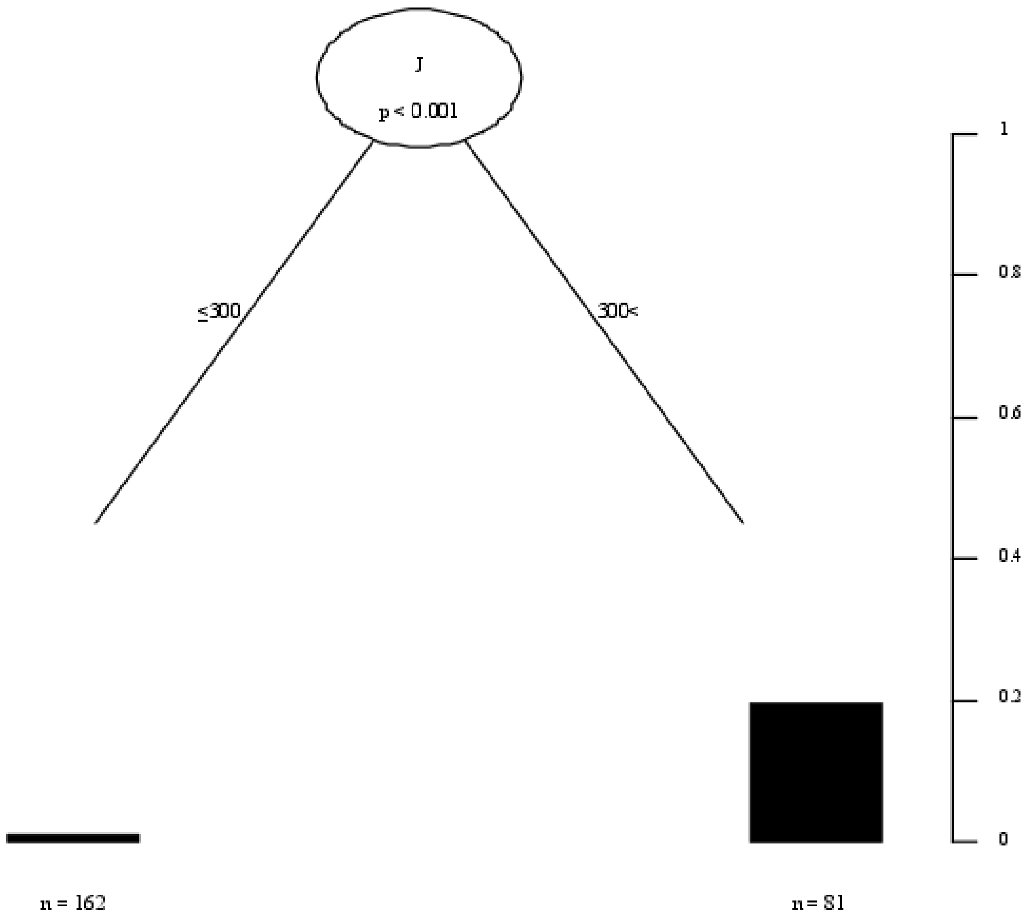
Figure 4: The power of the tests applying different indices. Dashed line indicates a threshold above which the statistical power of the test is generally accepted. (a) detecting trait convergence by using the between-plot randomization strategy, (b) detecting trait divergence by using the between-plot randomization strategy, (c) detecting trait divergence by using the within-plot randomization strategy, (d) detecting trait divergence by using the restricted randomization strategy


σ_A p < 0.001
≤0.05
0.05<
σ_A p < 0.001
index p < 0.001
≤0.01
0.01<
FRic
RaoQ
index p = 0.027
σ_B p = 0.003
σ_B p < 0.001
RaoQ
FRic
≤0.01
0.01<
≤0.05
0.05<
c p = 0.027
c p = 0.007
0
{-0.6, 0.6}
{-0.6, 0.6}
0
n = 81
n = 81
n = 54
n = 36
n = 72
n = 81
n = 36
n = 18
n = 27
1
0.8
0.6
0.4
0.2
0

a)


K p < 0.001
≤133
133<
σ_B p = 0.027
J p = 0.008
≤0.01
0.01<
≤300
300<
σ_B p = 0.035
≤0.05
0.05<
n = 54
n = 106
n = 211
n = 72
n = 36
1
0.8
0.6
0.4
0.2
0

b)



c)

Figure 5: Classification tree models for revealing the effect of parameters and indicator choice on the power of tests. Each internal node describes a statistically significant split (with splitting variable and p value given in the node, and split values given on the branches below), whereas terminal nodes give the number of complying cases (n) and a visual summary of the response variable (rate of cases above the 0.8 threshold). (a) trait convergence tests using between-plot randomization. Only FRic and RaoQ were involved into this analysis. (b) trait divergence tests using between-plot randomization Only RaoQ and Vd were involved into this analysis. (c) trait divergence tests using within-plot randomization Only RaoQ was involved into this analysis

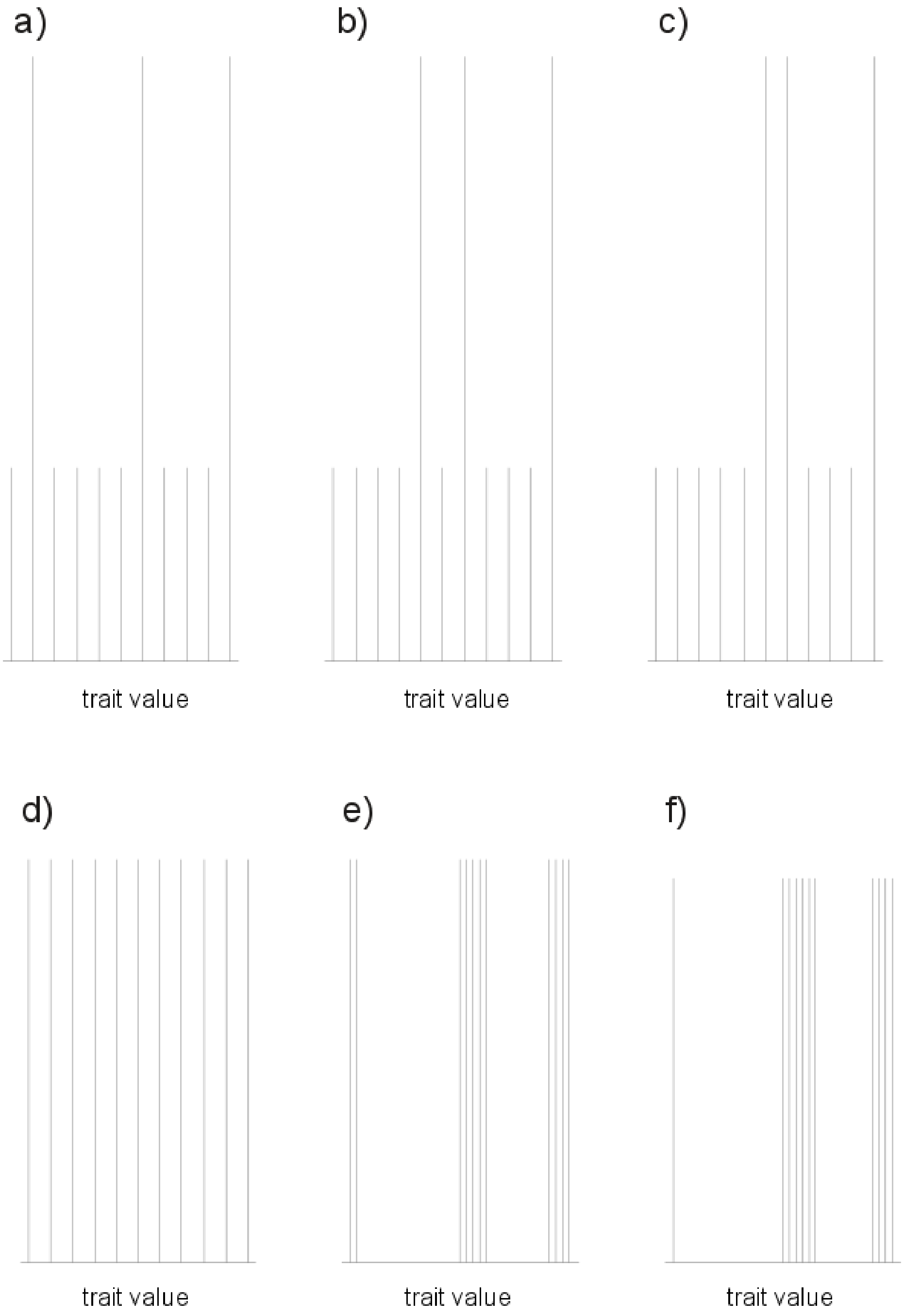


Figure 6: Hypothetical communities for illustrating that both FEve (a-c) and variance of nearest neighbor distance (Vnnd) (d-f) failed to correctly measure the functional evenness because they consider only the (nearest) neighbor species. Each vertical line represents a trait value of the species, while height shows the abundance. FEve is the same for community a) and b), while lower in c). Vnnd is zero in both communities d) and e), but high in community f). These examples illustrate the unfavorable property of these indices that considerably different communities may results in the same value of the index, while in other cases small community changes may cause major changes in the index values. Furthermore, pattern in communities a) and d) may be caused by limiting similarity, while the other patterns contradict this theory, but the values of the indices do not show this difference. Note that communities a-c differ only in position of

abundant and rare species, thus they can be considered as illustrations of a within-plot randomization strategy: if community a) is the field data, the same low values can be easily obtained during the randomization.

Appendix S1: Illustrative examples of the diversity of test statistics used for testing trait divergence/convergence in field case studies

Test statistic	Used for testing		
	trait convergence	trait divergence	both convergence and divergence
FRic	Cornwell *et al.* 2006; Kraft *et al.* 2008; Cornwell & Ackerly 2009; Bernard-Verdier *et al.* 2012; Raevel *et al.* 2012; Yan *et al.* 2012		Schamp *et al.* 2008; Schamp & Aarssen 2009
Vnnd[1]		Cornwell & Ackerly 2009; Schamp & Aarssen 2009; Yan *et al.* 2012; Montaña *et al.* 2013	Schamp *et al.* 2008
RaoQ[2]		Mason & Wilson 2006; Bernard-Verdier *et al.* 2012	Smith *et al.* 1994; Schamp *et al.* 2008; de Bello *et al.* 2009; Thompson *et al.* 2010; Paillex *et al.* 2013
FEve		Raevel *et al.* 2012	

[1.] includes both the variance and the standard deviation of nearest neighbor distances

[2] If only one trait is considered and Euclidean distance is applied, RaoQ is a weighted variance of trait values. Therefore case studies using variance of trait values are listed here.

De Bello, F., Thuiller, W., Lepš, J., Choler, P., Clément, J.-C., Macek, P., Sebastià, 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.

Bernard-Verdier, M., Navas, M.-L., Vellend, M., Violle, C., Fayolle, A. & Garnier, E. (2012). Community assembly along a soil depth gradient: contrasting patterns of plant trait convergence and divergence in a Mediterranean rangeland. *Journal of Ecology*, **100**, 1422–1433.

Cornwell, W.K. & Ackerly, D.D. (2009). Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecological Monographs*, **79**, 109–126.

Cornwell, W.K., Schwilk, D.W. & Ackerly, D.D. (2006). A trait-based test for habitat filtering: convex hull volume. *Ecology*, **87**, 1465–1471.

Kraft, N.J.B., Valencia, R. & Ackerly, D.D. (2008). Functional Traits and Niche-Based Tree Community Assembly in an Amazonian Forest. *Science*, **322**, 580–582.

Laliberté, E. & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, **91**, 299–305.

Mason, N.W.H. & Wilson, J.B. (2006). Mechanisms of species coexistence in a lawn community: mutual corroboration between two independent assembly rules. *Community Ecology*, **7**, 109–116.

Montaña, C.G., Winemiller, K.O. & Sutton, A. (2014). Intercontinental comparison of fish ecomorphology: null model tests of community assembly at the patch scale in rivers. *Ecological Monographs*, **84**, 91–107

Paillex, A., Doledec, S., Castella, E., Merigoux, S. & Aldridge, D.C. (2013). Functional diversity in a large river floodplain: anticipating the response of native and alien macroinvertebrates to the restoration of hydrological connectivity. *Journal of Applied Ecology*, **50**, 97–106.

Raevel, V., Violle, C. & Munoz, F. (2012). Mechanisms of ecological succession: insights from plant functional strategies. *Oikos*, **121**, 1761–1770.

Schamp, B.S. & Aarssen, L.W. (2009). The assembly of forest communities according to maximum species height along resource and disturbance gradients. *Oikos*, **118**, 564–572.

Schamp, B.S., Chau, J. & Aarssen, L.W. (2008). Dispersion of traits related to competitive ability in an old-field plant community. *Journal of Ecology*, **96**, 204–212.

Smith, B., Moore, S.H., Grove, P.B., Harris, N.S., Mann, S. & Wilson, J.B. (1994). Vegetation texture as an approach to community structure - community-level convergence in a New-Zealand temperate rain-forest. *New Zealand Journal of Ecology*, **18**, 41–50.

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

Yan, B., Zhang, J., Liu, Y., Li, Z., Huang, X., Yang, W. & Prinzing, A. (2012). Trait assembly of woody plants in communities across sub-alpine gradients: Identifying the role of limiting similarity. *Journal of Vegetation Science*, **23**, 698–708.

Appendix S2: Detailed description of the simulation model applied

The main steps

Step 1: A regional species pool consisting of S species is created. Each species is characterized by three numeric traits (A, B and C). Trait A is related to habitat matching, trait B regulates resource acquisition, while trait C is neutral. Trait values, which always range between 0 and 1, are sampled from the same prespecified symmetric beta distribution (uniform, arcsine or Wigner semicircle distribution). Correlations between traits A and B are added with the help of a correlation parameter (*c*) in the following way:

$$A = c\,S_B + (1-c)\,S_A, \text{ and } B = c\,S_A + (1-c)\,S_B \quad \text{if } c \geq 0, \text{ and}$$

$$A = |c|\,(1-S_B) + (1-|c|)\,S_A, \text{ and } B = |c|\,(1-S_A) + (1-|c|)\,S_B \quad \text{if } c \leq 0 \quad (1)$$

where S_A and S_B are the random variables from the specified beta distribution, and *A* and *B* are the trait values.

To initialize the environmental gradient, *n* local communities get positioned equidistantly along a pre-specified central part of the range of trait A. As the position of the local communities along the gradient actually determines the optimal value of trait A for each community, the position is characterized by this “optimal trait value” itself, rather than any arbitrarily defined artificial environmental data.

In the last part of the initialization process, the local communities get filled up with species from the species pool until the predefined total number of individuals (J) is reached. J is set to be equal for all of the communities. Propagules are assumed to arrive from the species pool at the same rate for all species, while their survival is considered to depend on the suitability of the local conditions for the species. Accordingly, the initial communities are constructed as a random sample of size J from the species pool with the survival probabilities (s_{jk}) for species *i* in local community *k* calculated in the following way (Gaussian survival functions):

$$s_{ik} = exp\left(-\frac{(A_k - A_i)^2}{2\sigma_A^2}\right) \tag{2}$$

where: A_k is the position of community *k* along the environmental gradient (i.e. the local optimum for trait A), A_i is the value of trait A for species *i*, and σ_A is the parameter determining the tolerance width of the species (considered to be constant for all species). The influence of σ_A on the outcome of a lottery competition for local adaptedness between two species is illustrated in Figure S2.1.

Step 2: In each local community a single individual dies in this step. To keep the simulation as simple as possible, each individual has the same probability for being selected, irrespective of their traits or their “age”.

Step 3: Each individual produces zero or one seed in this step. The probability of reproduction of species *i* in local community *k* depends on the competition for resources:

$$p_{ik} = b \max\left(\frac{K_0 - K_{ik}}{K_0},\ 0\right) \tag{3}$$

where K_{ik} is the competitive pressure for species *i* in community *j*, K_0 is the critical level of competitive pressure above which the probability of reproduction becomes zero, and *b* is the background reproduction probability at no competition (considered to be equal for all species). The competitive pressure is determined as:

$$K_{ik} = \sum_j n_{jk} C_{ij} \tag{4}$$

where C_{ij} is the level of competition between species *i* and *j*, and n_{jk} is the abundance of species *j* in the local community (*k*). The level of competition between two species depends on their similarity in trait B (the resource acquisition trait):

$$C_{ij} = \exp\left(-\frac{\left(B_i - B_j\right)^2}{\sigma_B}\right) \tag{5}$$

where σ_B is a scaling parameter for setting the strength of interspecific competition, set to be the same for each pair of species. The maximum value of C_{ij} is thus 1 (if *i* = *j*, i.e. intraspecific competition), while the minimum value possible over the whole range of B values (i.e. from 0 to 1) is set by σ_B as:

$$C_{min} = \exp\left(-\frac{1}{\sigma_B}\right) \qquad (6)$$

The effect of σ_B on the strength of competition for resources between two species is illustrated in Figure S2.2.

Step 4: To simulate simple metapopulation dynamics, each seed produced can spread to one of the other local communities with probability m. Nevertheless, the target locality is chosen randomly without considering any "spatial position" for the local communities.

Step 5: In this last step, a lottery competition[1] takes place among the seedlings germinating from the seeds produced locally or having arrived from the other localities in the previous step. This is set to be a stochastic process with the probability of winning being proportional to s_{ik} (equation (2)) for the seedlings of species i at location k, thus not necessary the best adapted species will win (Figure S2.1). In each turn only one seedling can survive, all other propagules are lost. This surviving seedling matures by the next iteration, and become a fully functional individual capable of reproduction in full competition with all other individuals of the local community.

Iteration cycles: After having completed step 1 once, steps 2-5 are repeated many times for each local community so that the final composition is independent from the colonization process. Since changes are slower at larger community sizes, the number of cycles was set to 60-150 times (*sim.len*, see Table 1) the number of individuals in a local community (J). The "convergence" of this iterative process is illustrated in Figures S2.3-5 with the help of a randomly selected simulation run.

---

[1] Chesson, P.L. & Warner, R.R. (1981). Environmental Variability Promotes Coexistence in Lottery Competitive Systems. *The American Naturalist*, **117**, 923–943.

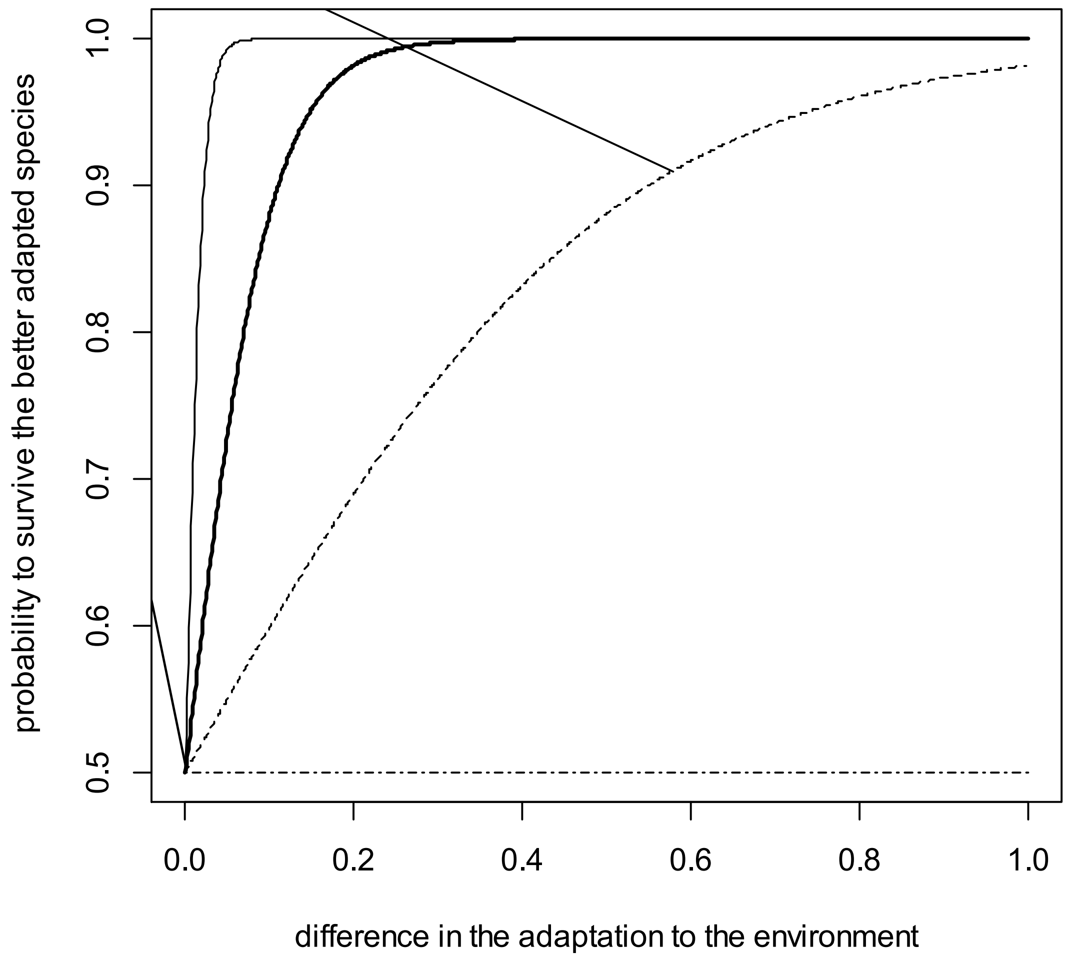


Figure S2.1. The effect of σ_A governing habitat filtering on the outcome of a lottery competition between two species for local adaptedness. The four lines shown correspond to the values tested in this study: $\sigma_A = 0.01$ (thin solid line); $\sigma_A = 0.05$ (thick solid line); $\sigma_A = 0.25$ (dashed line) and $\sigma_A = \infty$ (dashed-dotted line).

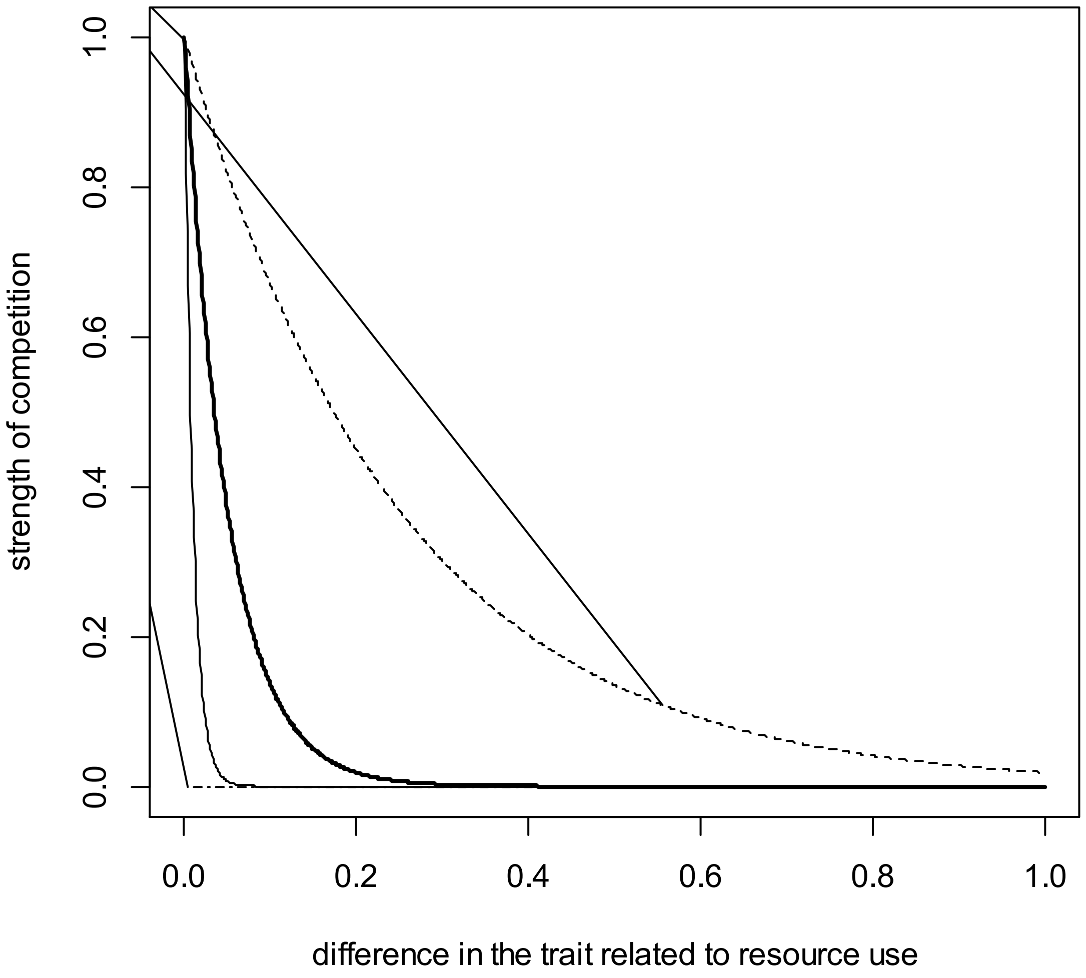


Figure S2.1. The effect of σ_B on the strength of competition for resources between two species in the simulation model. The four lines shown correspond to the values tested in this study: σ_B = 0.01 (thin solid line); σ_B =0.05 (thick solid line); σ_B = 0.25 (dashed line); and σ_B = 0 (dashed-dotted line).

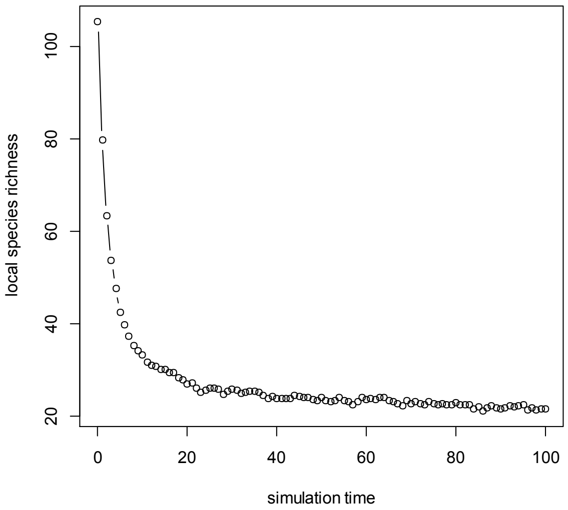


Figure S2.3: Typical change in local species richness (the mean number of species in each local community) during a simulation run

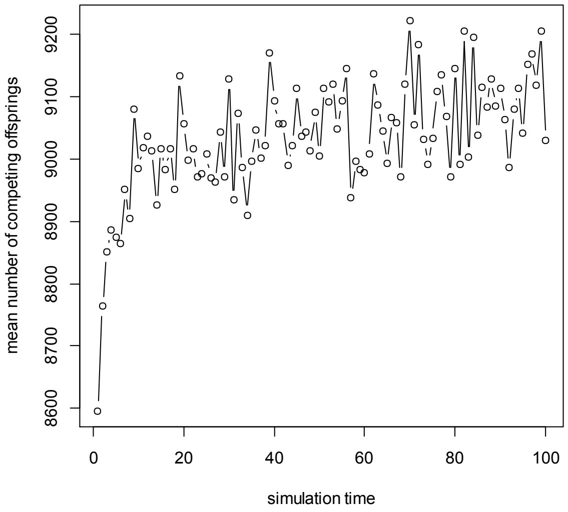
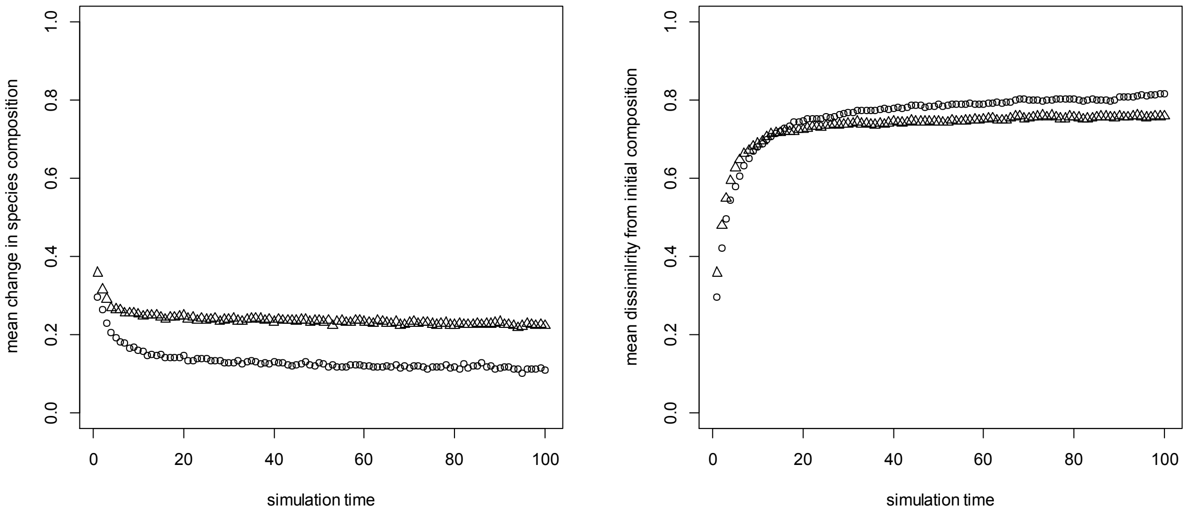


Figure S2.4. Changes in the mean number of competing seedlings during a simulation. Note that only one of them can grow up, all others die.



a) b)

Figure S2.5. Changes in species composition decrease during the simulation. (a) Mean Bray-Curtis dissimilarity between consecutive steps. (b) Mean Bray-Curtis dissimilarity from the starting composition. Circles: simulation with baseline parameter values, triangles: neutral dynamics with baseline parameter values.

Appendix S3: Comparison of trait distribution in the initial regional species pool, in the pool of observed species, and two local communities

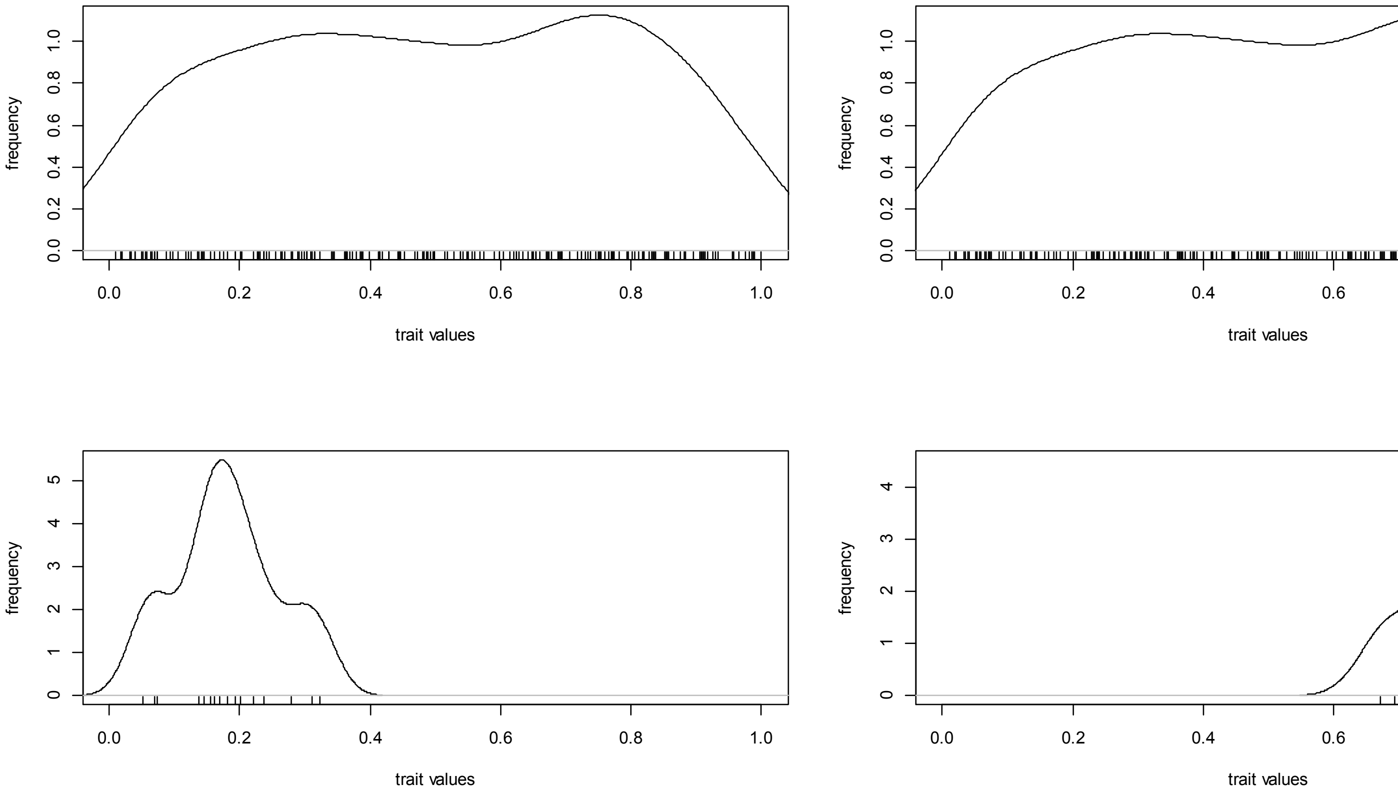


Figure S3.1 Distribution of trait A (related to environmental tolerance) in the initial regional species pool (a), and three set of species pools after a simulation run with baseline values: trait values in the pool of observed species (b), and species occurring in the first (c) and last local community (d) along the environmental gradient. The line is a smoothed frequency curve, while ticks indicate the actual values.

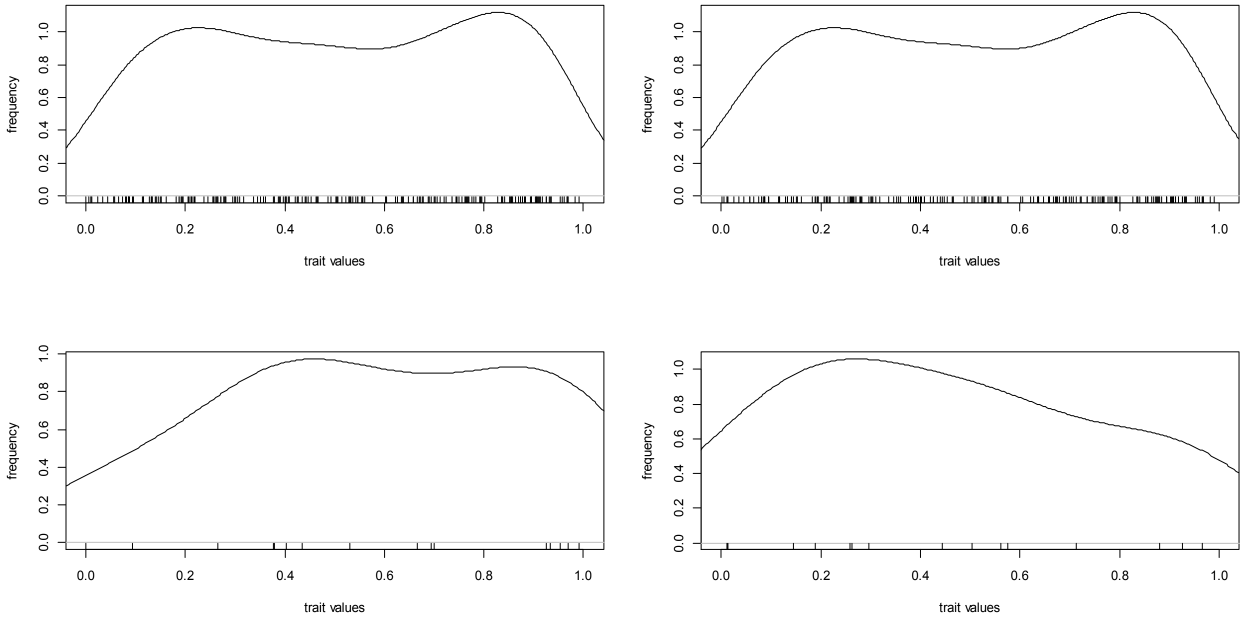
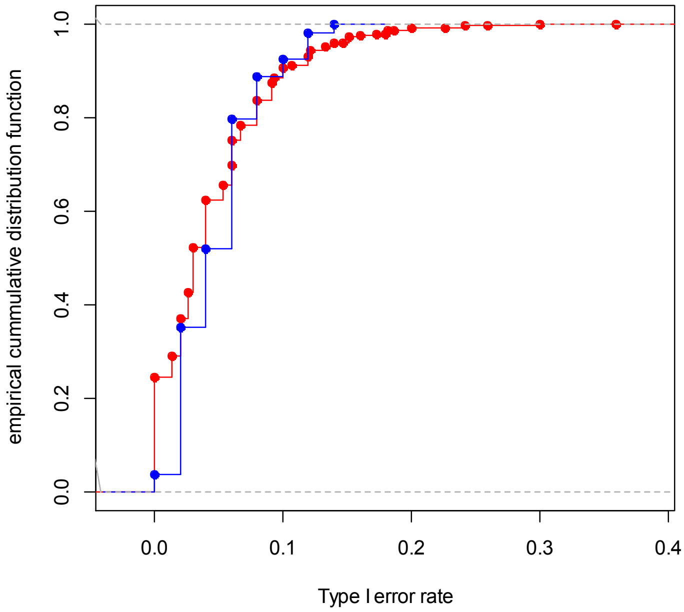


Figure S3.2 Distribution of trait B (related to competition) in the initial regional species pool (a), and three set of species pools after a simulation run with baseline values: trait values in the pool of observed species (b), and species occurring in the first (c) and last local community (d) along the environmental gradient. The line is a smoothed frequency curve, while ticks indicate the actual values.

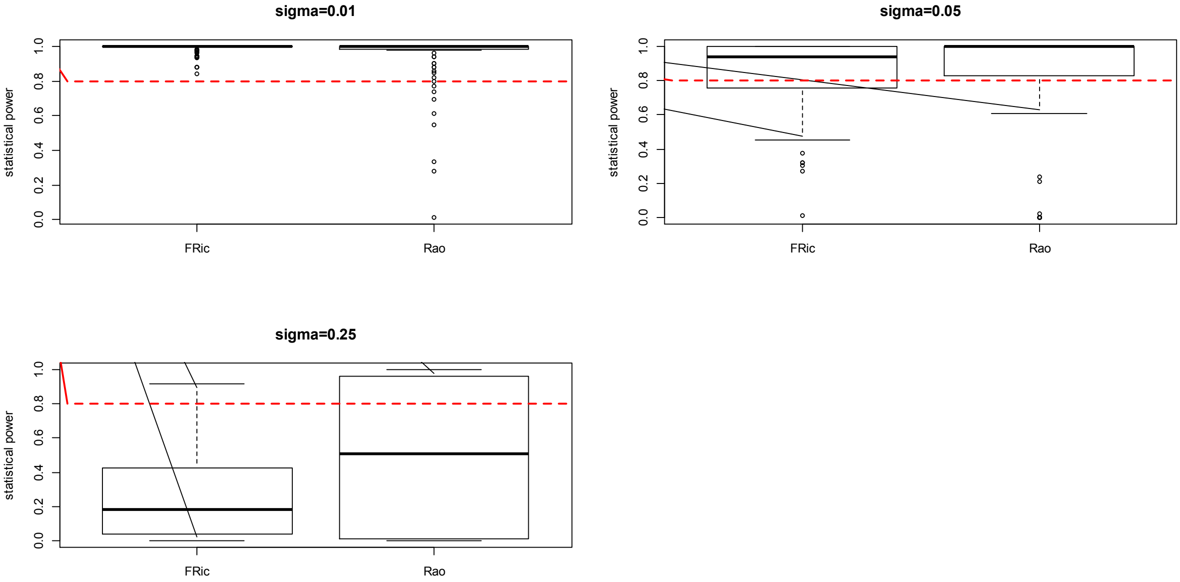
864

865 **Appendix S4:** Additional figures on the power of tests



866

867 Figure S4.1 Empirical cumulative distribution function of type I error rate estimated by switching off
868 both habitat filtering and interspecific competition (blue lines and dots; n=54) and by using neutral
869 traits (red lines and dots; n=1437).



870

871 Figure S4.2 Power of the test for revealing trait convergence when applying between-plot
872 randomization strategy at three different levels of the strength of environmental filtering (lower
873 sigma means stronger filtering, see Figure S2.1).

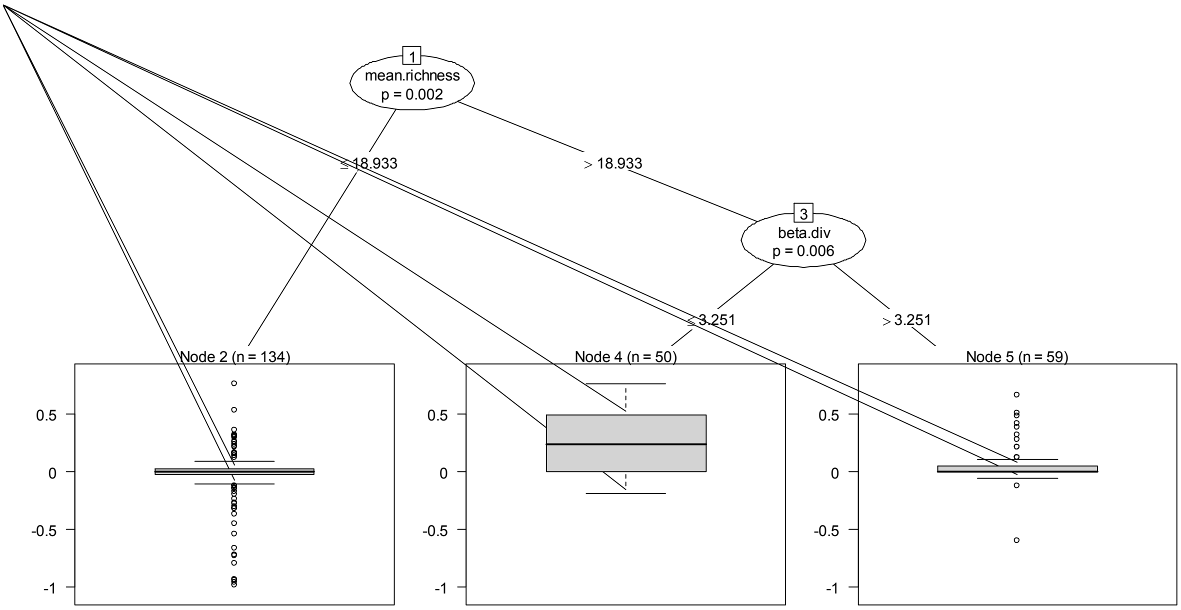


Figure S4.3 Conditional inference tree for differences in power between RaoQ and FRic for detecting habitat filtering applying between-plot randomization. Positive values indicate that RaoQ performs better.

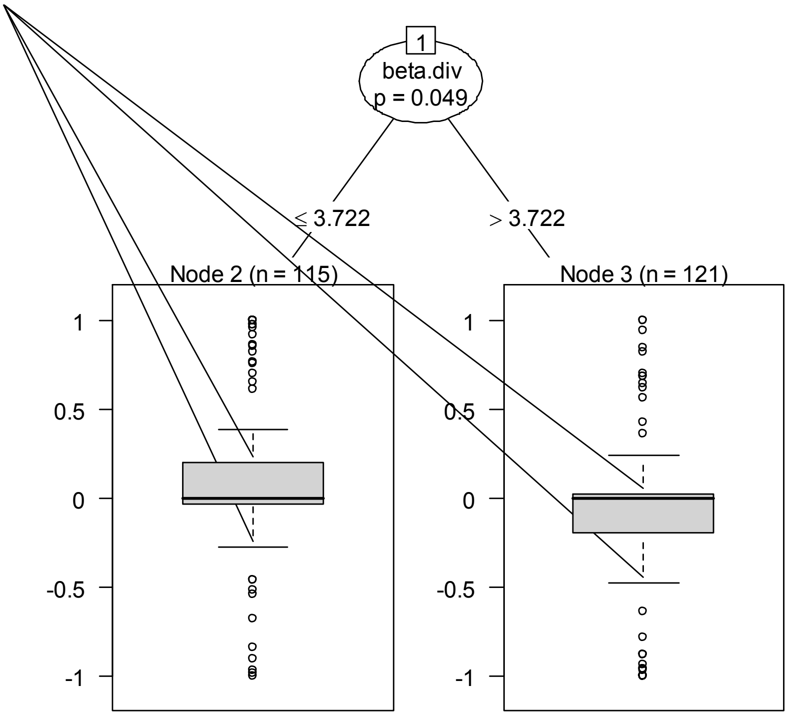
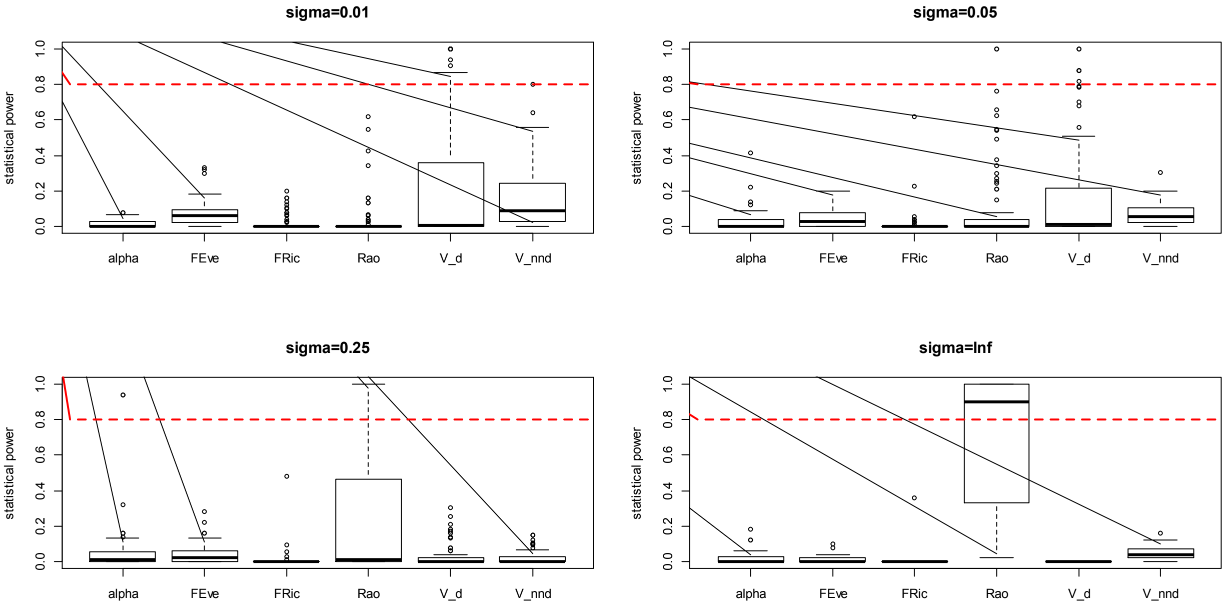
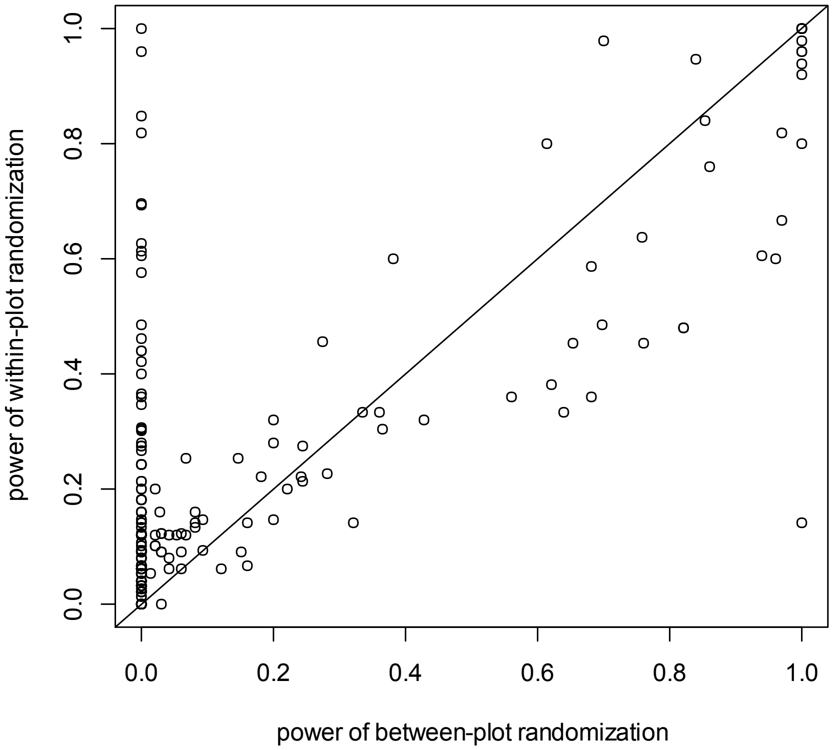


Figure S4.4 Conditional inference tree for differences in power between RaoQ and Vd for detecting trait convergence applying between-plot randomization. Positive values indicate that RaoQ performs better.



883

884 Figure S4.5 Power of trait divergence tests using between-plot randomization at four different levels
885 of the strength of environmental filtering (lower sigma means stronger filtering).



886

887 Figure S4.6 Comparing the performance of tests using between- and within-plot randomization
888 strategies for detecting trait divergence using *RaoQ* as test statistic.

Appendix S5: Dependence of the range of Vd on the range of trait values in one dimension

V_d is the variance of values in the lower (or upper) half-matrix of distance values. Let us denote the vector of these distance values by *d*. The length of this vector, hereafter denoted by *n*, depends on the number of species (*S*):

$$n = \frac{S(S-1)}{2}$$

Using these symbols the variance of distances can be calculated by the following form:

$$V_d = \left(\sum_{i=1}^{n} d_i^2 - \frac{\left(\sum_{i=1}^{n} d\right)^2}{n}\right) \Big/ (n-1)$$

Let us consider first the equidistant placing of species in range R. In this case the distance between neighbouring species is $\Delta = R/(S-1)$. In this case the possible distance values are: Δ, 2Δ, 3Δ, …,$(S-j)\Delta$,…, $(S-1)\Delta$. The number of species pairs with these values are $S-1$, $S-2$, $S-3$,…, $S-j$,…, 1, respectively. Thus,

$$\left(\sum_{i=1}^{n} d\right)^2 = \left(\sum_{j=1}^{S-1} (S-j) j \Delta\right)^2 = \Delta^2 \left(\sum_{j=1}^{S-1} (S-j) j\right)^2 = \frac{R^2}{(S-1)^2} \left(\sum_{j=1}^{S-1} (S-j) j\right)^2$$

$$\sum_{i=1}^{n} d_i^2 = \sum_{j=1}^{S-1} (S-j) j^2 \Delta^2 = \Delta^2 \left(\sum_{j=1}^{S-1} (S-j) j^2\right) = \frac{R^2}{(S-1)^2} \left(\sum_{j=1}^{S-1} (S-j) j^2\right)$$

$$V_d = \frac{R^2}{(S-1)^2 n(n-1)} \left[n \left(\sum_{j=1}^{S-1} (S-j) j^2\right) - \left(\sum_{j=1}^{S-1} (S-j) j\right)^2 \right]$$

In this case, Vd equals the squared *FRic* multiplied by a value that depends only on species richness.

When V_d is maximal, species are grouped into two clusters in the ends of trait gradients. The distances within groups are zero (i.e. species in the group has the same value), while the distance between groups is *R*. Vd is maximal, if the number of between-group distances, *B,* is as close to *n*/2 as possible. The formulas for calculating depend on if *S* is even or odd, but value of *B* that maximizes V_d depends on *S* only. In this case:

$$\left(\sum_{i=1}^{n} d\right)^2 = R^2 B^2$$

$$\sum_{i=1}^{n} d_i^2 = R^2 B$$

$$V_d = R^2 \frac{nB - B^2}{n(n-1)}$$

Again, V_d equals the squared *FRic* multiplied by a value that depends only on species richness.

Between plot randomization does not change the species richness, but it can possibly change the trait values and this way the possible minimum and maximum of V_d.

Appendix S6: Annotated R script for individual based simulation

```
##################################################
#                                                #
#           Community simulation                 #
#                                                #
##################################################
#
# Input parameters:
#
# S = number of species in the regional species pool
# m = probability of colonization from meta-community
# n = number of local communities
# J = number of individuals in a local community
# sigma = tolerance width (equal for all species)
#         has to be positive
#         lower values means more specialist species
#         sigma=Inf means that species are maximally generalist,
#               thus abiotic conditions
#                   do not influence their abundance
#         sigma=0 would mean that species are maximally specialist,
#                 they can occur at only one point of the
#                 environmental gradient(s)
# sigma.b =  width of competition kernel
#            sigma.b=0 means no interspecific competition (no effect
#                       of trait B on competition)
#            sigma.b=Inf leads to equally strong inter- and
#                        intraspecific competition
# If both sigma and sigma.b equal to Inf, species are neutral,
# community composition influenced by random drift only
# b0 = probability of birth without competition
# K = carrying capacity
# xrange =  the range of the environmental gradient, along which the
#           simulated sites lie
# distrib = parameter influencing the shape of distribution. It
#           should be positive!
#             distrib<1 U-shaped distribution
#             distrib=1 uniform  distribution
#             distrib>1 bell-shaped distribution
# correl = correlation between traits, no ceoorelation if correl=0
# (default)
# rand.seed = seed for random number generation
#               the default NULL initialize the random number
#               genarator using current time
# sim.length = length of the simulation
#
#
#Output:
#
#  List of parameters +
#          Y = plot-by-species matrix of abundances
#          trait.env, trait.compet, trait.neutr = three vector of
#          trait values

# S=200; m=0.1; n=50; J=300; sigma=0.05; sigma.b=0.03; b0=1;
```

```
# xrange=.8;distrib=1;
# correl=0;rand.seed=NULL; sim.length=100; K=200
# correl=-.6
# sigma = 0.05; sigma.b = 0.25; S = 200; J = 450; n = 50; distrib =
1; correl = 0; m = 0.1; b0 = 1; sim.length = 100; K = 200; xrange =
0.8; n.random = 1000; sig.level = 0.05
# rand.seed=NULL
traitsimul<-function(S=200, n=50, J=300, sigma=0.05, sigma.b=0.03,
                     m=0.1, b0=1, xrange=.8, distrib=1, correl=0,
                     rand.seed=NULL, sim.length=100, K=200, ...)
{
  # the position of the sites along the environmental gradient:
  #  the xrange long central part of the gradient, sampled at
equidistant points
  x <- seq((1-xrange)/2,1-(1-xrange)/2,xrange/(n-1))
  set.seed(rand.seed)
  cat("Generating species pool... \n")
  trait.a <- rbeta(S, distrib, distrib) # values of trait1 (related
to tolerance)
  tmp <- rbeta(S, distrib, distrib) # values of trait2 (related to
resource use)
  trait.b <- switch(sign(correl)+2,
    abs(correl)*(1-trait.a)+(1-abs(correl))*tmp,  # correl: negative
    tmp,                                          # correl=0 (no
correlation)
    correl*trait.a+(1-correl)*tmp)                # correl:
posistive
  # correl values of +/-.6 result in cor(trait.a,trait.b) of ~.8 for
distrib=1 using this algorithm
  trait.c <- rbeta(S, distrib, distrib) # values of trait3 (neutral)

  dist.b <- as.matrix(dist(trait.b))
  dist.a <- as.matrix(dist(trait.a))
  compet <- matrix(0,S,S)
  if (sigma.b==0) diag(compet) <- 1
  if (sigma.b==Inf) compet <- matrix(1,S,S)
  if ((sigma.b>0) & (sigma.b<Inf)) compet <- exp(-dist.b^2/sigma.b)


  Y<-matrix(NA,n,S) # species abundances
  off.spring<-vector()
  X<-matrix(rep(x,S),ncol=S)  #position along the gradient
  A<-t(matrix(rep(trait.a,n),ncol=n))

  survive <- if (sigma<Inf) pmax(exp(-((X-A)^2)/sigma)-0.01,0) else
matrix(0.99,nrow(X),ncol(X))

  cat("Generating starting community composition...\n")
  for (i in 1:n) Y[i,]<-
table(c(sample(1:S,J,replace=T,prob=survive[i,]),seq(1,S)))-1

  cat("Community assembly...\n")
  pb <- txtProgressBar (min = 0, max = sim.length, char = ".", width
= 45, style = 3)

# epoch=1; j=1
```

```
  for (epoch in 1:sim.length) {
    for (j in 1:J) {
      seed<-matrix(0,nrow=n,ncol=S)
      for (i in 1:n) {
        death<-sample(1:S,1,prob=Y[i,])
        Y[i,death] <- Y[i,death]-1
        NE <- compet %*% Y[i,]
        birth.limit <- b0*(K-NE)/K
        birth.limit[birth.limit<0] <- 0
        occurrence <- (Y[i,]>0)
        seed[i,occurrence] <- rbinom(sum(as.numeric(occurrence)),
           Y[i,occurrence],birth.limit[occurrence])
        }
      off.spring <- matrix(rbinom(n*S,size=seed,prob=(1-
m)),nrow=n,ncol=S)
      seed <- seed-off.spring
      p <- matrix(1/(n-1),nrow=n,ncol=n)
      diag(p) <- 0

      for (i in 1:n)
        for (k in 1:S)
          if (seed[i,k]>0) off.spring[,k] <- off.spring[,k] +
             rmultinom(1, size=seed[i,k], prob=p[,i])

      for (i in 1:n) {
        if (sum(off.spring[i,]*survive[i,])>0) {
          birth <- sample(1:S,1,prob=off.spring[i,]*survive[i,])
         } else {
          birth <- sample(1:S,1,prob=as.numeric(Y[i,]>0))
          }
        Y[i,birth] <- Y[i,birth]+1
        }
      }
    setTxtProgressBar(pb, epoch)
    }
  res <- list(S=S, m=m, n=n, J=J, sigma=sigma, sigma.b=sigma.b,
b0=b0, x=x,
            distrib=distrib, correl=correl, rand.seed=rand.seed,
            sim.length=sim.length,K=K, Y=Y,trait.env=trait.a,
            trait.compet=trait.b,trait.neutr=trait.c)
  close(pb)
  return(res)
  }
```