

**Queries for ecol-93-09-11**

**This manuscript/text has been typeset from the submitted material. Please check this proof carefully to make sure there have been no font conversion errors or inadvertent formatting errors. Allen Press.**

# Functional species pool framework to test for biotic effects on community assembly

FRANCESCO DE BELLO,<sup>1,2,6</sup> JODI N. PRICE,<sup>3</sup> TAMARA MÜNDEMÜLLER,<sup>2</sup> JAAN LIIRA,<sup>3</sup> MARTIN ZOBEL,<sup>3</sup>  
WILFRIED THULLER,<sup>2</sup> PILLE GERHOLD,<sup>3,4</sup> LARS GÖTZENBERGER,<sup>3</sup> SÉBASTIEN LAVERGNE,<sup>2</sup> JAN LEPEŠ,<sup>5</sup>  
KRISTJAN ZOBEL,<sup>3</sup> AND MEELIS PÄRTEL<sup>3</sup>

<sup>1</sup>*Institute of Botany, Czech Academy of Sciences, Dukelská 135 37982 Třeboň, Czech Republic*

<sup>2</sup>*Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Université Joseph Fourier, BP 53 2233 Rue de la Piscine, 38041 Grenoble, Cedex 9, France*

<sup>3</sup>*Institute of Ecology and Earth Sciences, University of Tartu, Lai 40 51005 Tartu, Estonia*

<sup>4</sup>*Department of Experimental Plant Ecology, Institute for Water and Wetland Research, Radboud University Nijmegen, P.O. Box 9010 6500 GL Nijmegen, The Netherlands*

<sup>5</sup>*Department of Botany, Faculty of Science, University of South Bohemia—Institute of Entomology, Biology Centre of Czech Academy of Science, Na Zlaté stoce 1 and Branišovská 31 37005 České Budějovice, Czech Republic*

**Abstract.** Functional trait differences among species are increasingly used to infer the effects of biotic and abiotic processes on species coexistence. Commonly, the trait diversity observed within communities is compared to patterns simulated in randomly generated communities based on sampling within a region. The resulting patterns of trait convergence and divergence are assumed to reveal abiotic and biotic processes respectively. However, biotic processes such as competition can produce both trait divergence and convergence, through either excluding similar species (niche differences, divergence), or through excluding dissimilar species (weaker competitor exclusion, convergence). Hence, separating biotic and abiotic processes that can produce identical patterns of trait diversity, or even patterns that neutralize each other, is not feasible with previous methods. We propose an operational framework in which the functional trait dissimilarity within communities (FDcomm) is compared to the corresponding trait dissimilarity expected from the species pool (i.e., functional species pool diversity, FDpool). FDpool includes the set of potential species for a site delimited by the operating environmental and dispersal limitation filters. By applying these filters, the resulting pattern of trait diversity is consistent with biotic processes, i.e., trait divergence (FDcomm > FDpool) indicates niche differentiation, while trait convergence (FDcomm < FDpool) indicates weaker competitor exclusion. To illustrate this framework, with its potential application and constraints, we analyzed both simulated and field data. The functional species pool framework more consistently detected the simulated trait diversity patterns than previous approaches. In the field, using data from plant communities of typical Northern European habitats in Estonia, we found that both niche-based and weaker competitor exclusion influenced community assembly, depending on the traits and community considered. In both simulated and field data, we demonstrated that only by estimating the species pool of a site is it possible to differentiate the patterns of trait dissimilarity produced by operating biotic processes. The framework, which can be applied with both functional and phylogenetic diversity, enables a reinterpretation of community assembly processes. Solving the challenge of defining an appropriate reference species pool for a site can provide a better understanding of community assembly.

**Key words:** coexistence; competitive exclusion; habitat selection; limiting similarity; niche and neutrality; species pool; trait convergence; trait divergence.

## INTRODUCTION

Understanding how local communities are assembled from the available pool of species is a central goal for both theoretical and applied ecology (MacArthur and Levins 1967, Lortie et al. 2004, Kraft and Ackerly 2010). Classical niche theory predicts there is a limit to how

similar coexisting species can be, often expressed as the principle of limiting similarity (MacArthur and Levins 1967, Adler et al. 2010). According to this principle, coexisting species should be more different than expected at random (divergence). The opposite pattern of increasing similarity (convergence) is usually attributed to the effects of abiotic filters reflecting shared ecological tolerances between coexisting species (Cornwell et al. 2006, Grime 2006). Competition can, however, also increase similarity among coexisting species by excluding species bearing traits associated

Manuscript received 2 August 2011; revised 9 December 2011; accepted 12 March 2012; final version received 11 April 2012. Corresponding Editor: J. Elith.

<sup>6</sup> E-mail: fradebello@ctfc.es

TABLE 1. Glossary of the terminology used.

Term	Definition
Trait divergence and convergence	Patterns observed when coexisting species are respectively more, or less, functionally different than expected by chance (Grime 2006). Divergence indicates niche differentiation. Convergence can indicate either abiotic effects selecting for species with shared ecological tolerances or biotic effects such as competitive exclusion eliminating weakest competitors (Mayfield and Levine 2010), i.e., weaker competitor exclusion.
Niche differentiation	The principle, often expressed as “limiting similarity” (MacArthur and Levins 1967), by which coexistence is maintained by functionally different species exploiting different alpha niches.
Alpha niche	The region of a species’ realized niche corresponding to species diversity at the local scale where interactions among species occur (Silvertown et al. 2006).
Beta niche	The region of a species’ niche that corresponds to the habitat(s) where it is found (Silvertown et al. 2006), determining its belonging to the species pool.
Species pool	The set of species that could potentially co-occur in a site given operating abiotic processes and dispersal filters (Pärtel et al. 2011). It includes those species that can disperse in a site and whose environmental preferences, or beta niche, falls within the range of environmental conditions prevailing in that community.
Dark diversity	The set of species that are currently absent from a site but which belong to its species pool (Pärtel et al. 2011).
Functional diversity	The extent of trait differences between species
Random community	Approach by which the functional diversity observed within communities (FDcomm) is compared to the functional diversity simulated in randomly generated communities based on sampling within a region (FDcommRand).
Functional species pool	The set of trait values possessed by the species comprising a given species pool.
Functional species pool diversity	The extent of trait differences expected between species in the species pool (FDpool). Depending on the mathematical formulation, the FDpool correspond to the expected functional diversity in communities within the species pool (FDcommExp) without null models. Biotic interactions can define the difference between FDcomm and FDpool.

with low competitive ability (Chesson 2000, Grime 2006, Mayfield and Levine 2010). According to this view, biotic interactions may cause species to coexist either because they are functionally dissimilar (divergence, due to niche differentiation), or because they are functionally similar (convergence due to weaker competitor exclusion; Table 1). Little is known about the prevalence of biotic processes causing trait convergence and divergence in community assembly, because of methodological constraints in separating the effect of abiotic and biotic processes that can both produce convergence (Mayfield and Levine 2010, Mason et al. 2011) or patterns that neutralize each other (i.e., divergence counterbalanced by abiotic convergence; Götzenberger et al. 2012).

Functional traits reveal ecological differentiation between species and thus represent one of the most relevant components of biodiversity that can be considered to infer processes governing community assembly (Grime 2006, Swenson and Enquist 2009). An increasing number of studies have used community functional diversity (the extent of trait differences between coexisting species; Table 1) to understand patterns of trait convergence and divergence (Cornwell and Ackerly 2009, Thompson et al. 2010). Commonly, the functional diversity observed within communities (FDcomm) is compared to the functional diversity simulated in randomly assembled communities (FDcommRand). Several methods have been proposed to simulate random communities, but which is the most appropriate is still widely debated (Kraft and Ackerly 2010, Thompson et al. 2010, de Bello 2012). The results of these “random community” approaches give, basi-

cally, three mutually exclusive outputs: (1) trait convergence ( $FDcomm < FDcommRand$ ), (2) trait divergence ( $FDcomm > FDcommRand$ ), or (3) random patterns ( $FDcomm$  not significantly different from  $FDcommRand$ ).

Attention needs to be given to how the patterns are interpreted with random community approaches (de Bello 2012). In random community approaches, the species that could potentially coexist in a given site are often defined as all the species found within a set of samples in a region. As such, species from different habitat types are commonly used to generate the random communities (either randomizing species composition or trait values across samples; Fig. 1). This makes it more likely to detect trait convergence in target communities because species from different habitat types generally have different traits (Grime 2006), hence random communities will tend to have greater functional diversity, i.e.,  $FDcomm < FDcommRand$  (see *Results* for a demonstration). The random community approach could also detect convergence because species with certain similar dispersal syndromes can be overrepresented in the target communities due to geographical dispersal filters. Indeed, most null-model studies using functional diversity have often found evidence of trait convergence rather than trait divergence (Götzenberger et al. 2012). However, without excluding abiotic effects and dispersal limitations, incorrect conclusions about biotic assembly processes can be made (Cornwell and Ackerly 2009).

Removing the effect of abiotic filters on community assembly is essential to distinguish between the opposing biotic filters producing both divergence and convergence

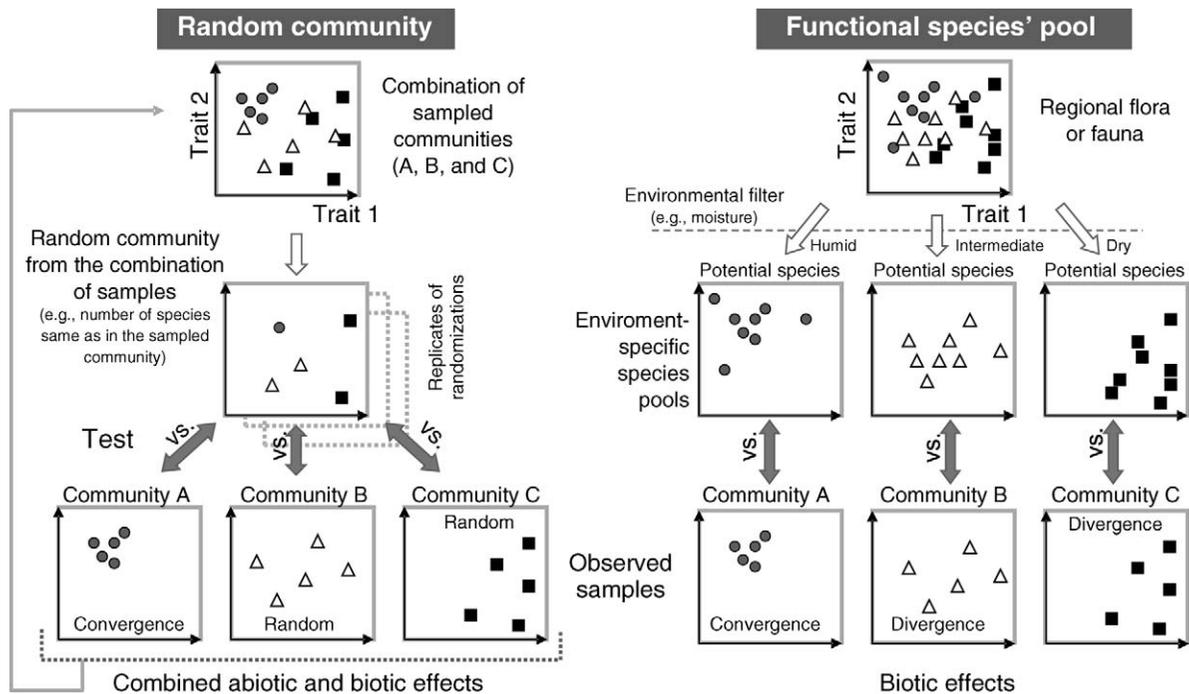


FIG. 1. Conceptual comparison between the random community and the functional species pool approaches. This is a hypothetical example with three sampled communities in different habitats (e.g., humid, dry, and intermediate environmental conditions) and five species per sample (different symbols denote species from different habitats) each with a combination of two trait values. In the random community, the observed samples are combined and randomizations are run to estimate the trait dissimilarity expected by chance from the whole data set. This inevitably produces results that often fail to exclude abiotic effects on community assembly. In the functional species pool, the observed trait dissimilarity in each community is compared, for each community, to the trait dissimilarity expected from the corresponding species pool (Table 1), independently from the composition of other samples.

(Table 1; Mayfield and Levine 2010). In an attempt to remove the effect of abiotic filters, constrained null-models have increasingly been considered where only species from similar habitats, or a given site, are included (Peres-Neto et al. 2001, Kraft and Ackerly 2010, Thompson et al. 2010). This creates, however, another problem: species included in such randomizations are derived only from observed samples (Fig. 1) but biotic filters have already excluded species during community assembly (Tofts and Silvertown 2000, Lortie et al. 2004). The environmentally constrained random community approach, therefore, does not take into account the “absent” part of biodiversity (“dark diversity” sensu Pärtel et al. 2011; Table 1) that theoretically could inhabit a particular site but does not appear in a given sample.

In this paper, we merge modern community assembly theory with the species pool and niche concepts, and reinterpret assembly processes that have been inferred from patterns of trait diversity. We formalize the approach of adding dark diversity effectively into this type of analysis (Tofts and Silvertown 2000) and explore how important it is in the context of assembly patterns detection. We propose a framework, which we call functional species pool that can effectively separate abiotic and biotic processes and helps to distinguish

between opposing biotic effects on community assembly. We define the functional species pool as the set of trait values possessed by the species comprising a given species pool, and the functional species pool diversity (FDpool) as the extent of trait differences expected between species in the species pool (Table 1). With the proposed approach, the species pool is defined independently for each site by including those species that could potentially co-occur given operating abiotic processes, i.e., including the “dark diversity” of a given site. From the regional flora or fauna (i.e., 1 in Fig. 2), the species pool for each target community includes those species whose environmental preferences, or beta niche (Table 1), falls within the range of multiple environmental conditions prevailing in that community (i.e., 2 in Fig. 2). As such all species present in the pool have already “passed” the prevailing environmental and dispersal limitation filters.

It is recognized that only a portion of the species in the species pool will coexist in a site (Pärtel et al. 1996, 2011). We ask, given such a reduction, are species selected from the species pool functionally more or less similar to each other? To answer this question, the functional diversity of each community (FDcomm) is compared to the corresponding functional diversity in the species pool (FDpool; Fig. 2). The difference

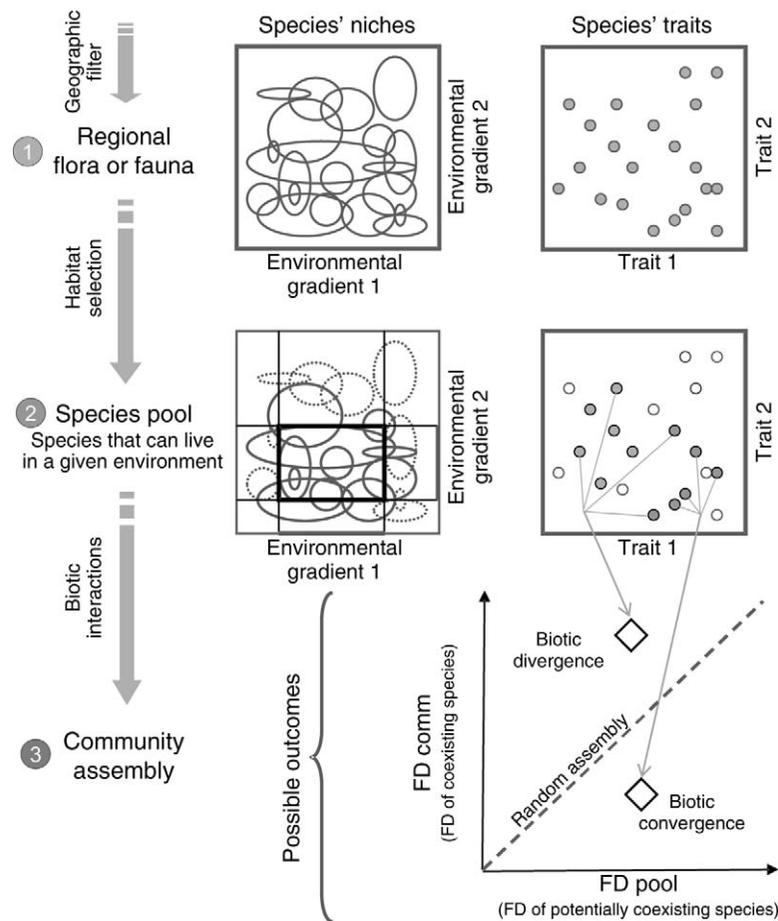


FIG. 2. Given a particular regional flora or fauna, that was formed according to different geographical and historical filters (i.e., 1), habitat selection will filter out species whose environmental preference (or beta niche) falls outside the range of environmental conditions available in a given site, thus creating a convergence in trait values (e.g., compare trait ranges in 1 and 2). This convergence is removed when using the functional species pool approach. The test for assessing the relevance of biotic interactions on community assembly in each single site (i.e., 3) is performed by comparing the functional diversity in the sampled community (FDcomm) with the functional diversity expected within the corresponding functional species pool (FDpool; species that potentially coexist denoted by filled circles). The deviance from a 1:1 relationship (i.e., random assembly) between FDcomm and FDpool corresponds to prevailing biotic assembly processes (FDcomm > FDpool biotic divergence vs. FDcomm < FDpool indicating biotic convergence).

between the FDcomm and FDpool determines if biotic processes produce trait convergence or divergence. The first pattern, namely *biotic convergence* ( $FD_{pool} > FD_{comm}$ ), is an increase in trait similarity as a consequence of certain traits being selected against others, for example through weaker competitor exclusion, or a narrow range of different biotic interactions within and across trophic levels. The second pattern, namely *biotic divergence* ( $FD_{pool} < FD_{comm}$ ), suggests that the most dissimilar traits are selected from the functional species pool, for example through ecological niche differences, facilitation, or a broad range of direct and indirect biotic interactions (sensu Lortie et al. 2004). We illustrate our framework using both simulated and empirical data. We considered mainly functional diversity of species but, as we demonstrate, the approach can be also used with phylogenetic dissimilarity. We discuss

the potential and constraints of this framework in comparison to existing random community approaches.

## METHODS

### Mathematical basis

Different indices of functional diversity (Mouchet et al. 2010, Pavoine and Bonsall 2011) can be used to estimate trait differences between species; the only limitation for our framework is that functional diversity must be independent from the number of species. A suitable index is the mean dissimilarity between species (MPD; Pavoine and Bonsall 2011). This index is mathematically similar to other mainstream indices of functional diversity that can include species relative abundance, such as the widely used Rao's quadratic entropy (Pavoine and Bonsall 2011). The MPD index calculates the mean trait dissimilarity among all possible

pairs of species, which can be weighted (or not) by species relative abundances. Most importantly, with presence/absence data, the MPD index indicates the expectation of trait dissimilarity between two randomly chosen species within a set of species. This way, the FDcomm expected by chance from the species pool (i.e., if species are taken randomly from the species pool; FDcommExp; Table 1) can be estimated by the functional diversity of the species pool (FDpool) with the MPD index (Appendix A).

When using MPD index with presence/absence data, in particular, the FDpool is the reference of the trait dissimilarity expected by chance at all levels of species richness (i.e.,  $FD_{pool} = FD_{commExp}$ ; Appendix A). Hence, the difference between the FDcomm and FDpool expresses an effect size, indicating the strength and direction of biotic processes (i.e., 3 in Fig. 2), which can be compared across communities. We expect a positive relationship between FDcomm and FDpool, and by comparing this observed relationship with a theoretical 1:1 line expected under the assumption of no biotic assembly, we can infer biotic processes across different communities. This is done by testing for differences in slopes and checking if most communities are above or below the expected 1:1 line, and can be done without the use of randomizations. Additionally, randomizations can be used to determine if the deviation of FDcomm vs. FDpool is statistically significant for each particular local community. Randomizations would be particularly important when considering the drivers of uneven abundance of species in a community, or when considering species from the species pool as having an uneven likelihood to occur in a site.

#### Simulations

We used simulated data to compare the functional species pool with random community approaches. Our premise, similarly to Mouchet et al. (2010), was that the approach applied (being the FDpool or the random community) should be able to distinguish communities with known assembly patterns. We simulated communities ( $n = 90$ ) with known trait diversity patterns: trait divergence, trait convergence, or random assembly (scenarios hereafter, with 30 communities/scenario). Communities were assembled out of three species pools to account for the case of species pools from different habitats and including species with different traits (with 10 communities/scenario and species pool). Species pools contained 70 species taken from a total list of 150 species. We selected partially overlapping groups of species for each of three species pools: species pool 1 had species 1–70, species pool 2 had species 40–110, and species pool 3 had species 80–150. Species trait values were spaced at regular increasing intervals from species 1 to 150, with dissimilarity between species 1 and 150 equal to 1. For the random scenario, we randomly selected species within a given species pool. For the convergence scenario, we selected the 30 species with

lowest trait value within a given species pool. For the divergence scenario, we selected species equally spaced in the trait gradient within a pool. In all scenarios, we set community species richness between 10 and 14 species (around 20% of the species pool size; Pärtel et al. 1996). Next we applied the two approaches (random community and functional species pool) in different configurations (Fig. 3a–g). For the random community approach, we used four types of randomizations that are, implicitly or explicitly, commonly applied in the literature: (a) randomizing sample composition across all samples, (b) randomizing across all samples within a species pool, (c) randomizing within a scenario, but across all species pools, and (d) randomizing within a species pool and scenario. In all randomizations, we kept the number of species in a community the same.

For the FDpool approach, we considered three cases: the species pool (e) being correctly estimated (considering all species within a given species pool), (f) being underestimated (considering a lower number of species randomly chosen within a pool), and (g) being overestimated (considering additional species randomly chosen from other pools). For the cases in panels f and g, we considered different levels of misestimation, i.e., 10, 20, 30, 40, and 50 species removed or added to the species pool, respectively, (which corresponds to a misestimation of 14%, 28%, 42%, 57%, and 71% of the species in a species pool). Cases f and g were to reflect a situation in which the environmental requirements of species, which is essential information for the functional species pool framework (Fig. 2), were underestimated or overestimated (see *Empirical data* and *Discussion*).

#### Empirical data

We further tested our framework with empirical data from 27 plant communities characteristic of Northern European habitat types in Estonia. A variety of forest and grassland vegetation types were sampled. The species composition records include presence/absence data for a total of 333 vascular species over the 27 sites (Zobel and Liira 1997). Species composition in each site (only understory species in the forest communities) was recorded in a homogeneous area of ~1 ha. The study region is climatically relatively uniform, with mean temperature in February around  $-6^{\circ}\text{C}$ , mean temperature in July around  $16^{\circ}\text{C}$ , and mean annual precipitation around 600–750 mm (Zobel and Liira 1997). We defined the species pool, for each community independently, by accounting for species habitat preferences from the regional flora of 1073 species and their known distribution in the region. We used species environmental requirements, according to Ellenberg et al. (1991) that place each species on a semi-quantitative scale reflecting habitat preferences along environmental gradients (Appendix C). This database is considered among the most suitable for characteristics of species distributions, i.e., species beta niches (Silvertown et al. 2006). We are aware, however, that the Ellenberg values can partly include

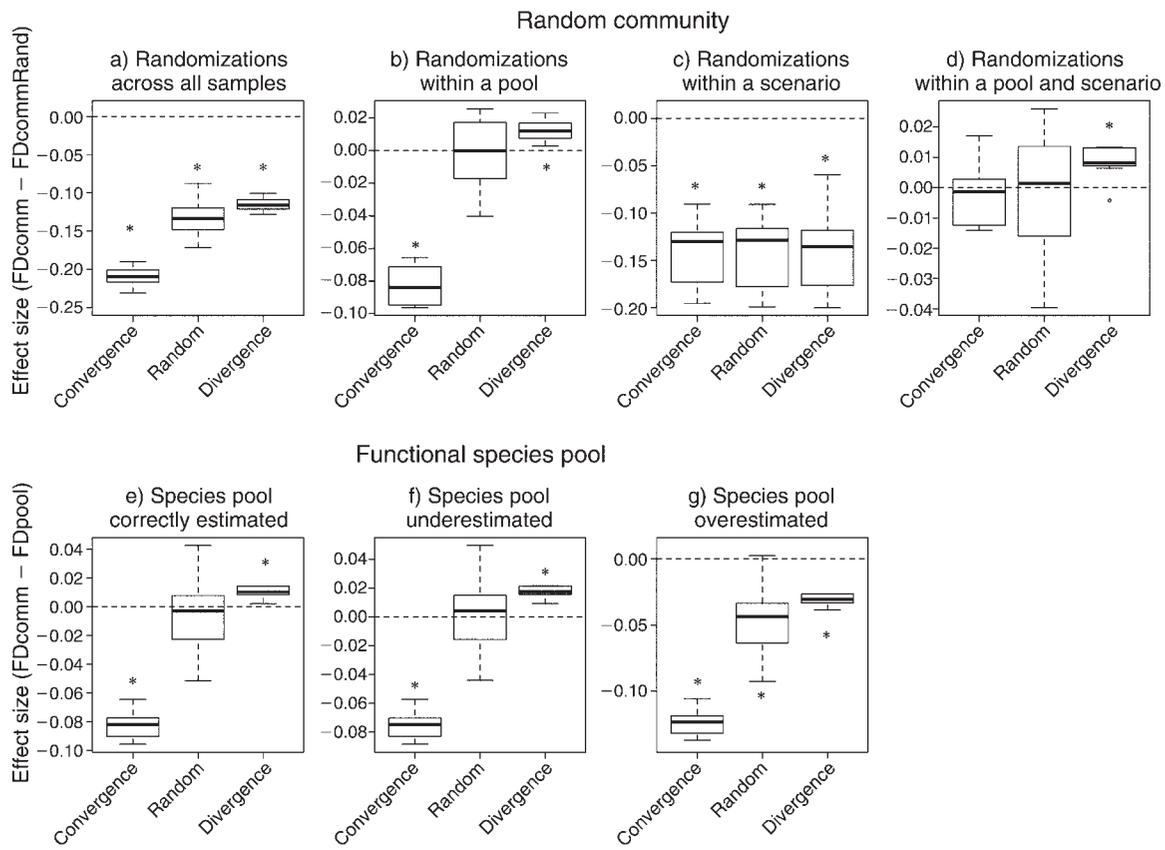


FIG. 3. Results of the simulated communities comparing the random community approach with the functional species pool approach. Simulated communities were assembled within a given species pool and with a known trait dissimilarity scenario (trait convergence, divergence, or random assembly; see text for details). The two approaches were applied with different configurations. For the random community approach, panels refer to (a) randomizations across all samples, (b) randomizations within a species pool, (c) randomizations within a scenario but across species pools, and (d) randomizations within a species pool and scenario. For the functional species pool approach, panels refer to (e) correctly, (f) under-, and (g) overestimated species pools. The effect size is the difference between the observed FD in a community against that expected by chance (which corresponds to  $FD_{commRand}$  for panels a–d and  $FD_{pool}$ , or  $FD_{commExp}$ , for panels e–g; see Table 1). Asterisks indicate cases when the observed FD in samples within a scenario were significantly different from those expected by chance. In box plots, the median for each data set is indicated by the heavy center line, and the first and third quartiles are the lower and upper edges of each box, which is known as the interquartile range (IQR). The extreme values (within 1.5 times the inter-quartile range from the upper or lower quartile) are the ends of the lines extending from the IQR. Points at a greater distance from the median than 1.5 times the IQR are plotted individually. These points represent potential outliers.

effects of dominant biotic interactions as they are largely based on observed species assemblages. As such, we assume that the species pool could be either under- or overestimated (Pulliam 2000 and Lortie et al. 2004), but more likely underestimated (see *Discussion*). We explored the influence of this bias with the simulated data.

For the majority of species in this data set (902 from 1073), plant trait values were freely available from databases (BiolFlor [i.e., Kühn et al. 2004], LEDA [i.e., Kleyer et al. 2008]). We focused on four traits recognized as representing key axes of ecological strategies among vascular plant species: height, specific leaf area, seed mass, and the extent of lateral spread due to different clonal growth organs (Westoby 1998, Kleyer et al. 2008). We computed functional diversity in terms of each trait alone and by combining the traits together (Mouchet et al. 2010). Functional diversity was com-

puted using the `mpd` function in the package *picante* in R (R Development Core Team 2009). The same function was also used to compute phylogenetic diversity in communities ( $PD_{comm}$ ) and in the respective species pool ( $PD_{comm}$ ), as another integrated approach to characterize differences between species in communities and species pools. The phylogeny of the species was based on the BiolFlor database (Kühn et al. 2004; Appendix C).

## RESULTS

### Simulations

When the species pool of a site was correctly estimated, or underestimated, then the functional species pool framework provided the most consistent results in detecting the simulated trait dissimilarity patterns (Fig.

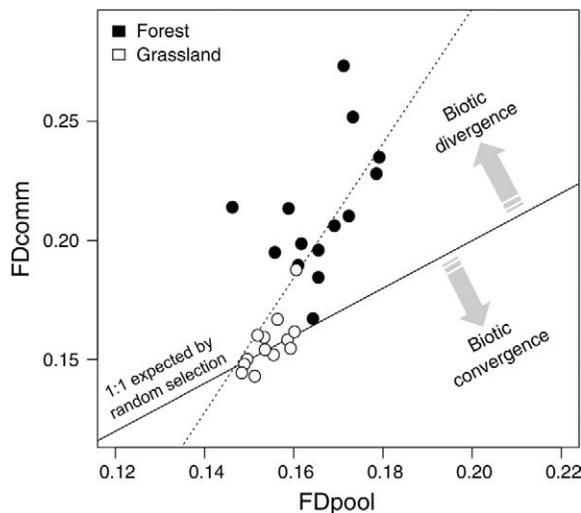


FIG. 4. The observed functional diversity of communities (FDcomm) is compared with the functional diversity of the species pool corresponding to each site (FDpool). The 1:1 relationship (solid line) indicates the diversity expected by a random selection of species from the species pool. Different traits are combined for the analysis (LHS traits, which stands for specific leaf area, height, and seed mass, were combined with plant lateral spread, denoting different clonal modalities). The dotted line denotes the line fitted according to a standardized major axis regression for all samples. The Pearson correlation between FDcomm and FDpool was  $R = 0.73$  for all samples,  $R = 0.69$  for forest communities and  $R = 0.45$  for grassland communities (all correlations  $P < 0.05$ ).

3e). Only one randomization scheme in the random community approach gave as reliable outputs as those obtained when correctly estimating the species pool. This was the case of randomizing samples collected within a given species pool, with samples being assembled under different scenarios (Fig. 3b). In the other three randomization schemes, the random community approach was either not able to distinguish the simulated scenarios or a given scenario was not correctly identified. As predicted, randomizing samples across species pools resulted in trait convergence in all randomization schemes (Fig. 3a, c; but biotic and abiotic effects can not be separated). Importantly, running randomizations across convergent samples within a given species pool, i.e., without considering the dark diversity, indicates random assembly when biotic convergence has already occurred (Fig. 3d).

Increasing the extent of misestimation of species pools produced contrasting patterns. Increasing the level of underestimation of the species pool, which is the most likely situation in practice, made the FDpool approach slightly less reliable (Appendix B). Particularly random communities were increasingly detected as nonrandomly assembled by neglecting a greater number of species in the dark diversity (around 0%, 1%, 3%, 5%, and 11% of cases with missing 10, 20, 30, 40, and 50, respectively; Appendix B). In contrast, in the less likely case of overestimating the species pool of a site, it was sufficient

to include only 10 species from different species pools to find trait convergence in all samples (Fig. 3g leading to a similar case as in Fig. 3a).

#### Empirical data

We detected that FDpool and FDcomm were positively related, rather linearly (both when using all four traits, Fig. 4, and single traits, except for plant height, Fig. 5). However, the relationship between FDpool and FDcomm deviated considerably from the expected 1:1 relationship (Figs. 4 and 5). The two lines showed different slopes according to standardized major axes regressions (test of shift in slope  $P < 0.05$  indicating different slopes; Warton et al. 2006). The greater slope indicates greater trait divergence for higher FDpool values. This pattern suggests that divergence might become prominent when a more functionally diverse species pool allows for the selection of a greater combination of functional dissimilar species. The same patterns, although less marked, were observed with phylogenetic diversity (Appendix C).

We found different results in forest understorey communities compared to grasslands, and different patterns depending on traits used (Figs. 4 and 5). When all traits were considered (Fig. 4), we observed closer-to-random patterns in grassland vegetation (FDcomm similar to FDpool) compared to forest understoreys. Nonetheless, in grasslands convergence was found for plant height and seed weight (significant  $FDcomm < FDpool$ ; Fig. 5), while divergence was reported for lateral spread (significant  $FDcomm > FDpool$ ). Similarly, several grassland sites showed  $PDpool > PDcomm$ , confirming convergent patterns in these habitats (Appendix C). In forest understoreys, a different pattern emerged—the FDpool was often lower than the FDcomm suggesting more prevalence of trait divergence for all four traits and single traits (Figs. 4 and 5).

Finally, we compared the results obtained with the random community approaches with those obtained using our framework (see details in Fig. 6 caption). The random community approach, as expected, detected more evidence for trait convergence, i.e., due to environmental filtering, and less evidence for biotic divergence (Fig. 6). Several sites classified as convergent in the random community approach were classified as randomly assembled using the FDpool. Also, while several cases detecting trait divergence were similarly detected by both randomization types, the number of significant divergent communities was lower with the random community approach.

#### DISCUSSION

##### *Random community or functional species pool?*

Our simulations demonstrate that only by estimating the species pool of a site, with its dark diversity, is it possible to differentiate the patterns of trait diversity produced by operating biotic processes, i.e., biotic

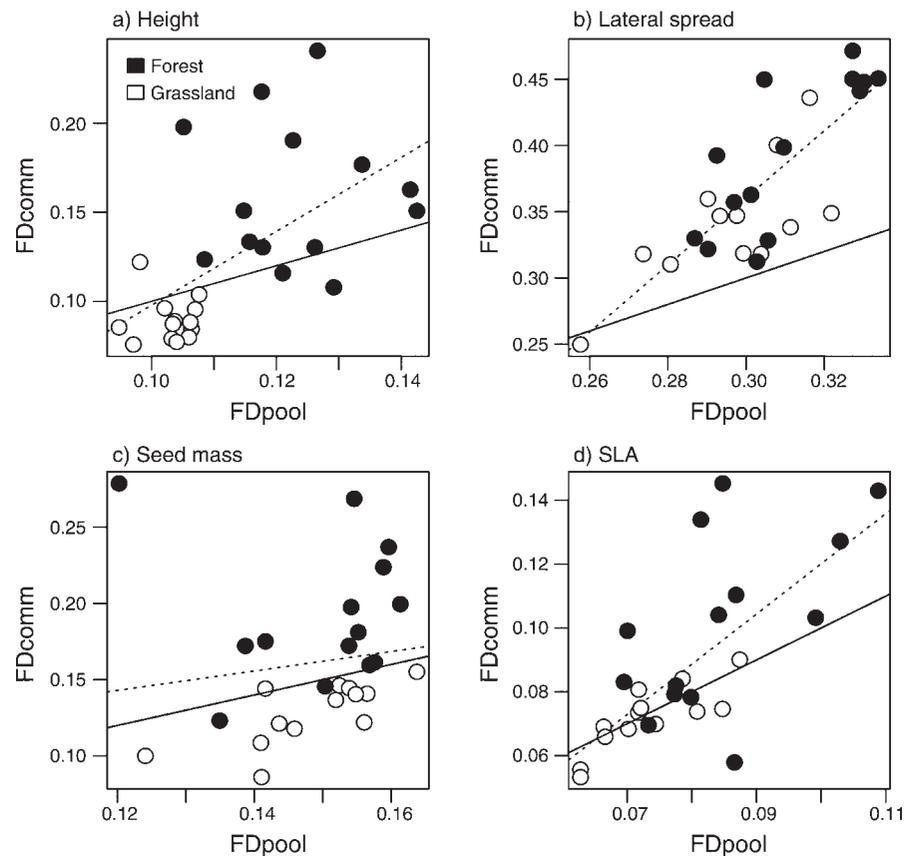


FIG. 5. Relationship between FDpool and FDcomm for the traits considered. Solid circles represent forest and open circles represent grassland. Convergence and divergence (as reported above each panel) were detected using paired tests, indicating cases when the FDcomm was significantly lower or higher than the FDpool (one-tailed  $P$  value  $< 0.05$ ; similar results were obtained with Wilcoxon and  $t$  tests). (a) For height, there was divergence in forest and convergence in grasslands; (b) for lateral spread, there was divergence in both forest and grassland; (c) for seed mass, there was divergence in forest and convergence in grassland; and (d) for specific leaf area (SLA) there was divergence in forest and in all plots together. The paired tests were run on all plots and within forests and grasslands separately. The correlation between FDcomm and FDpool was  $R > 0.65$  in all cases except for height ( $R = 0.58$  for all samples and  $R < 0.2$  within grasslands and forests). See Fig. 4 legend for more details.

convergence vs. divergence. Using environmentally constrained randomizations is not sufficient to detect community assembly patterns, and in particular biotic convergence, unless they take into account the dark diversity of a site (i.e., therefore accounting for the process of species exclusion during assembly). Only one random community approach (Fig. 3b) detected the simulated assembly patterns as consistently as the functional species pool approach. This case, however, would require the same data as the FDpool approach, i.e., it requires the dark diversity of a site to be included in the randomizations (i.e., all species in the species pool are considered in the randomization scheme in Fig. 3b).

Notably this specific case (Fig. 3b) corresponds to a frequently used randomization approach, where constraints in the randomization are imposed by limiting the range of trait values of the species that could coexist together, i.e., based on the trait range observed within a habitat (Cornwell and Ackerly 2009). Our results indicate that this approach is indeed useful, especially

if the entire range of trait values in the species pool has been considered. Using only the observed trait range in a given sample (i.e., when the whole species pool is not considered) to build randomizations can produce false random assembly patterns when biotic convergence has occurred. This case is shown in the convergent scenario in Fig. 3d, which simulates the situation where (1) only samples from homogeneous environmental conditions are considered and (2) competitive species with similar traits have excluded less competitive (and functionally different) species during community assembly. In this case, the observed trait range in each sample is lower than the trait range in the species pool. The random community approach here is only randomizing functionally similar species across samples, while neglecting the functionally different species that were excluded during assembly and this produces false random patterns. As such randomizations within the range of observed trait values in a community could overlook biotic convergence patterns when the observed trait

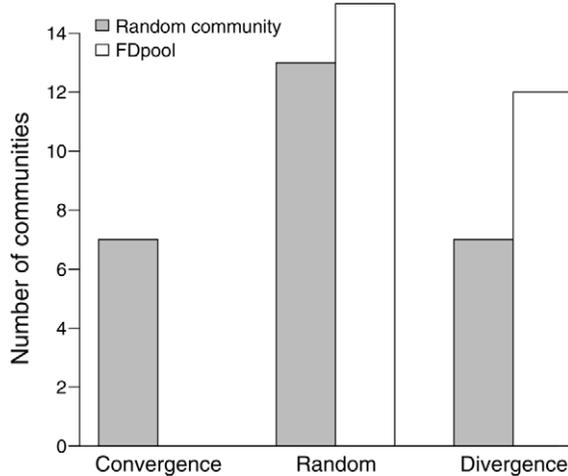


FIG. 6. Results obtained with the random community were compared with those produced with the functional species pool, using the data of the illustrative example presented in the main text (i.e., with all four functional traits). Bars indicate the number of communities for which significant trait convergence, trait divergence or random patterns were detected. The statistical significance of assembly patterns was considered as the number of times that the observed functional diversity in a given community was lower, or higher, than the expected functional diversity (with 999 randomizations for each approach). With the random community approach, species were shuffled across sampled communities with classic randomization procedures. The results presented refer to randomizations obtained with the quasi-swap method of Miklos and Podani (2004). Using other randomization approaches (Götzenberger et al. 2012) gave similar results. The number of samples showing convergence, random and divergence patterns was different between the two approaches (Fisher exact test,  $P = 0.011$ ).

range in the samples is lower than that in the species pool.

Overall, the FDpool approach more consistently detected both biotic convergence and divergence and separated these from abiotic effects. This was especially so when the species pool was not overestimated by including species from different habitats. Overestimating species pools is, however, a less likely scenario (Dupré 2000, Ewald 2002; see next section). Notably, if the random community approach detects trait divergence it is most likely valid, because the probability of detecting divergence is smaller than with the FDpool approach. The field data confirmed these patterns. When the overriding effects of habitat filtering are not removed then trait divergence might be masked. Comparing the results of the FDpool with those obtained with the random community approach can further assess to what extent the species pool has been correctly estimated. We have shown that trait divergence detected by the random community approach should be robust (Fig. 3), and then the trait divergence obtained with the two methods (Fig. 6) actually validates indirectly the estimation of the species pool used for the empirical test.

Both simulated and field data demonstrated that the FDpool approach is a useful tool to detect biotic convergence through weaker competitor exclusion, and divergence through niche differences. These patterns, however, are not mutually exclusive; in some communities competition can produce divergence of some traits and convergence in others (Grime 2006). For example, both biotic convergence and divergence can occur in grasslands, but they might operate differently for different suites of vegetative and regenerative traits (Grime 2006, Swenson and Enquist 2009). This pattern could remain undetected when tested on multiple traits (compare Figs. 4 and 5). Our approach may be more effective when focusing on one trait at a time and with carefully chosen trait combinations. Trait databases are becoming increasingly available for different biomes and in the future it will be possible to use site-specific measures in some places, including intraspecific trait variability (Mason et al. 2011). Finally, we suggest that phylogenetic diversity could be also applied within the framework. However, as pointed out by Swenson and Enquist (2009), while phylogenetic diversity may be a good general proxy for ecological similarity between species, it also has a reduced capacity to depict the functional mechanisms behind species coexistence, especially when coexisting species simultaneously converge and diverge in different traits, as we show here.

#### *Challenges and potential of the framework*

The results of this study indicate that estimating the dark diversity of a site is essential for effectively separating abiotic and biotic processes and distinguishing opposing biotic effects in community assembly. An important step in applying the method is a good estimation of the species pool; species pools can be determined in different ways (Dupré 2000, Pärtel et al. 2011), but importantly for our approach defining an environmentally and geographically delimited species pool is essential to remove abiotic and dispersal effects (see Tofts and Silvertown 2000) and avoid overestimation of the species pool. We recommend excluding species that could potentially disperse to the focal community, but would never establish a viable population if no other (plant) species were present (Weiher and Keddy 1995). One way to do so is through extensive sampling of one habitat type within a region, but this requires a lot of time when many communities are considered, as it is difficult to find all potential species (Sadlo et al. 2007). Alternatively, species pools can be predicted by species co-occurrence matrices (as a likelihood of species to inhabit a community; Ewald 2002), spatial modeling of species occurrence (Mokany and Paine 2011), or with monographs where species habitat preferences are described for regional floras or faunas (Sadlo et al. 2007). Moreover, recent advances in ecoinformatics have increased the possibilities of determining environmentally and geographically delimited species pools from large data-bases. For instance,

ecological habitat characteristics and species occurrence data are increasingly available for many locations within several open repositories (e.g., NCEAS and USDA NRCS as well as a European-wide database DAISIE depicting alien species distribution). The level of precision on habitat type distribution is also increasing at both the country level and at larger scales (e.g., CORINE habitat types created for Europe; *available online*).<sup>7</sup>

We mentioned (*Methods*) that the range of environmental conditions that could be potentially occupied by a species could be both underestimated (Pulliam 2000) or overestimated (Lortie et al. 2004), but most likely the estimations of the species pool will produce an underestimation of the theoretical species pool because of the difficulty in estimating the dark diversity of a site (Dupré 2000). However, we demonstrated that potential underestimations of the species pool will not dramatically alter the results obtained with our approach, and that the results obtained are generally more reliable than those obtained with the random community approach. In fact, in our simulations, only by missing more than 50% of the species in the species pool (which seems a rather extreme case in practice), the results became slightly less consistent and resulted in nonnegligible mistakes (i.e., in around 11% of randomly assembled communities could be detected as nonrandom with a 70% underestimation; Appendix B).

#### CONCLUSIONS

The fundamental mechanisms driving the coexistence of species remains a central question in ecology (Swenson and Enquist 2009, Mayfield and Levine 2010). Despite conceptual and empirical attempts to reconcile opposing views, it is still unclear which biotic processes have a primary impact on different community types (Adler et al. 2010, Mayfield and Levine 2010, Mason et al. 2011). Here, we show that several existing methods do not, both conceptually and mathematically, separate the opposing effects of biotic processes on community assembly and could lead to misinterpretations in the patterns of trait diversity. By considering the set of potentially coexisting species, the functional species pool framework can detect both the effect of niche differentiation (leading to biotic trait divergence) and weaker competitor exclusion (leading to biotic trait convergence). We demonstrate that the framework can be applied to assess the importance of biotic processes influencing species coexistence, across ecosystems and different functional traits and presents opportunities to examine community assembly in other ecological systems. We also demonstrate that underestimating the species pool of a site, which is likely due to the difficulty in estimating the entire dark diversity of a site, will still provide reliable results with our framework. As with

other existing approaches, the main challenge with the FDpool approach is in how the processes are inferred from the observed patterns (Mouchet et al. 2010). Importantly, we demonstrate that (1) the functional species pool approach is particularly reliable in detecting the expected patterns of both trait convergence and divergence caused by biotic assembly processes and, (2) without considering the dark diversity site, it is not possible to accurately differentiate the patterns of trait dissimilarity produced by operating biotic processes. The application of the framework will help understanding the opposing effects of competition in plant communities, and contribute to a greater understanding of the mechanisms underlying local species coexistence.

#### ACKNOWLEDGMENTS

This research was supported by the EU funded EcoChange project (FP6 European Integrated project 2007–2011, contract No 066866 GOCE), the ANR DIVERSITALP project (ANR 2008–2011, contract No. ANR 07 BDIV 014), the CNRS APIC RT PICs 4876 action, the GACR 206/09/1471, the European Regional Development Fund (Center of Excellence FIBIR), the European Social Fund (Mobilitas grant MJD47), the Estonian Science Foundation (8323, 8613), the European Union 7th framework project SCALES (FP7-226852) and by the University of Tartu (SF0180098s08 and SF0180012s0). We thank several anonymous reviewers for their help with the manuscript.

#### LITERATURE CITED

- Adler, P. B., S. P. Ellner, and J. M. Levine. 2010. Coexistence of perennial plants: an embarrassment of niches. *Ecology Letters* 13:1019–1029.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* 31:343–366.
- Cornwell, W. K., and D. D. Ackerly. 2009. Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecological Monographs* 79:109–126.
- Cornwell, W. K., D. W. Schwilk, and D. D. Ackerly. 2006. A trait-based test for habitat filtering: convex hull volume. *Ecology* 87:1465–1471.
- de Bello, F. 2012. The quest for trait divergence and convergence: are null-models the magic wand? *Global Ecology and Biogeography* 21:312–317.
- Dupré, C. 2000. How to determine a regional species pool: a study in two Swedish regions. *Oikos* 89:128–136.
- Ellenberg, H., H. E. Weber, R. Düll, V. Wirth, W. Werner, and D. Paulißen. 1991. Zeigerwerte von Pflanzen in Mitteleuropa. *Scripta Geobotanica* 18:1–248.
- Ewald, J. 2002. A probabilistic approach to estimating species pools from large compositional matrices. *Journal of Vegetation Science* 13:191–198.
- Götzenberger, L., et al. 2012. Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biological Reviews* 87:111–127.
- Grime, J. P. 2006. Trait convergence and trait divergence in herbaceous plant communities: mechanisms and consequences. *Journal of Vegetation Science* 17:255–260.
- Kleyer, M., et al. 2008. The LEDA Traitbase: a database of life-history traits of the Northwest European flora. *Journal of Ecology* 96:1266–1274.
- Kraft, N. J. B., and D. D. Ackerly. 2010. Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecological Monographs* 80:401–422.

<sup>7</sup> <http://www.eea.europa.eu/publications/COR0-landcover>

- Kühn, I., W. Durka, and S. Klotz. 2004. BiolFlor—a new plant-trait database as a tool for plant invasion ecology. *Diversity and Distributions* 10:363–365.
- Lortie, C. J., R. W. Brooker, P. Choler, Z. Kikvidze, R. Michalet, F. I. Pugnaire, and R. M. Callaway. 2004. Rethinking plant community theory. *Oikos* 107:433–438.
- MacArthur, R., and R. Levins. 1967. Limiting similarity, convergence, and divergence of coexisting species. *American Naturalist* 101:377–385.
- Mason, N. W. H., F. de Bello, J. Doležal, and J. Lepš. 2011. Niche overlap reveals the effect of competition, disturbance and constraining assembly processes in experimental grassland communities. *Journal of Ecology* 99:788–796.
- Mayfield, M. M., and J. M. Levine. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters* 13:1085–1093.
- Miklos, I., and J. Podani. 2004. Randomization of presence-absence matrices: comments and new algorithms. *Ecology* 85:86–92.
- Mokany, K., and D. R. Paine. 2011. Dark diversity: adding the grey. *Trends in Ecology and Evolution* 26:264–265.
- Mouchet, M. A., S. Villegger, N. W. H. Mason, and D. Moullot. 2010. Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. *Functional Ecology* 24:867–876.
- Pärtel, M., R. Szava-Kovats, and M. Zobel. 2011. Dark diversity: shedding light on absent species. *Trends in Ecology and Evolution* 26:124–128.
- Pärtel, M., M. Zobel, K. Zobel, and E. van der Maarel. 1996. The species pool and its relation to species richness: evidence from Estonian plant communities. *Oikos* 75:111–117.
- Pavoine, S., and M. B. Bonsall. 2011. Measuring biodiversity to explain community assembly: a unified approach. *Biological Reviews* 4:792–912.
- Peres-Neto, P. R., J. D. Olden, and D. A. Jackson. 2001. Environmentally constrained null-models: site suitability as occupancy criterion. *Oikos* 93:110–120.
- Pulliam, H. R. 2000. On the relationship between niche and distribution. *Ecology Letters* 3:349–361.
- R Development Core Team. 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>
- Sadlo, J., M. Chytrý, and P. Pyšek. 2007. Regional species pools of vascular plants in habitats of the Czech Republic. *Preslia* 79:303–321.
- Silvertown, J., M. Dodd, D. Gowing, C. Lawson, and K. McConway. 2006. Phylogeny and the hierarchical organization of plant diversity. *Ecology* 87:S39–S49.
- Swenson, N. G., and B. J. Enquist. 2009. Opposing assembly mechanisms in a Neotropical dry forest: implications for phylogenetic and functional community ecology. *Ecology* 90:2161–2170.
- Thompson, K., O. L. Petchey, A. P. Askew, N. P. Dunnett, A. P. Beckerman, and A. J. Willis. 2010. Little evidence for limiting similarity in a long-term study of a roadside plant community. *Journal of Ecology* 98:480–487.
- Tofts, R., and J. Silvertown. 2000. A phylogenetic approach to community assembly from a local species pool. *Proceedings of the Royal Society B* 267:363–369.
- Warton, D. I., I. J. Wright, D. S. Falster, and M. Westoby. 2006. Bivariate line-fitting methods for allometry. *Biological Reviews* 81:259–291.
- Weither, E., and P. A. Keddy. 1995. The assembly of experimental wetland plant communities. *Oikos* 73:323–335.
- Westoby, M. 1998. A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil* 199:213–227.
- Zobel, K., and J. Liira. 1997. A scale-independent approach to the richness vs biomass relationship in ground-layer plant communities. *Oikos* 80:325–332.

## SUPPLEMENTAL MATERIAL

**Appendix A**

Validation of the mathematical framework proposed (*Ecological Archives* XXXXX).

**Appendix B**

Results for different levels of species pool misestimation (*Ecological Archives* XXXXX).

**Appendix C**

Empirical test: species pool estimation and results for phylogenetic diversity (*Ecological Archives* XXXXX).

**Supplement 1**

R code for the simulations in Fig. 3 and Appendix B (*Ecological Archives* XXXXX).

**Supplement 2**

R code for tests in Appendix A (*Ecological Archives* XXXXX).