

A signal of competitive dominance in mid-latitude herbaceous plant communities

José A. Capitán^{1,2}, Sara Cuenda³, Alejandro Ordóñez⁴, and David Alonso²

¹Department of Applied Mathematics. Technical University of Madrid. Av. Juan de Herrera, 6. 28040 Madrid, Spain.

²Theoretical and Computational Ecology Group. Center for Advanced Studies (CEAB-CSIC). C. Accés Cala St. Francesc 14. 17300 Blanes, Catalonia, Spain.

³Universidad Autónoma de Madrid. Facultad de Ciencias Económicas y Empresariales. Depto. Análisis Económico: Economía Cuantitativa. C. Francisco Tomás y Valiente 5, 28049 Madrid, Spain

⁴Department of Bioscience. Aarhus University, Aarhus. Ny Munkegade 114, DK-8000, Aarhus C, Denmark.

January 27, 2020

Corresponding author	José A. Capitán / David Alonso
Address	Department of Applied Mathematics Av. Juan de Herrera, 6 28040 Madrid, Spain
Phone / Fax	+34-913 36 7594 / +34-913 36 7806
E-mail José A. Capitán	ja.capitan@upm.es
E-mail Sara Cuenda	sara.cuenda@uam.es
E-mail Alejandro Ordóñez	alejandro.ordonez@bios.au.dk
E-mail David Alonso	dalonso@ceab.csic.es
Type of Article	Letter
Running title	Significant dominance in plant communities
Words Abstract	150
Words Main Text	≈3100
Figures / Tables	6 / 0
References	67
Text Boxes	0

Statement of authorship: JAC and DA designed the research; JAC and SC performed the statistical analysis; JAC, SC, AO and DA analyzed data and results; AO prepared data files and contributed materials; JAC and DA wrote the paper.

Data accessibility statement: The LEDA Traitbase is an open internet data base. The data base of Atlas Florae Europaeae that supports the findings of this study is available from the Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo (<https://www.luomus.fi/en/publishing-atlas-florae-europaeae>). Should the manuscript be accepted, the specific data files that were assembled for this article from this data base will be archived in an appropriate public repository (Dryad, Figshare or Hal).

Code availability: Computer code to analyze data and support the claims reported in the manuscript will be made available upon request with no restrictions.

Abstract

Understanding the main determinants of species coexistence across space and time is a central question in ecology. However, ecologists still know little of the scales and conditions at which biotic interactions matter and how these interact with the environment to structure species assemblages. Here we use recent theory developments to analyze plant distribution and trait data across Europe and find that plant height clustering is related to both evapotranspiration and gross primary productivity. This clustering is a signal of interspecies competition between plants, which is most evident in mid-latitude ecoregions, where conditions for growth (reflected in actual evapotranspiration rates and gross primary productivities) are optimal. Away from this optimum, climate severity likely overrides the effect of competition, or other interactions become increasingly important. Our approach bridges the gap between modern coexistence theory and large-scale species distribution data analysis.

Keywords: Ecological community dynamics Plant diversity Species coexistence Biogeographic pattern Null hypotheses testing Stochastic processes in continuous time.

13 Introduction

14 Modern coexistence theory (Chesson, 2000, HilleRisLambers *et al.*, 2011, Mayfield & Levine, 2010) is based
15 on species difference and their interplay to determine effective competitive (biotic) interactions among species in
16 natural communities. The balance between stabilizing trait differences and species dominance among competitors
17 is crucial to understand species coexistence under this framework. In communities driven by fitness differences,
18 species turn out to be clustered around similar trait values selected through competitive dominance. Therefore, trait
19 clustering may be interpreted as a fingerprint of competition even in the absence of environmental filtering (Kraft
20 *et al.*, 2015, Mayfield & Levine, 2010). Community ecology, however, still needs a comprehensive theoretical
21 framework able to describe quantitatively the role of biotic, species-to-species interactions that are relevant to
22 determine species composition and diversity across large spatial scales. Empirical studies, while they may be
23 able to independently assess environmental stress and species competitive abilities, are often limited to small
24 community sizes (Violle *et al.*, 2011) or restricted to single habitats (Kunstler *et al.*, 2012). Very few studies have
25 explored the idea of competition as a driver of community assembly across biogeographic regions (Kunstler *et al.*,
26 2016, Swenson *et al.*, 2012). Here we attempted a continent-wide macro-ecological study of species assemblage
27 patterns based on theoretical predictions from modern coexistence theory (Capitán *et al.*, 2020, Chesson, 2000,
28 Mayfield & Levine, 2010) at large geographical scales.

29 Light and water availability (Fig. 1) impose significant limitations on gross primary productivity which is re-
30 flected in actual evapotranspiration rates (Garbulsky *et al.*, 2010). These two resources vary at regional scales,
31 placing strong, sometimes opposing constraints on how tall a plant can grow within the limits of structural stabil-
32 ity. Plant height is a fundamental trait that reflects the ability of the individual to optimize its own growth within its
33 local biotic environment and regional physical constraints (see Falster & Westoby (2003), Holmgren *et al.* (1997)
34 and references therein). How plant height adapts to these opposing constraints has been studied in trees (King,
35 1990, Law *et al.*, 1997, Midgley, 2003) and herbaceous plants (Givnish, 1995, 1982). Here we analyzed presence-
36 absence matrices of floral herbaceous taxa across different European ecoregions to determine if competitive ability
37 (reflected in maximum stem height) could help explain assemblage patterns at local scales across gradients of
38 relevant environmental factors such as evapotranspiration. We examined how well observed plant assemblages at
39 macro-ecological scales match theoretical predictions generated by a synthetic, stochastic framework of commu-
40 nity assembly (Capitán *et al.*, 2015, 2017, Haegeman & Loreau, 2011, McKane *et al.*, 2000), which we described
41 in full detail in Capitán *et al.* (2020). By assuming that competition between hetero-specifics is driven by signed
42 height differences, we found a significant positive correlation between the degree of clustering and actual evapo-
43 transpiration rates. Across Europe, actual evapotranspiration is lower at more southern latitudes (due to reduced
44 precipitation levels) as well as at more northern latitudes (due to colder temperatures and low levels of sunlight).
45 Herbaceous plant height clustering is significant only over a latitudinal band where environmental constraints to
46 plant growth are weaker, which suggests that the signature of competitive dominance can only be detected in the

47 assemblage patterns of mid-latitude ecoregions.

48 **Theoretical predictions**

49 In order to make this contribution self-contained, we first provide a summary of the main predictions
50 derived by our suite of models. Recent theoretical approaches have focused on predicting analytically the ex-
51 pected fraction of species that survive in competitive scenarios (Serván *et al.*, 2018). A spatially-implicit model
52 of Lotka-Volterra type (Capitán *et al.*, 2020) allowed us to predict on average how many species are expected to
53 survive as a function of mean competitive strengths. We observed that the fraction of extant species p_c , which we
54 called “coexistence probability”, decays with the average competitive strength $\langle \rho \rangle$ as a power law above a certain
55 threshold in competition, and curves for different pool sizes S can be collapsed into the same curve following the
56 mathematical dependence,

$$p_c \sim (\langle \rho \rangle S)^{-\gamma}, \quad (1)$$

57 which was observed numerically and justified analytically (see Capitán *et al.* (2020)). We showed that the exponent
58 γ is controlled by the immigration rate μ . This is the first prediction of the spatially implicit model.

59 In order to explore the significance of competitive dominance in empirical communities, we applied first ran-
60 domization tests to model communities. In this way, we established a second prediction for this model. Null models
61 for community assembly (Chase *et al.*, 2011, Gotelli *et al.*, 2010, Webb *et al.*, 2002) compare the properties of
62 actual communities against random samples of the same size extracted from a species pool (observed diversity
63 at the ecoregion level). This approach assumes that realized communities are built up through the independent
64 arrival of equivalent species from the pool (Alonso *et al.*, 2015, MacArthur & Wilson, 1967) regardless of species
65 preferences for particular environments or species interactions. Our randomization tests were based on a single
66 statistic, the competitive strength averaged over species present in realized model communities, which were then
67 compared to random samples of the same size drawn from the species pool. The null hypothesis (i.e., empirical
68 communities are built as random assemblages from the ecoregion) can be rejected in both sides of the distribution,
69 implying signals of ‘significant trait overdispersion’ (‘clustering’) if average trait differences are larger (smaller)
70 than expected at random. In the low immigration regime, the model predicts a significant signal of clustering. This
71 regime is characterized by a low non-dimensional immigration rate ($\lambda = \mu/(\alpha K)$ much lower than 0) —here α
72 stands for the average species growth rate in isolation, and K is the carrying capacity of the environment.

73 The spatially-explicit model incorporates a trade-off between potential growth and alternative mechanisms
74 other than growth that allow shorter individuals to overcome being out-competed by taller plants (see Capitán
75 *et al.* (2020)). While the latter are better competitors for light, the former allocate more energy in allelopathic
76 compounds (Fig. 1). Height hierarchies alone, as assumed in our spatially-implicit model, lead to the selection of
77 taller plants in species assemblages. In the more realistic spatially-explicit model, species processes take place on

a lattice where locally taller plants grow faster than neighbors because they are less shaded, but in the presence of heterospecific neighbors, they are also more prone to die. Computer simulations show that the balance of these two mechanisms can end up selecting plant sizes characterized by an optimal potential height that can be either shifted toward lower or higher values depending on the choice of model parameters. This is the first prediction of the spatially-explicit model: species abundance distributions are not necessarily biased towards taller individuals, and they can peak at species at intermediate or even shorter heights. In any case, and consistently, in this more complex scenario, a balance between the gains of potential growth and the gains of energy allocation in allelopathy (as an example of a non-size-related, alternative mechanism) may result in a selection for plants exhibiting significant height clustering at stationarity.

A second result that can be derived from the spatially-explicit model is related to the persistence of trait clustering when species are aggregated over spatial scales larger than local interaction distances. Our spatially-explicit model can help explain why clustering patterns persist over large scales. The distributions of species within a region may reveal more information about the underlying assembly processes than the co-occurrence of species at any given location (Ricklefs, 2008). As species are aggregated over lattice cells of increasing size, clustering patterns hold even at scales much larger than local interaction distances. The model predicts consistent clustering patterns regardless of the aggregation scale used to define species communities. This was the second prediction, derived and carefully analyzed in Capitán *et al.* (2020), from our spatially-explicit model.

Materials and methods

Plant community data were drawn from Atlas Florae Europaeae (Jalas & Suominen, 1964–1999). The distribution of flora is geographically described using equally-sized grid cells ($\sim 50 \times 50$ km) based on the Universal Transverse Mercator projection and the Military Grid Reference System, see Fig. 2. Each cell was assigned to a dominant habitat type based on the WWF Biomes of the World classification (Olson *et al.*, 2001), which defines different ecoregions, i.e., geographically distinct assemblages of species subject to similar environmental conditions. We consider each cell in an ecoregion to represent a species aggregation.

Each herbaceous species in an ecoregion was characterized by its maximum stem height H , an eco-morphological trait that relates to several critical functional strategies among plants (Díaz *et al.*, 2015). It represents an optimal trade-off between the gains of accessing light (King, 1990, Law *et al.*, 1997), water and nutrient transport from soil (Midgley, 2003, Ryan & Yoder, 1997), and additional constraints posed by the local biotic environment of each individual plant, such as competition, facilitation, or herbivory.

Mean height values were obtained from the LEDA database (Kleyer *et al.*, 2008) for as many species as there were available in the database. Missing values were taken from (Ordóñez *et al.*, 2010) or inferred using a MICE (Multivariate Imputation by Chained Equations) approach (Buuren & Groothuis-Oudshoorn, 2011) together with a predictive mean matching algorithm based on other available traits (leaf and seed traits), genus, and growth

forms as predictors. Based on plant growth forms, 2610 herbaceous species (aquatic, herbs, or graminoid) were considered in this work.

Maximum stem height values spanned several orders of magnitude, so we used a log-transformed variable ($h = \log H$) to measure species differences (using non-transformed heights yielded comparable results, here not shown). The values of h were standardized within ecoregions as $t = (h - h_{\min}) / (h_{\max} - h_{\min})$ so that $0 \leq t \leq 1$.

Results

For all the species reported in an ecoregion, we formed an empirical competition matrix with pairwise ρ_{ij} signed height differences $\rho_{ij} = \hat{\rho}(t_j - t_i)$, where t_i are height values standardized across ecoregions and sorted in increasing order. The advantage of having these values represent trait differences between pairs of species is that any trend in competitive strengths can be immediately translated into patterns of functional clustering or overdispersion. As in Capitán *et al.* (2020), we calculated the average competitive strength as $\langle \rho \rangle = \frac{2}{S(S-1)} \sum_{i=1}^S \sum_{j=i+1}^S |\rho_{ij}|$, S standing for ecoregion richness.

In an ecoregion with richness S , a number $s_k \leq S$ of species will form a species assemblage at cell k . The coexistence probability was calculated from data as the average fraction of species that survive per cell,

$$p_c = \frac{\langle s \rangle}{S} = \frac{1}{SN_C} \sum_{k=1}^{N_C} s_k, \quad (2)$$

with N_C representing the number of cells in the ecoregion. This quantity, together with the distribution of trait differences in cells, was used to compare model predictions with real data.

Larger plants capture more resources. Therefore, evolution should favor investment in potential growth (maximum height) as a competitive mechanism. However, investment in alternative mechanisms, such as allelopathy, may help smaller plants stave off competitors, reducing local heterospecific plant cover and giving them a competitive advantage over potentially taller plant species. As a consequence, the maximum species stem height can be regarded as the outcome of an evolutionary game (Givnish, 1982) that balances opposing constraints, both physical (Craine & Dybzinski, 2013, Falster & Westoby, 2003) and biotic (King, 1990, Law *et al.*, 1997). To explore these opposing constraints, we analyzed plant data in the light of the two community assembly models. The first one is a spatially-implicit model of Lotka-Volterra type, and the second one is a straightforward spatially-explicit extension including height-driven competition and allelopathic effects. Both have been carefully defined and studied in Capitán *et al.* (2020).

Two predictions from the implicit model tested against data

The collapse of curves predicted by Eq. (1) helps eliminate the variability in S , so that empirical coexistence probabilities, which arise from different ecoregion sizes, can be fitted together (Fig. 3). Confirming the first

prediction of the spatially-implicit model, we found a significant correlation between the probability of coexistence and the scaled competitive overlap based on empirical data (Fig. 3), indicating that a model driven solely by dominant competitive interactions reliably predicts the average richness of plant communities across ecoregions. In addition, this theoretical prediction allowed an indirect estimation of the relative importance $\hat{\rho}$ of average inter- vs. intraspecific effects: the average ratio of inter- to intraspecific competition strength is about 5% (see Supporting Information, section A for details on the estimation procedure).

As a second prediction, the implicit model implies high levels of trait clustering for low immigration rates and high carrying capacity values. Importantly, this parameter regime precisely emerges from the data. In Capitán *et al.* (2020) we derived a deterministic prediction for the exponent, $\gamma = 1$, which does not match the one obtained from data ($\gamma = 0.61$). As shown in that paper, it is a non-zero (but small) value of the immigration rate that determines the power-law exponent γ being lower than 1. Indeed, for a realistic fit in Fig. 3, the exponent of the empirical power law is obtained for $\mu/\alpha \sim 0.1$ individuals per generation. Since plant communities operate in a low-immigration regime, the non-dimensional immigration rate $\lambda = \mu/(\alpha K)$ must satisfy $\lambda = 0.1/K \ll 1$, hence the carrying capacity must be large. In a regime of low immigration rate and high carrying capacity, which best fits empirical coexistence probabilities, the implicit model predicts a significant degree of species clustering [see Fig. 3 in Capitán *et al.* (2020)].

Following Triadó-Margarit *et al.* (2019), our randomization tests applied to empirical communities were based on the average competitive strength observed in a cell C formed by s species,

$$\langle \rho \rangle_C = \frac{2}{s(s-1)} \sum_{i=1}^s \sum_{j=i+1}^s |\rho_{ij}^C|, \quad (3)$$

where (ρ_{ij}^C) is the submatrix of the ecoregion competition matrix restricted to the species present in the cell. Compared to ecoregion samples, the lower (higher) the empirical community average $\langle \rho \rangle_C$ is, the higher (lower) is the degree of species clustering in the cell. For each cell we calculated the probability $p = \Pr(\langle \rho \rangle_Q \leq \langle \rho \rangle_C)$ that the competition average $\langle \rho \rangle_Q$ randomly-sampled from the pool is smaller than the empirical average. At a 5% significance level, if $p > 0.95$ the empirical competition average is significantly larger than the average measured for random pool samples, which implies that average trait differences in realized communities are larger than would be expected at random. On the other hand, if $p < 0.05$, observed trait differences are significantly smaller than would be expected at random. Therefore, if $p > 0.95$, the community exhibits ‘significant trait overdispersion’, whereas if $p < 0.05$, there is evidence for ‘significant trait clustering’ in the observed species assemblage.

Testing the second prediction against empirical observations yields a mixed picture. We calculated p -values for randomization tests applied to every cell in each ecoregion, which represent the empirical distribution of p -values (Fig. 4). At the parameter values that make plant data consistent with the first prediction, the spatially-implicit model predicts significant trait clustering. We observe that some ecoregions are consistent with this theoretical

172 expectation. However, other ecoregions clearly do not comply with this prediction. In addition, no ecoregion is
173 consistent with trait overdispersion (Fig. 4). Selecting species in randomization tests according to species dispersal
174 abilities portrays the same picture (results not shown).

175 **Ecoregion clustering and actual evapotranspiration rates**

176 In order to better quantify the propensity of an ecoregion to exhibit clustering in maximum stem height, we defined
177 a clustering index q for an ecoregion as the fraction of its cells that lie within the 5% range of significant clustering
178 (randomization tests yield p -values smaller than 0.05 for those cells). An ecoregion for which significant clustering
179 is found in most of its cells will tend to score high in the q index. We examined how the clustering index varied
180 across the continent in terms of the geographical location of ecoregion centroids as well as with actual evapotran-
181 spiration (Fig. 5). Evapotranspiration maps were obtained from data estimated through remote sensing (Mu *et al.*,
182 2011).

183 Water availability acts as a factor limiting plant growth at geographical scales (Fig. 1a), and correlates with
184 gross primary productivity (Garbulsky *et al.*, 2010), see Fig. 5d. Therefore, for a given region, mean annual evap-
185 otranspiration is a reliable measure of environmental constraints on plant growth (Garbulsky *et al.*, 2010). Panels a
186 and b of Fig. 5 show a clear latitudinal trend: there is an intermediate range of ecoregion latitudes where both clus-
187 tering indices and evapotranspiration are large, indicating that evapotranspiration measures can robustly predict
188 clustering indices (Fig. 5c). The same pattern can also be seen in the relation between mean relative height differ-
189 ences and actual evapotranspiration across individual grid cells. The intensity of the clustering pattern increases
190 with actual evapotranspiration rates across Europe, not only at the ecoregional level (Fig. 5c), but also at the lower
191 spatial scale of grid cells (see Fig. C1, Supporting Information). More importantly, since evapotranspiration is
192 a powerful proxy of environmental constraints on plant growth, this clustering in maximum stem height appears
193 to be stronger at ecoregions less limited by environmental conditions. As environments become harsher and less
194 optimal for plant growth, these clustering patterns disappear. This is particularly true for the severe climatic con-
195 ditions characteristic in the Mediterranean (with erratic rainfall, limited water availability and drought), as well as
196 of boreal zones (with low radiation incidence and cold temperatures). According to model predictions, the overall
197 clustering patterns found at middle-range latitudes are consistent with species competitive dominance controlling
198 species height differences.

199 **Two predictions from the explicit model tested against data**

200 The spatially-explicit model allows for either the dominance of tall, mid-sized or short plants, as a consequence
201 of the trade-off between investment in either potential growth or alternative mechanisms other than growth (see
202 Fig. 5 in Capitán *et al.* (2020)). We have tested whether taller or shorter plants are most commonly represented
203 in ecoregions via the correlation of cell-averaged heights and evapotranspiration (Fig. 6a), which shows a mixed

204 picture. With few exceptions, mid-latitude ecoregions exhibit positive correlation (taller plants are selected in
205 regions favoring plant growth), whereas negative dependencies are often observed in latitudinal extremes (Fig. 6b).
206 Correlations are significant but, in some cases, very weak. These results are consistent with our interpretation in
207 terms of a signal of competitive dominance in mid-latitude ecoregions.

208 Our spatially-explicit model predicts the persistence of trait clustering as species are aggregated at larger spa-
209 tial scales (much larger than the typical range of species interactions). This is important because real individual
210 plants interact at much lower spatial scales (1 to 1000ha) compared to the spatial resolution of our dataset (grid cell
211 sizes about 50 km). To assess the robustness of our results, we further investigated the effect of aggregation scales
212 on clustering patterns using plant data. In line with the spatially-implicit model, the analysis of herbaceous plant
213 communities from mid-latitude ecoregions reveals that our results are robust to both up- and down-scaling com-
214 munity sizes (see Fig 6c). Height clustering remains significant in a range of aggregated scales, and extrapolates to
215 smaller areas (under a random placement hypothesis, communities of smaller sizes were built by randomly select-
216 ing a number of species as predicted by the empirical species-area relation, see Supporting Information, section B.
217 We conclude that clustering patterns at large scales is an emerging pattern that can be interpreted as a signature of
218 competitive dominance operating at much smaller spatial scales.

219 Discussion

220 In this work we have tested predictions from modern coexistence theory (Chesson, 2000) and the competition-
221 similarity paradigm (Mayfield & Levine, 2010) using macro-ecological trait data at large spatial scales (Kunstler
222 *et al.*, 2016). While potential evapotranspiration decreases with latitude, actual evapotranspiration peaks at in-
223 termediate latitudes, and is strongly associated with higher levels of trait clustering. Critically, actual evapotran-
224 spiration is positively correlated with gross primary productivity (GPP) across terrestrial ecosystems [see Fig. 5d
225 and Garbulsky *et al.* (2010)], which also peaks at intermediate latitudes across Europe. Consistently, our results
226 were reproduced using GPP instead of ET, although both variables yield similar results. The agreement of model
227 predictions with plant community data can be interpreted as a signature of competitive dominance in empirical
228 communities in the environmentally conducive middle-range latitudes. Significant height clustering would be the
229 trace that competition leaves on community assembly pattern by filtering out subdominant species. This result
230 does not necessarily mean that competition is the main driver of community assembly. It rather highlights the
231 potential role of competitive dominance, along with other processes, in the assembly of herbaceous communities
232 at intermediate latitudes. On the contrary, as environmental conditions get increasingly extreme, no significant
233 clustering in plant height is observed. Although the interplay between facilitation and competition is far from sim-
234 ple (Hart & Marshall, 2013), the harshness of extreme conditions likely override the effects of competition, and
235 other processes such as species tolerances and facilitation (Maestre *et al.*, 2009, Valiente-Banuet & Verdú, 2007)
236 may be critical community drivers at climatic extremes.

237 Throughout this work, species assemblages within each grid cell ($\sim 50 \times 50$ km) have been defined as distinct
238 communities. Current consensus about the concept of ecological community emphasizes the importance of biotic
239 interactions. An ecological community is defined as a set of species that live in the same area and can potentially
240 interact (Stroud *et al.*, 2015). In spite of the size and heterogeneity within each grid cell at the 50×50 km spatial
241 scale, cells are much smaller than the ecoregion they belong to, and are, of course, much more homogeneous, both
242 in species composition and in environment, than the the ecoregion itself. Therefore, in principle, grid cells could
243 be regarded as communities in an operational and relative sense. In addition, we assumed that the European Flora
244 database represents species composition at a steady state, this is, we examined the stationary patterns resulting
245 from eco-evolutionary processes associated to long time scales. Although real individual plants interact at much
246 lower spatial scales, two species from the same ecoregion will eventually interact within a grid cell given enough
247 time. The larger the temporal scale, the larger is the area where two species will have a chance to interact through
248 generations and repeated dispersal events. The scale at which a set of local communities reveal information
249 about underlying assembly processes is very often the regional scale (Diniz-Filho *et al.*, 2009, Olalla-Tárraga &
250 Rodríguez, 2007, Ricklefs, 2015), which has led to the "regional community concept" (Ricklefs, 2008, 2011).

251 It is important to make a clear distinction between actual plant size and the species-level trait, "maximum
252 stem height". While a species-level trait is shaped by evolutionary constraints at longer temporal scales, actual
253 plant size is determined by a host of contingent ecological constraints operating over shorter temporal scales.
254 Although there is a large body of theory and experiments positively co-relating actual plant size and individual
255 plant competition ability (Gaudet & Keddy, 1988, Weiner, 1993), there has been considerably less attention paid
256 to the evolutionary establishment of functional trade-offs between different species-level traits (Adler *et al.*, 2014,
257 Stearns, 1989). The common wisdom that competition favors taller plants may not always hold [for instance, in
258 low-nutrient, competition-intensive, undisturbed habitats, see Tilman & Wedin (1991)]. Our analysis shows that
259 height clustering at middle-range latitudes is a fingerprint of a balance between energy invested in either potential
260 growth or other mechanisms that may help plants overcome competitors. For instance, when competitors are
261 close relatives in dense herbaceous communities, selection may favor the evolution of a low leaf height. In these
262 situations, "for short conspecific herbs to exclude competitors from a highly productive site, they must possess
263 alternative mechanisms to overcome competition, such as root competition or allelochemicals" (Givnish, 1982).
264 More generally, we argue that functional trade-offs tend to evolve in regions of higher primary productivity, where
265 the relative role of biological interactions (competition, parasitism, herbivory) is expected to be higher.

266 Competitive hierarchies have been theoretically investigated (Tilman, 1982, 2004), and empirically demon-
267 strated in herbaceous plant communities at much smaller spatial scales (Stanley Harpole & Tilman, 2006, Tilman,
268 1994, Tilman & Wedin, 1991). In these studies, a trade-off between competitive and colonization abilities has been
269 shown to maintain plant diversity, although other hierarchies have been also suggested (Muller-Landau, 2010). In-
270 terestingly, the relevant role of competitive dominance driven by species trait hierarchies has been also reported at
271 much smaller spatial scales for forest trees along an altitudinal gradient in the French Alps (Kunstler *et al.*, 2012).

Moreover, a recent study of the assembly of forest communities across East Asia shows that a phylogenetic-based species similarity index tends to be smaller the higher the minimum temperature of the coldest month is (Feng *et al.*, 2015). Although traits are not generally related to competitive abilities, and they are diverse in their functionality and in their response to environmental stress, these studies, together with our results, suggest that trait clustering is generally likely to occur where conditions for plant growth are less restrictive. Our models indicate that the process underlying this pattern is competitive dominance rather than Darwin's competition-similarity hypothesis, although it is likely that community assembly for other taxa may be driven by other biotic or environmental filters. For instance, phytoplankton communities from estuarine ecosystems (Segura *et al.*, 2012) are more consistent with Darwin's seminal hypothesis since they appear to be driven by limiting similarity creating clumpy species coexistence (Pigolotti *et al.*, 2007, Scheffer & van Nes, 2006). Competitive hierarchies are, of course, not hard-wired in nature. Intransitivities may still play a key role in maintaining diversity in some systems (Allesina & Levine, 2011, Soliveres *et al.*, 2015, Zhang & Lamb, 2012).

In Capitán *et al.* (2020) we demonstrated how different coexistence *vs.* competition curves can be collapsed into a single curve. Here we showed that model predictions were quantitatively consistent with the observed decaying behavior of the probability of local coexistence as overall competition intensity increases. This general scaling behavior is typical for stochastic community models in the presence of both symmetrical (Capitán *et al.*, 2015, 2017) and asymmetrical competition, as presented here. The scaling allowed us to give a rough estimate of $\hat{\rho}$, an average ratio of inter- *vs.* intraspecific competition (see Fig 3a). Our indirect method is only able to estimate an average $\hat{\rho}$ across ecoregions. Whenever direct empirical estimates of the ratio of inter- *vs.* intra-competition are obtained, a few similar species are typically studied using small-scale field experiments (Goldberg & Barton, 1992, Schoener, 1983). It is, therefore, unsurprising that empirical estimations of this parameter tend to be higher than ours (Kraft *et al.*, 2015), but see also Volkov *et al.* (2009) and Wang *et al.* (2016). Being able to provide rough estimates of this parameter at regional scales is also a novel result from our analysis. Our results are in agreement with a recent study of trees across six forest biomes where the authors found that trait variation is mostly related to competitive imbalances tending to drive inferior competitors to extinction (Kunstler *et al.*, 2016). Further work is required to better relate the average ratio of inter- *vs.* intraspecific competition, which stabilizes species co-existence, to plant traits, and analyze how this aggregated parameter changes at increasing spatial scales and across taxa.

In this paper we have explored several predictions from theoretical models aimed at describing plant dynamics, which have been derived and carefully studied in Capitán *et al.* (2020). In total, we have contrasted four model predictions against reported herbaceous plant diversity across Europe. As usual, our theoretical models represent a strong over-simplification of real plant community dynamics. However, in spite of disregarding the true complexity of these communities, our theory approach is useful, not only because it can reproduce macro-ecological, observational patterns with a small number of meaningful aggregated variables, but also because it provides new quantitative or qualitative predictions than may lead to new both empirical and observational studies. Finding a

307 theoretically robust and ecologically meaningful rapprochement between theory and data at relevant scales remains
308 a challenge for ecology, and we trust that our work will inspire new contributions in this direction.

309 **Acknowledgments**

310 The authors thank Mercedes Pascual for her insightful comments, and are indebted to Rohan Arthur, Han Olff,
311 Joaquín Hortal, and Fernando Valladares for their constructive criticism on earlier versions of this manuscript. This
312 work was funded by the Spanish ‘Ministerio de Economía y Competitividad’ under the projects CGL2012-39964
313 and CGL2015-69043-P (DA, JAC) and the Ramón y Cajal Fellowship program (RYC-2010-06545, DA). JAC
314 acknowledges partial financial support from the Department of Applied Mathematics (Universidad Politécnica de
315 Madrid).

316 **References**

317 1.

318 Adler, P. B., Salguero-gómez, R., Compagnoni, A., Hsu, J. S., Ray-mukherjee, J., Adler, P. B., Salguero-gómez,
319 R., Compagnoni, A., Hsu, J. S. & Ray-mukherjee, J. (2014). Correction for Adler et al., Functional traits explain
320 variation in plant life history strategies. *Proceedings of the National Academy of Sciences*, 111, 10019–10019.

321 2.

322 Allesina, S. & Levine, J. M. (2011). A competitive network theory of species diversity. *Proc. Nat. Acad. Sci.*
323 *USA*, 108, 5638–5642.

324 3.

325 Alonso, D., Pinyol-Gallemí, A., Alcoverro, T. & Arthur, R. (2015). Fish community reassembly after coral mass
326 mortality: higher trophic groups are subject to increased rates of extinction. *Ecol. Lett.*, 18, 451–461.

327 4.

328 Buuren, S. & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in r. *J. Stat.*
329 *Softw.*, 45(3).

330 5.

331 Capitán, J. A., Cuenda, S. & Alonso, D. (2015). How similar can co-occurring species be in the presence of
332 competition and ecological drift? *J. R. Soc. Interface*, 12, 20150604.

333 6.

334 Capitán, J. A., Cuenda, S. & Alonso, D. (2017). Stochastic competitive exclusion leads to a cascade of species
335 extinctions. *J. Theor. Biol.*, 419, 137–151.

- 336 7.
337 Capitán, J. A., Cuenda, S. & Alonso, D. (2020). Competitive dominance in ecological communities: Modeling
338 approaches and theoretical predictions. <http://arxiv.org/abs/2001.07284>, Under review.
- 339 8.
340 Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M. & Inouye, B. D. (2011). Using null models to disentangle
341 variation in community dissimilarity from variation in α -diversity. *Ecosphere*, 2, 24.
- 342 9.
343 Chesson, P. L. (2000). Mechanisms of maintenance of species diversity. *Ann. Rev. Ecol. Syst.*, 31, 343–366.
- 344 10.
345 Craine, J. R. & Dybzinski, R. (2013). Mechanisms of plant competition for nutrients, water and light. *Funct.*
346 *Ecol.*, 27, 833–840.
- 347 11.
348 Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C.,
349 Prentice, I. C., Garnier, E., Bönsch, G., Westoby, M., Poorter, H., Reich, P. B., Moles, A. T., Dickie, J., Gillison,
350 A. N., Zanne, A. E., Chave, J., Wright, S. J., Sheremet'ev, S. N., Jactel, H., Christopher, B., Cerabolini, B., Pierce,
351 S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J. S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M. D. &
352 Gorné, L. D. (2015). The global spectrum of plant form and function. *Nature*, 529, 1–17.
- 353 12.
354 Diniz-Filho, J. A. F., Rodríguez, M. Á., Bini, L. M., Olalla-Tarraga, M. Á., Cardillo, M., Nabout, J. C., Hortal,
355 J. & Hawkins, B. A. (2009). Climate history, human impacts and global body size of Carnivora (Mammalia:
356 Eutheria) at multiple evolutionary scales. *Journal of Biogeography*, 36, 2222–2236.
- 357 13.
358 Falster, D. S. & Westoby, M. (2003). Plant height and evolutionary games. *Trends. Ecol. Evol.*, 18, 337–343.
- 359 14.
360 Feng, G., Mi, X., Eiserhardt, W. L., Jin, G., Sang, W., Lu, Z., Wang, X., and B. Li, X. L., Sun, I., Ma, K. &
361 Svenning, J.-C. (2015). Assembly of forest communities across East-Asia. Insights from phylogenetic community
362 structure and species pool scaling. *Scientific Reports*, 5, 9337.
- 363 15.
364 Garbulsky, M. F., Peñuelas, J., Papale, D., Ardö, J., Goulden, M. L., Kiely, G., Richardson, A. D., Rotenberg,
365 E., Veenendaal, E. M. & Filella, I. (2010). Patterns and controls of the variability of radiation use efficiency and
366 primary productivity across terrestrial ecosystems. *Global Ecology and Biogeography*, 19, 253–267.

- 367 16.
368 Gaudet, C. L. & Keddy, P. A. (1988). A comparative approach to predicting competitive ability from plants trait.
369 *Nature*, 334, 242–243.
- 370 17.
371 Givnish, T. (1995). Plant stems: biomechanical adaptation for energy capture and influence on species distribu-
372 tions. In: *Plant stems: Physiology and Functional Morphology* (ed. Gartner, B.). Academic Press, Cambridge,
373 Massachusetts, pp. 3–49.
- 374 18.
375 Givnish, T. J. (1982). Adaptive significance of leaf height in forest herbs. *Am. Nat.*, 112, 279–298.
- 376 19.
377 Goldberg, D. E. & Barton, A. M. (1992). Patterns and Consequences of Interspecific Competition in Natural
378 Communities : A Review of Field Experiments with Plants. *The American naturalist*, 139, 771–801.
- 379 20.
380 Gotelli, N. J., Graves, G. R. & Rahbek, C. (2010). Macroecological signals of species interactions in the danish
381 avifauna. *Proc. Nat. Acad. Sci. USA*, 107, 5030–5035.
- 382 21.
383 Haegeman, B. & Loreau, M. (2011). A mathematical synthesis of niche and neutral theories in community
384 ecology. *J. Theor. Biol.*, 4, 263–271.
- 385 22.
386 Hart, S. P. & Marshall, D. J. (2013). Environmental stress, facilitation, competition, and coexistence. *Ecology*,
387 94, 2719–2731.
- 388 23.
389 HilleRisLambers, J., Adler, P., Harpole, W., Levine, J. & Mayfield, M. M. (2011). Rethinking Community As-
390 sembly Through the Lens of Coexistence Theory. *Annual Review of Ecology, Evolution, and Systematics*, 43,
391 120830113150004.
- 392 24.
393 Holmgren, M., Scheffer, M. & Huston, M. A. (1997). The interplay of facilitation and competition in plant
394 communities. *Ecology*, 78, 1966–1975.
- 395 25.
396 Jalas, J. & Suominen, J. (1964–1999). *Atlas Florae Europaeae. Distribution of vascular plants in Europe, Vol.*
397 *1–12*. Societas Biologica Fennica Vanamo, Helsinki.

398 26.

399 King, D. A. (1990). The adaptive significance of tree height. *Am. Nat.*, 135, 809–828.

400 27.

401 Kleyer, M., Bekker, R. M., Knevel, I. C., Bakker, J. P., Thompson, K., Sonnenschein, M., Poschlod, P., Groe-
402 nendael, J. M. V., Klimes, L., Klimesová, J., Klotz, S., Rusch, G. M., Hermy, M., Adriaens, D., Boedeltje, G.,
403 Bossuyt, B., Dannemann, A., Endels, P., Götzenberger, L., Hodgson, J. G., Jackel, A.-K., Kühn, I., Kunzmann,
404 D., Ozinga, W. A., Römermann, C., Stadler, M., Schlegelmilch, J., Steendam, H. J., Tackenberg, O., Wilmann, B.,
405 Cornelissen, J. H. C., Eriksson, O., Garnier, E. & Peco, B. I. (2008). The leda traitbase: a database of life-history
406 traits of the northwest european flora. *Journal of Ecology*, 96, 1266–1274.

407 28.

408 Kraft, N. J. B., Godoy, O. & Levine, J. M. (2015). Plant functional traits and the multidimensional nature of
409 species coexistence. *Proceedings of the National Academy of Sciences*, 112, 797–802.

410 29.

411 Kunstler, G., Falster, D., Coomes, D. A., Hui, F., Kooyman, R. M., Laughlin, D. C., Poorter, L., Vanderwel,
412 M., Vieilledent, G., Wright, S. J., Aiba, M., Baraloto, C., Caspersen, J., Cornelissen, J. H. C., Gourlet-Fleury,
413 S., Hanewinkel, M., Herault, B., Kattge, J., Kurokawa, H., Onoda, Y., Peñuelas, J., Poorter, H., Uriarte, M.,
414 Richardson, S., Ruiz-Benito, P., Sun, I.-F., Stahl, G., Swenson, N. G., Thompson, J., Westerlund, B., and
415 M. A. Zavala, C. W., Zeng, H., Zimmerman, J. K., Zimmermann, N. E. & Westoby, M. (2016). Plant functional
416 traits have globally consistent effects on competition. *Nature*, 529, 204–207.

417 30.

418 Kunstler, G., Lavergne, S., Courbaud, B., and G. Vieilledent, W. T., Zimmermann, N. E., Kattge, J. & Coomes,
419 D. A. (2012). Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic
420 or functional similarity: implications for forest community assembly. *Ecol. Lett.*, 15, 831–840.

421 31.

422 Law, R., Marrow, P. & Dieckmann, U. (1997). On evolution under asymmetric competition. *Evol. Ecol.*, 11,
423 485–501.

424 32.

425 MacArthur, R. H. & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton University Press,
426 Princeton.

427 33.

428 Maestre, F. T., Callaway, R. M., Valladares, F. & Lortie, C. J. (2009). Refining the stress-gradient hypothesis for
429 competition and facilitation in plant communities. *Journal of Ecology*, 97, 199–205.

- 430 34.
- 431 Mayfield, M. M. & Levine, J. M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure
432 of communities. *Ecol. Lett.*, 13, 1085–1093.
- 433 35.
- 434 McKane, A. J., Alonso, D. & Solé, R. V. (2000). A mean field stochastic theory for species rich assembled
435 communities. *Phys. Rev. E*, 62, 8466–8484.
- 436 36.
- 437 Midgley, J. J. (2003). Is bigger better in plants? The hydraulic costs of increasing size in trees. *Trends. Ecol.*
438 *Evol.*, 18, 5–6.
- 439 37.
- 440 Mu, Q., Zhao, M. & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration
441 algorithm. *Remote Sens. Environ.*, 115, 1781–1800.
- 442 38.
- 443 Muller-Landau, H. (2010). The tolerance-fecundity trade-off and the maintenance of diversity in seed size. *Proc.*
444 *Nat. Acad. Sci. USA*, 107, 4242–4247.
- 445 39.
- 446 Olalla-Tárraga, M. Á. & Rodríguez, M. Á. (2007). Energy and interspecific body size patterns of amphibian
447 faunas in Europe and North America: Anurans follow Bergmann’s rule, urodeles its converse. *Global Ecology*
448 *and Biogeography*, 16, 606–617.
- 449 40.
- 450 Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D’amico,
451 J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux,
452 J. F., Wettengel, W. W., Hedao, P. & Kassem, K. R. (2001). Terrestrial ecoregions of the world: A new map of life
453 on Earth. *BioScience*, 51, 993–938.
- 454 41.
- 455 Ordóñez, A., Wright, I. J. & Olff, H. (2010). Functional differences between native and alien species: A global-
456 scale comparison. *Fun. Ecol.*, 24, 1353–1361.
- 457 42.
- 458 Pigolotti, S., López, C. & Hernández-García, E. (2007). Species clustering in competitive Lotka-Volterra models.
459 *Phys. Rev. Lett.*, 98, 1–4.
- 460 43.
- 461 Ricklefs, R. E. (2008). Disintegration of the ecological community. *American Naturalist*, 172, 741–750.

- 462 44.
- 463 Ricklefs, R. E. (2011). Applying a regional community concept to forest birds of eastern North America. *Pro-*
464 *ceedings of the National Academy of Sciences of the United States of America*, 108, 2300–2305.
- 465 45.
- 466 Ricklefs, R. E. (2015). Intrinsic dynamics of the regional community. *Ecology Letters*, 18, 497–503.
- 467 46.
- 468 Ryan, M. G. & Yoder, B. J. (1997). Hydraulic limits to tree height and tree growth. *Bioscience*, 47, 235–242.
- 469 47.
- 470 Scheffer, M. & van Nes, E. H. (2006). Self-organized similarity, the evolutionary emergence of groups of similar
471 species. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 6230–6235.
- 472 48.
- 473 Schoener, T. W. (1983). Field Experiments on Interspecific Competition. *The American naturalist*, 122, 240–285.
- 474 49.
- 475 Segura, A. M., Kruk, C., Calliari, D., García-Rodríguez, F., Conde, D., Widdicombe, C. E. & Fort, H. (2012).
476 Competition drives clumpy species coexistence in estuarine phytoplankton. *Scientific Reports*, 3, 1037.
- 477 50.
- 478 Serván, C. A., Capitán, J. A., Grilli, J., Morrison, K. E. & Allesina, S. (2018). Coexistence of many species in
479 random ecosystems. *Nature Ecology and Evolution*, 2, 1237.
- 480 51.
- 481 Soliveres, S., Maestre, F. T., Ulrich, W., Manning, P., Boch, S., Bowker, M. A., Prati, D., Delgado-Baquerizo,
482 M., Quero, J. L., Schöning, I., Gallardo, A., Weisser, W., Müller, J., Socher, S. A., García-Gómez, M., Ochoa, V.,
483 Schulze, E. D., Fischer, M. & Allan, E. (2015). Intransitive competition is widespread in plant communities and
484 maintains their species richness. *Ecology Letters*, 18, 790–798.
- 485 52.
- 486 Stanley Harpole, W. & Tilman, D. (2006). Non-neutral patterns of species abundance in grassland communities.
487 *Ecology Letters*, 9, 15–23.
- 488 53.
- 489 Stearns, S. C. (1989). Evolution in life-history Trade-offs. *Functional Ecology*, 3, 259–268.
- 490 54.
- 491 Stroud, J. T., Bush, M. R., Ladd, M. C., Nowicki, R. J., Shantz, A. A. & Sweatman, J. (2015). Is a community still
492 a community? Reviewing definitions of key terms in community ecology. *Ecology and Evolution*, 5, 4757–4765.

- 493 55.
- 494 Swenson, N. G., Enquist, B. J., Pither, J., Kerkhoff, A. J., Boyle, B., Weiser, M. D., Elser, J. J., Fagan, W. F.,
 495 Forero-Montaña, J., Fyllas, N., Kraft, N. J. B., Lake, J. K., Moles, A. T., Patiño, S., Phillips, O. L., Price, C. A.,
 496 Reich, P. B., Quesada, C. A., Stegen, J. C., Valencia, R., Wright, I. J., Wright, S. J., Andelman, S., Jorgensen,
 497 P. M., Lacher, T. E., Monteagudo, A., Núñez-Vargas, M. P., Vasquez-Martínez, R. & Nolting, K. M. (2012). The
 498 biogeography and filtering of woody plant functional diversity in North and South America. *Global Ecology and*
 499 *Biogeography*, 21, 798–808.
- 500 56.
- 501 Tilman, D. (1982). *Resource Competition and Community Structure*. Princeton University Press, Princeton, NJ.
- 502 57.
- 503 Tilman, D. (1994). Competition and Biodiversity in Spatially Structured Habitats. *Ecology*, 75, 2–16.
- 504 58.
- 505 Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource compe-
 506 tition, invasion, and community assembly. *Proc. Nat. Acad. Sci. USA*, 101, 10854–10861.
- 507 59.
- 508 Tilman, D. & Wedin, D. (1991). Plant Traits and Resource Reduction For Five Grasses Growing on a Nitrogen
 509 Gradient. *Ecology*, 72, 685–700.
- 510 60.
- 511 Triadó-Margarit, X., Capitán, J. A., Menéndez-Serra, M., Ortiz-Álvarez, R., Ontiveros, V. J., Casamayor, E. O.
 512 & Alonso, D. (2019). A Randomized Trait Community Clustering approach to unveil consistent environmental
 513 thresholds in community assembly. *The ISME Journal*, 13.
- 514 61.
- 515 Valiente-Banuet, A. & Verdú, M. (2007). Assembly through facilitation can increase the phylogenetic diversity of
 516 plant communities. *Ecol. Lett.*, 10, 1029–1036.
- 517 62.
- 518 Violle, C., Nemegut, D. R., Pu, Z. & Jiang, L. (2011). Phylogenetic limiting similarity and competitive exclusion.
 519 *Ecol. Lett.*, 14, 782–787.
- 520 63.
- 521 Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. (2009). Inferring species interactions in tropical forests.
 522 *Proc. Nat. Acad. Sci. USA*, 106, 13854–13859.
- 523 64.
- 524 Wang, X. G., Wiegand, T., Kraft, N. J. B., Swenson, N. G., Davies, S. J., Hao, Z. Q., Howe, R., Lin, Y. C., Ma,

- 525 K. P., Mi, X. C., Su, S. H., Sun, I. F. & Wolf, A. (2016). Stochastic dilution effects weaken deterministic effects
526 of niche-based processes in species rich forests. *Ecology*, 97, 347–360.
- 527 65.
- 528 Webb, C. O., Ackerly, D. D., McPeck, M. A. & Donoghue, M. J. (2002). Phylogenies and community ecology.
529 *Ann. Rev. Ecol. Syst.*, 33, 475–505.
- 530 66.
- 531 Weiner, J. (1993). Competition among plants. *Treballs de la Societat Catalana de Biologia*, 44, 99–109.
- 532 67.
- 533 Zhang, S. & Lamb, E. G. (2012). Plant competitive ability and the transitivity of competitive hierarchies change
534 with plant age. *Plant Ecology*, 213, 15–23.

Figure captions

Figure 1. Conceptual framework for maximum height resulting from a trade-off between investing energy either in potential growth, or in any other alternative, non-size-related strategy. In panel **a**, we illustrate latitudinal patterns of potential light and water availability. The latitudinal gradient of actual evapotranspiration (ET) is also shown along with the expected role of biotic interactions in determining community dynamics. At middle-range latitudes, we expect competitive hierarchies to be at their maximum due to a greater relative role of species interactions. Panel **b** shows how the trade-off between potential growth and any alternative mechanism not related to size can be included in a spatially-explicit model: species that are either good at growing taller or in investing energy in allelopathy remain short, but cause incremental death of their heterospecific neighbors. As an outcome of this trade-off, the model predicts the dominance of taller, mid-sized, or shorter plants at stationarity (panel **c**).

Figure 2. Geographical description of plant data across European ecoregions. **a**, 25 different habitats covering most of Europe are shown in the map and listed below. Ecoregions are regarded as a pool comprising all plant species observed in that region. **b**, The Military Grid Reference System divides ecoregions in grid cells, each one considered as an assemblage formed by a species sample of the pool.

Figure 3. The implicit model predicts a power-law decay regardless of the ecoregion size S , which permits fitting a power law to data ($r^2 = 0.51$, $p < 10^{-3}$, 95% confidence lines are shown). In order to match the empirical exponent γ we need to choose the immigration rate $\mu = 5$, the net growth rate $\alpha = 50$ and the carrying capacity $K = 1000$. To match the starting point of the decay we need to set $\hat{\rho} = 0.04$ in the calculation of ρ_{ij} . For completeness, we have reproduced here model expectations (triangles) for different pool sizes. Data colors match ecoregion codes in Fig. 2.

Figure 4. Empirical randomization tests.. The majority of ecoregions are consistent with model predictions as the distributions (Tukey boxplots) lie in the 5% range of significant clustering (Methods). Data colors in panels **a** and **c** match codes in Fig. 2.

Figure 5. Linking height clustering to geographical and environmental variables. **a**, Variation in the clustering index (q) with latitude (φ). Quadratic fit: $r^2 = 0.63$, $p < 10^{-3}$. **b**, Latitudinal variation in mean annual actual evapotranspiration (ET) data. Quadratic weighted regression: $r^2 = 0.63$, $p < 10^{-3}$. The shaded areas in panels **a** and **b** represent the latitudinal range for which the adjusted dependence $q(\varphi) \geq 0.7$, where both height clustering and evapotranspiration are maximal. **c**, Linear weighted regression for ET as a function of the clustering index; $r^2 = 0.49$, $p < 10^{-3}$. **d**, Correlation between mean gross primary productivity (GPP) and mean annual ET; linear weighted fit: $r^2 = 0.73$, $p < 10^{-3}$. In the first four panels, the radius of each circle is proportional to the clustering index. Symbol colors refer to ecoregions (Fig. 2). All the fits show the 95% confidence bands. **e**, Geographical distribution of clustering indices for ecoregions across Europe.

Figure 6. Two predictions of the explicit model tested against data. **a**, Correlation of cell-averaged height (relative to ecoregion means) and mean annual ET by ecoregion (colors used for data match codes in Fig. 2). **b**, Correlation coefficient obtained in **a** vs. latitude. Circle radii are proportional to clustering indices. Observe that positive correlations tend to associate with high clustering index (with some exceptions) and middle-range latitude (quadratic fit: $r^2 = 0.44$, $p = 0.001$). **c**, Clustering patterns of an ecoregion characterized by high clustering index (Atlantic mixed forests) were analyzed at increasing aggregation scales. Communities were defined by increasingly aggregating contiguous 50×50 km cells. Below a critical aggregation scale (eleventh log-area bin, which corresponds to 10^5 km²), randomization tests show strong signals of clustering. The inset in **c** represents a down-scaling of randomization tests. Clustering patterns robustly persist at smaller spatial scales.

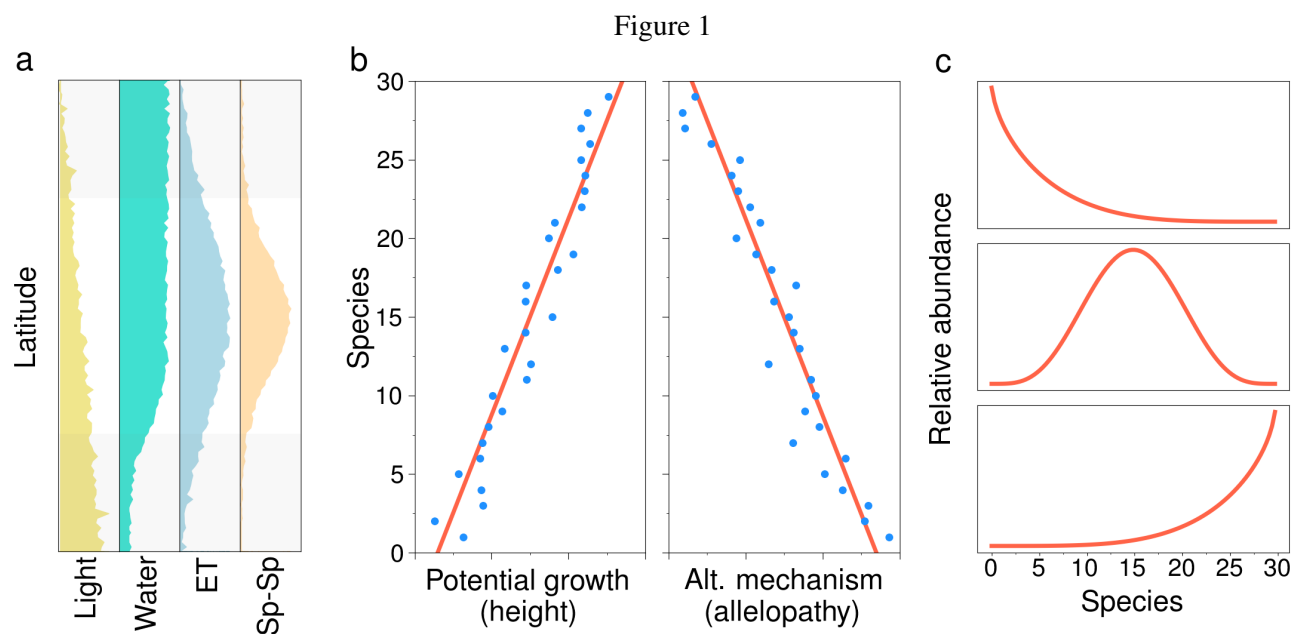


Figure 2

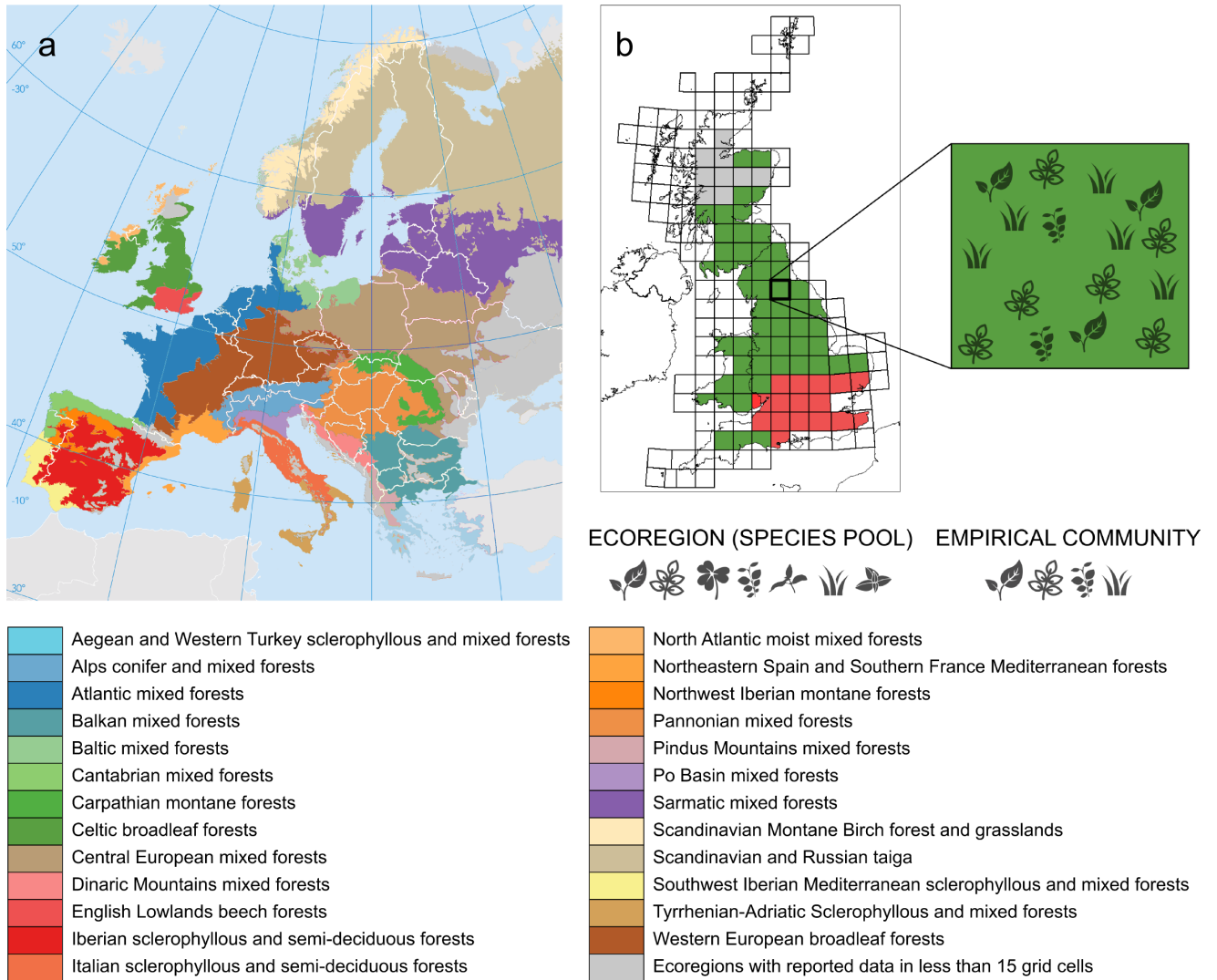


Figure 3

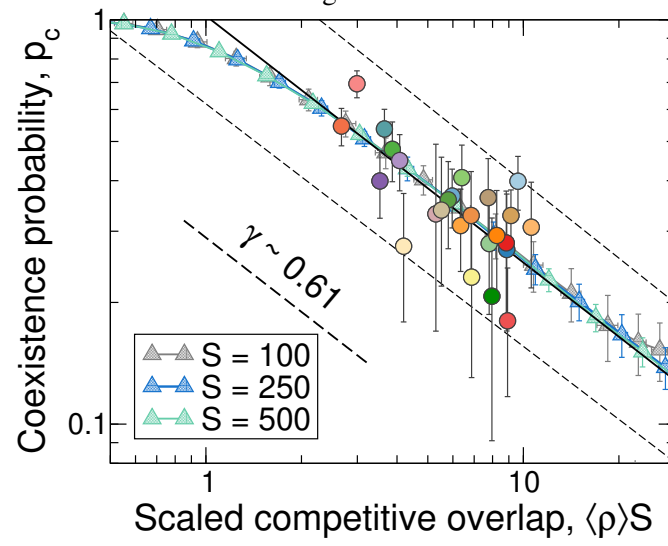


Figure 4

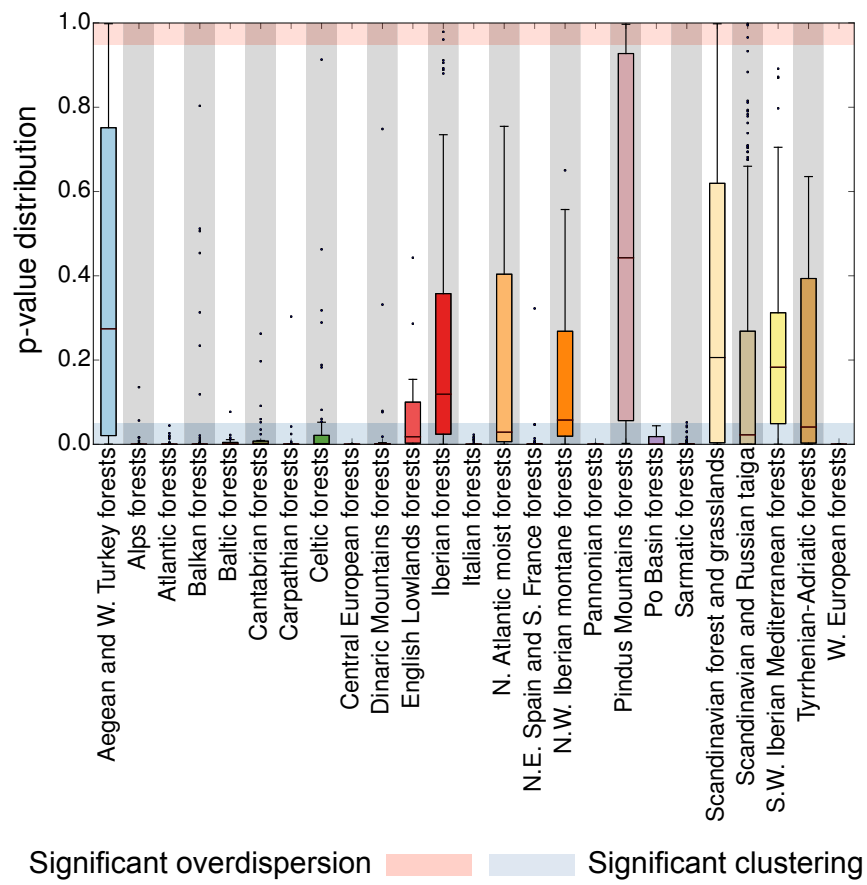


Figure 5

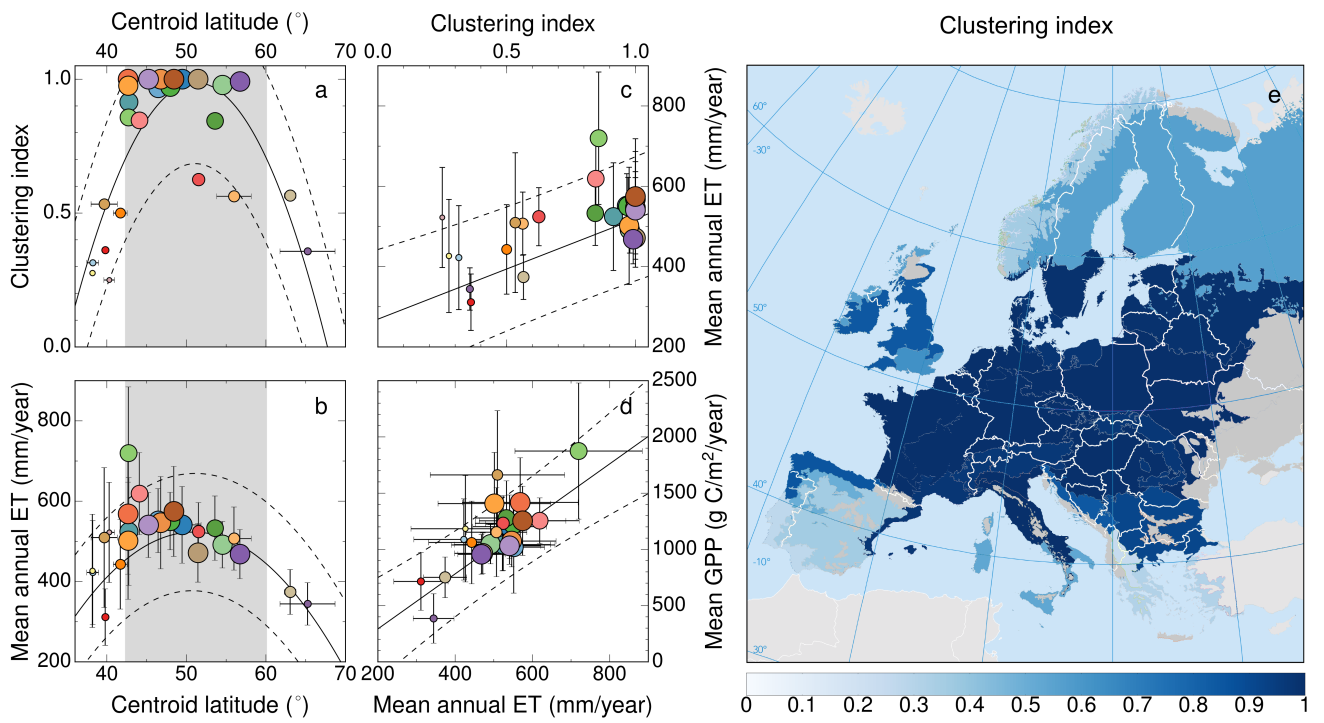


Figure 6

