

1 Original Article

2 **The role of seed traits in determining the phylogenetic structure of temperate plant**  
3 **communities**

4

5 Filip Vandeloos<sup>1,2,\*</sup>, Miguel Verdú<sup>3</sup>, Olivier Honnay<sup>2</sup>

6 <sup>1</sup> Laboratory of Plant Ecology, Philipps-Universität Marburg, Karl-von-Frisch-Strasse 8, D-  
7 35043 Marburg, Germany.

8 <sup>2</sup> Laboratory of Plant Ecology, Katholieke Universiteit Leuven, Kasteelpark Arenberg 31,  
9 Heverlee 3001, Belgium

10 <sup>3</sup> Centro de Investigaciones sobre Desertificación (CIDE, CSIC-UV), Carretera Moncada -  
11 Náquera, Km. 4,5 46113 Moncada (Valencia) Spain

12

13 Running title: Phylogenetic structure of temperate plant communities

14

15 \* e-mail: [filip.vandeloos@biologie.uni-marburg.de](mailto:filip.vandeloos@biologie.uni-marburg.de)

16

17

18

19

20

21

22

23

24

25

1 **Abstract**

2 • *Background and Aims* Phylogenetic clustering of species within plant communities can be  
3 expected to result from environmental filtering acting on an evolutionary conserved plant  
4 trait. One such a candidate trait is the embryo to seed size ratio (E:S). A high E:S may allow  
5 faster germination immediately after imbibition, and is therefore assumed to be advantageous  
6 in dry habitats.

7 • *Methods* We tested the hypothesis that habitat filtering driven by soil moisture conditions  
8 and acting on seed germination and seedling establishment is an important ecological  
9 mechanism in structuring temperate plant communities. Taxa with similar E:S are expected to  
10 co-occur in the same habitat types, resulting in both a phenotypically and phylogenetically  
11 clustered community structure.

12 • *Key Results* Genera with similar E:S tend to co-occur, as low and high E:S genera dominate  
13 in moist and dry habitats respectively. A phylogenetically clustered pattern of community  
14 structure was evident, and dispersion of E:S was positively related to phylogenetic dispersion.

15 • *Conclusions* The phenotypically and phylogenetically clustered pattern indicates that E:S  
16 mediated habitat filtering is an important assembly process structuring the plant community of  
17 the temperate climate habitats studied.

18

19 **Key words**

20 Calcareous grassland, dune slacks, embryo size, habitat filtering, phylogenetic structure, plant  
21 community structure, seed mass, temperate forest

22

23

24

25

## 1 **Introduction**

2 Together with the availability of well resolved species phylogenies and of quantitative tools to  
3 analyse them, awareness has risen that the phylogenetic structure of local communities may  
4 give insight in processes governing community assembly (Hardy and Senterre, 2007; Vamosi  
5 *et al.*, 2009; Pausas and Verdú, 2010). Coexistence of closely related species within plant  
6 communities, i.e. phylogenetic clustering, can be expected to result from environmental  
7 filtering acting on a conserved plant trait (Webb *et al.*, 2002) or from divergence of trait-states  
8 (Prinzing *et al.*, 2008). If trait-states, the values a trait can take, diverge among closely related  
9 species, then they will not compete and their coexistence will be allowed (Beltrán *et al.*, 2012;  
10 Verdú *et al.*, 2012).

11  
12 The framework of Webb *et al.* (2002) also implies that competitive exclusion of closely  
13 related species can result in phylogenetic evenness. Recent insights, however, have shown  
14 that competitive exclusion may also result in phylogenetic clustering, depending on  
15 competitive ability and the magnitude of niche differences between species (Mayfield and  
16 Levine, 2010). Competition can exclude distantly related species and result in a clustered  
17 community structure when traits advantageous in competition are phylogenetically conserved.  
18 Moreover, habitat filtering and competitive exclusion might interact, i.e. species that are not  
19 adapted to the prevailing habitat conditions are filtered out and this is confounded by  
20 increased competition (Grime, 2006). To overcome these complicating issues, it has been  
21 suggested to broaden the definition of environmental filtering by including competitive  
22 exclusion due to competitive ability differences (Mayfield and Levine, 2010). Other processes  
23 such as facilitation, mutualism and predation have also been shown to result in non-random  
24 phylogenetic community structure (Valiente-Banuet and Verdú, 2007; Vamosi and Vamosi,  
25 2007; Elias *et al.*, 2008). Not only phylogenetic, but also trait-based tests may inform us about

1 the underlying community assembly processes, but have seldom been applied to the same  
2 community (Kraft and Ackerly, 2010).

3

4 The regeneration phase, and seedling establishment in particular, is a major bottleneck in the  
5 plants' life history, potentially affecting the species composition of established plant  
6 communities (Hanley, 1995). The inability of seedlings to establish under certain  
7 environmental conditions can be expected to act as a strong habitat filter (Vandervalk, 1981;  
8 Lord *et al.*, 1995). This filtering process may result in a phenotypically and phylogenetically  
9 clustered pattern of community structure, as was clearly shown in Mediterranean plant  
10 communities, where filtering of species in fire prone habitats depends partially on the ability  
11 of seeds to germinate after fires (Verdú and Pausas, 2007; Ojeda *et al.*, 2010). The size,  
12 composition and location (in the cotyledons or in extra-embryonic tissues) of the food  
13 reserves in seeds play an important role in the timing of seed germination, the seedling growth  
14 rate and competitiveness of seedlings (Hanley *et al.*, 2004; Soriano *et al.*, 2011). Therefore,  
15 two functional seed traits that can profoundly affect seed germination and seedling  
16 establishment are seed mass and the embryo to seed size ratio (E:S). The latter trait could very  
17 well affect plant community composition, since storing food reserves in the cotyledons may  
18 allow faster germination and increase seedling vigour (Stebbins, 1974). Faster germination  
19 would be particularly beneficial in dry habitats, enabling seedlings to make maximal use of  
20 short periods of moisture (Hodgson and Mackey, 1986; Vandelook *et al.*, 2012). Additionally,  
21 many species with endospermic seeds require a prolonged period of post-dispersal embryo  
22 growth before seeds can germinate, excluding these species from habitats subject to extensive  
23 dry periods (Fenner and Thompson, 2005).

24

1 A strong phylogenetic signal and considerably more variation in E:S among families than  
2 within families, both indicate that E:S is an evolutionarily conserved trait among angiosperms  
3 (Forbis *et al.*, 2002; Verdú, 2006). A low E:S is considered an ancestral feature in  
4 angiosperms, while an evolutionary trend towards increasing E:S resulted in high values in  
5 several derived taxa. Analysis of the tempo of evolution indicated that E:S changed very fast  
6 early in angiosperm evolution, and that it remained stable later (Verdú, 2006).

7

8 The role of seed mass in structuring plant communities has already been extensively explored.  
9 Many studies showed that seed mass is related to light availability, whereby large seeded  
10 species have an advantage under low light conditions (e.g. Leishman and Westoby, 1994;  
11 Hodkinson *et al.*, 1998; Thompson and Hodkinson, 1998). Relationships between seed mass  
12 and soil moisture, nutrients and acidity have also been observed, but evidence is more limited  
13 (Leishman *et al.*, 2000). Moreover, the relationship between seed mass and nutrient  
14 availability has been shown to depend on habitat-specific differences (Hanley *et al.*, 2007).  
15 Reports about the extent of phylogenetic signal in seed size are contradicting, as similarities  
16 in seed sizes between relatives lower (Zanne, 2005) and higher (Mazer, 1989; Prinzing *et al.*,  
17 2008) than expected by chance have been recorded.

18

19 Here we test the general hypothesis that habitat filtering acting on seedling establishment is an  
20 important ecological mechanism in structuring temperate plant communities. Since E:S is a  
21 highly conserved trait, we can expect two patterns emerging if our hypothesis is true. First,  
22 species with similar E:S will tend to co-occur in the same habitat type, resulting in a  
23 phenotypically clustered community structure. Second, under the habitat filtering hypothesis,  
24 a phylogenetically clustered plant community structure can be expected (Webb *et al.*, 2002).

1 Additionally we expect that seed mass, assuming that it has a lower phylogenetic signal, also  
2 contributes significantly to the phenotypic, but less to the phylogenetic community structure.

3

4 We tested our hypotheses over three very different habitats: wet dune slacks, mesic deciduous  
5 forests and dry calcareous grasslands, representing a large gradient in light and soil moisture  
6 conditions. This approach has the advantage that co-occurrence of phylogenetically related  
7 species under similar climatic conditions provides additional evidence for habitat filtering.

8 We also examined the alternative hypothesis that coexistence of closely related species is  
9 possible due to divergence of trait-states (Prinzing *et al.*, 2008). If this is true, the measures of  
10 phylogenetic and trait-state dispersion per plot should be negatively correlated indicating that  
11 closely related species have evolved divergent trait states. Finally, we tested the additional  
12 hypothesis that, as a small E:S is an ancestral condition that changed very fast early in the  
13 angiosperm evolution, the evolutionary signal in the communities with a non-random  
14 phenotypic and phylogenetic structure can be traced back towards the base of the phylogeny.

15

## 16 **Materials and Methods**

### 17 **Data collecting**

#### 18 *Study area and vegetation sampling*

19 Vegetation samplings were retrieved from previous studies, more details on sampling  
20 methodology can be found in the respective references (Table 1). To exclude large scale  
21 biogeographical effects, the selected habitats were all located within 200 km of each other  
22 (50°02'N - 51°05'N; 2°32'E - 4°38'E). Although the three habitats studied are situated in the  
23 same climate region, they are clearly distinct with respect to the soil moisture content and  
24 light availability. The dune slacks sampled are located in four nature reserves along the  
25 Belgian and north-western French coasts (Bossuyt *et al.*, 2003). A total of 83 dune slacks,

1 ranging from 20 m<sup>2</sup> to 0.13 km<sup>2</sup>, were surveyed by sampling a variable number of plots in  
2 each dune slack. The calcareous grasslands surveyed are situated in the Calestienne region  
3 along the Viroin valley. The vegetation was sampled in multiple plots distributed over 54  
4 fragments of calcareous grassland scattered over a region of approximately 4 km wide and 20  
5 km long (Butaye *et al.*, 2005). Two deciduous forests, considered as one habitat type, were  
6 included in this study. Langerodebos is located about 8 km southwest of Leuven (Honnay,  
7 unpublished data). The second forest, Vorte Bossen, is located 20 km to the southwest of  
8 Brugge (Lameire *et al.*, 2000).

9  
10 The vegetation in the different habitats was sampled in plots ranging from 1 x 1 m in the dune  
11 slacks and the calcareous grasslands, up to 10 x 10 m and 10 x 15 meter in the forests. Plot  
12 size was not correlated with any of the metrics for phylogenetic structure and therefore  
13 confounding effects of plot size with phylogenetic structure can be discarded. A variable  
14 number of plots from the different habitats were included in the study (Table 1). Planted  
15 *Populus x canadensis* trees were not taken into account, not were ferns and two  
16 gymnosperms, because of differences related to spore and seed morphology. Ultimately, a  
17 total of 773 plots containing presence/absence data for 233 genera (402 species) were  
18 included in the analysis [**Supplementary information**].

19  
20 *Trait data*

21 E:S values of the genera sampled were obtained from Forbis *et al.*, (2002). These are  
22 measurements of seed and embryo size from illustrations of the internal morphology of seeds  
23 drawn by Martin (1946). For genera in our study that were also represented in Martins' data-  
24 set (74.8 %), we used the average E:S of all species in the genus [**Supplementary**  
25 **information**]. In the case of missing genera, we assigned the average E:S of all species from

1 the same family. This generalization will have only a minor effect on the final results, since  
2 Forbis *et al.*, (2002) already calculated that variation in E:S among families ( $> 92.5\%$ ) is  
3 significantly larger as compared to within families ( $< 7.5\%$ ). Including genera for which E:S  
4 values were based on family means had a negligible effect on the phylogenetic signal (sensu  
5 Blomberg *et al.*, 2003). For 177 genera we were also able to retrieve E:S data from other  
6 literature sources. A linear regression of the alternative data with the Martins' data set  
7 confirmed the validity of the latter ( $t = 54.5$ ,  $r = 0.97$ ,  $P < 0.001$ ). Seed mass data were  
8 obtained from the Seed Information Database (<http://data.kew.org/sid/>). These data were  
9 converted from the species to the genus level by averaging data of all congeneric species  
10 sampled.

11

#### 12 *Environmental data*

13 We used Ellenberg indicator values to obtain an estimate of soil moisture, light availability  
14 and nutrient level for each of the 773 plots (Ellenberg *et al.*, 1992). Ellenberg values have  
15 been successfully used in other studies of phylogenetic plant community structure in Central  
16 Europe because they correlate well with environmental variables (Dzwonko, 2001). Ellenberg  
17 values were calculated using presence/absence data for each plot.

18

#### 19 **Data analysis**

##### 20 *Phenotypic community structure*

21 Mean trait values per plot were calculated for all three habitat types and compared using a  
22 one-way ANOVA. We determined the phenotypic community structure by testing whether  
23 seeds of co-occurring genera have similar trait values (Cavender-Bares *et al.*, 2004). Pairwise  
24 values of co-occurrence in all plots were calculated using a binary distance index. A  
25 Euclidean distance index was used to calculate pairwise distances between trait values of all



1 species. Seed mass data were log transformed prior to analyses. The resulting co-occurrence  
2 and phenotypic distance matrices were correlated using a Mantel test with 999 permutations  
3 (Legendre *et al.*, 1994). The analyses were performed with the ade4 software (Thioulouse *et*  
4 *al.*, 1996) implemented in R version 2.7.2 (R Development Core Team).

5

#### 6 *Phylogenetic community structure*

7 For consistency with the phenotypic data, we chose the genus as the taxonomic level of study.  
8 We also ran all the analyses at the species level and results were qualitatively identical to  
9 those at the genus level, keeping the conclusions unaltered. A phylogenetic tree containing all  
10 233 genera was assembled using the Phylomatic 2 software implemented in Phylocom 4.1  
11 (Webb *et al.*, 2009). This tree was obtained by matching the genus and family names of the  
12 sampled taxa with the angiosperm mega-tree (APG3; version R20091110.new), which is  
13 based on the work of the Angiosperm Phylogeny Group (Stevens, 2001). Since phylomatic  
14 only provides good resolution at the family level, infra-familiar resolution was increased by  
15 manually resolving nodes based on literature data [**Supplementary information**]. Finally we  
16 obtained a phylogenetic megatree of which 97% of the nodes were resolved. This tree was  
17 dated according to Wikström *et al.* (2001), using the *bladj* function in phylocom (Webb *et al.*,  
18 2009).

19

20 The phylogenetic structure of the regional pool was tested using different metrics, enabling us  
21 to characterize multiple aspects of community structure (Vamosi *et al.*, 2009; Pausas and  
22 Verdú, 2010). The first metric we used to analyse the phylogenetic composition of individual  
23 plots was the net relatedness index (NRI; Webb *et al.*, 2002). The NRI is a standardized  
24 measure of the mean phylogenetic distance (MPD) of taxa in a sample, relative to a  
25 phylogeny of a specified regional pool of taxa. The NRI increases with an increase in

1 clustering and becomes negative in case of overdispersion. The regional pool containing the  
2 genera in all plots was used as the reference pool. Decreasing the regional pool to the habitat  
3 level did not significantly affect the results obtained. To test whether NRI differed  
4 significantly from random expectation, the results were compared to those obtained by a null  
5 model created by randomly shuffling genera across the entire phylogeny. The results obtained  
6 were consistent with those of other null models. Calculation of NRI and the comparison to the  
7 null model was performed using the “comstruct” function in the Phylocom package.

8

9 Phylogenetic dispersion and dispersion of trait states per plot were correlated to test whether  
10 co-occurrence of closely-related taxa is the product of trait divergence instead of habitat  
11 filtering (*sensu* Prinzing *et al.*, 2008). Trait dispersion was calculated as a standardized  
12 standard deviation (SD) of the E:S or seed size in each plot (following Prinzing *et al.*, 2008).  
13 Phylogenetic dispersion was estimated as the inverse of NRI. The relative importance of each  
14 trait in the model was assessed using the relimp package in R (Firth, 2008). This method  
15 estimates the relative importance of two predictors in a glm model as the square root of the  
16 variance-ratio quantity.

17

18 The second metric we applied to determine phylogenetic structure separates phylogenetic  
19 diversity into  $\alpha$ - (within-site) and  $\beta$ - (among-site) components (Hardy and Senterre, 2007).

20 Using the presence/absence data we calculated  $\Pi_{ST}$ , which represents the relative increase of  
21 the mean phylogenetic distance between species sampled among sites versus within sites. To  
22 determine whether phylogenetic clustering or overdispersion occurred, the  $\Pi_{ST}$  value was  
23 compared to the  $p\Pi_{ST}$  value that was obtained after randomizing all species among the tips of  
24 the phylogenetic tree (999 permutations). After complete randomization,  $p\Pi_{ST}$  has a statistical  
25 expectation equal to zero. Clustering and overdispersion are observed when  $\Pi_{ST} > 0$  and  $\Pi_{ST} <$

1 0, respectively. Partial tree randomizations were performed on the total species pool to detect  
2 whether the observed community structure at the regional level is due to early or late  
3 branching in the phylogenetic tree. We included 15 different age thresholds, ranging from 10  
4 Myr to 227 Myr, above which the tree was kept intact. The selected age threshold coincided  
5 with splits between major clades. We also randomized particular clades within the tree (Table  
6 3), which allows us to test whether particular clades contribute significantly to under- or  
7 overdispersion. All these analysis were performed using the Spacodi software (SPACoDi\_0-  
8 10) provided by Olivier Hardy (Hardy and Senterre, 2007; Hardy and Jost, 2008).

9

## 10 **Results**

11 Seed mass and E:S ratio are fully independent of each other for the genera studied (Pearson  
12 correlation:  $r = 0.04$ ;  $P > 0.05$ ). Genera with similar E:S tend to co-occur in plots from similar  
13 habitats, indicating that phenotypic clustering occurs (Mantel test:  $r = 0.02$ ,  $P = 0.023$ ), the  
14 same is true for genera with similar seed mass (Mantel test:  $r = 0.05$ ,  $P = 0.018$ ). A  
15 significantly negative relationship ( $t = -15.85$ ;  $r = 0.49$ ;  $P < 0.001$ ) exists between mean E:S  
16 per plot and mean Ellenberg indicator values for moisture (Table 2, Fig. 1a). The mean E:S  
17 per plot increased with increasing habitat dryness and was significantly different for the three  
18 habitats studied ( $F_{2,772} = 169.62$ ;  $P < 0.001$ ). Mean indicator values for moisture and light of  
19 plants in separate plots each both explained  $> 40\%$  of variation in mean seed mass (Fig. 1b).  
20 At the habitat level, mean seed mass per plot was significantly different for the three habitats  
21 studied ( $F_{2,772} = 592.225$ ;  $P < 0.001$ ) and increased with decreasing habitat light conditions.

22

23 Both E:S and seed size had significant phylogenetic signals ( $K = 2.1$  and  $K = 0.31$   
24 respectively,  $P < 0.05$  in both cases), indicating that closely-related species resemble each  
25 other more than expected by chance. However, the amount of signal was 6.5 times greater for

1 E:S than for seed size. The mean NRI of all plots combined was  $0.07 \pm 0.03$  (mean  $\pm$  s.e.),  
2 indicating a phylogenetically clustered community structure. A significant correlation exists  
3 between the mean E:S and NRI of the plots when a quadratic term was added ( $t = 17.78$ ;  $r =$   
4  $0.54$ ;  $P < 0.001$ ). This is because phylogenetic clustering tends to increase in plots with a  
5 mean E:S nearing either 0 or 1, while no significant phylogenetic structure appeared in plots  
6 with an intermediate E:S, except in a negligible number of plots where overdispersion arose  
7 (Fig. 2a). All plots with a significantly clustered structure and a low mean E:S ( $< 0.35$ ) are  
8 located in the wet habitat (dune slacks), while those with a high mean E:S ( $> 0.56$ ) are related  
9 to the drier calcareous grassland and forest habitats. NRI significantly increased with  
10 increasing seed mass per plot ( $t = 11.29$ ;  $r = 0.38$ ;  $P < 0.001$ ; Fig. 2b). Adding a quadratic  
11 term only marginally improved the correlation.

12

13 Contrary to the expectation of the hypothesis that trait-divergence may allow the coexistence  
14 of closely related species, we found that the dispersion of E:S ( $0.31 \pm 1.00$ ; mean  $\pm$  s.d.), is  
15 positively related to phylogenetic dispersion measured as the inverse of NRI ( $t = 20.54$ ;  $r =$   
16  $0.60$ ;  $P < 0.001$ ) in the study plots. Seed mass dispersion ( $-0.09 \pm 0.69$ ), on the other hand, is  
17 negatively related to phylogenetic dispersion ( $t = -8.80$ ;  $r = 0.30$ ;  $P < 0.001$ ). When looking at  
18 the relative importance of both traits in a combined model, we found that E:S dispersion was  
19 *c.* 4 times (ratio of effect standard deviations) more important than dispersion of seed mass in  
20 explaining phylogenetic structure.

21

22 The observed  $\Pi_{ST}$  (0.047) was significantly larger than zero ( $P = 0.02$ ) after complete  
23 randomization of the phylogenetic tree, again indicating phylogenetic clustering. Partial tree  
24 randomizations at specific time intervals did not result in a  $\Pi_{ST}$  significantly different from  
25 zero. Partial randomization of specific clades showed significant phylogenetic clustering

1 within Rosales, but not in the other major clades tested (Table 3). Genera of the Rosales clade  
2 in this study had a mean E:S of  $0.88 \pm 0.08$ . Rosales species were most frequent in the  
3 calcareous grassland with on average 3.7 species per plot, as compared to 0.5 and 2.7 species  
4 per plot in the dune slack and forest habitat, respectively.

5

## 6 **Discussion**

7 E:S is an evolutionarily conserved trait directly linked to rate of seed germination and thus  
8 plant recruitment (Verdú, 2006). E:S was already reported to vary with habitat moisture  
9 conditions (Hodgson and Mackey, 1986). Here we formally showed that genera with high E:S  
10 are over-represented in dry habitats, clearly indicating that community structure is  
11 phenotypically and, consequently, phylogenetically clustered, as we discuss below. Although  
12 seed mass, widely regarded as an important functional trait (Westoby, 1998; Weiher *et al.*,  
13 1999), had a clear relationship to the phenotypic community structure, with seed mass  
14 increasing with decreased light availability, it was much less important in determining  
15 phylogenetic community structure.

16

17 A phylogenetically clustered pattern of community structure emerged from two of the metrics  
18 calculated, NRI and  $\Pi_{ST}$ . Additionally, phenotypic clustering was observed for both seed  
19 traits, indicating that co-occurring genera tend to have similar trait states. The proposition that  
20 closely related species may coexist if they diverge in niche traits would result in a negative  
21 correlation between phenotypic and phylogenetic dispersion (Prinzing *et al.*, 2008). Although  
22 we also observed this relationship for seed mass, we did find that this relationship is positive  
23 for E:S. Additionally, dispersion of E:S is much more important than seed mass dispersion,  
24 confirming that E:S based filtering dominates over past divergence of trait-states. Since E:S is

1 a strongly conserved trait, we can conclude that filtering may be an important assembly  
2 process structuring the plant community studied.

3

4 The reduction in the range of E:S values we observed, especially in the calcareous grassland  
5 plots, provides an additional indication of habitat filtering (Cornwell *et al.*, 2006; Kraft *et al.*,  
6 2008). This is because environmental filtering should limit the range of strategies found in a  
7 community, as we find here. The observation that the Rosales clade, characterized by a high  
8 E:S and most frequent in dry calcareous grasslands, contributes significantly to  
9 phylogenetically clustering agrees with the E:S based habitat filtering process. Our design  
10 does not allow us to differentiate between seedlings failing to establish due to abiotic  
11 conditions in the habitats studied, i.e. habitat filtering s.str., or due to competitive ability  
12 differences, habitat filtering s.l. (Mayfield and Levine, 2010).

13

14 A reduced range of E:S values in plots with low mean Ellenberg indicator values for moisture  
15 suggests that soil moisture availability may be one important factor that drives filtering of  
16 species in temperate climate habitats (Hodgson and Mackey, 1986; Vivrette, 1995). It has  
17 been hypothesized that a low E:S is the ancestral state, and that early angiosperms were  
18 mainly restricted to moist habitats, while the subsequent evolution towards an increased E:S  
19 enabled angiosperms to colonize new habitats (Forbis *et al.*, 2002; Verdú, 2006). Partial tree  
20 randomizations showed that the clustered pattern of community structure was not observed  
21 until the whole tree was randomized, thus the evolutionary signal in the communities with a  
22 non-random phenotypic and phylogenetic structure can be traced back towards the root of the  
23 phylogeny. Therefore, the rapid change of E:S early in angiosperm evolution appears to be  
24 still reflected in the habitats studied. The negative relationship between seed size and the  
25 Ellenberg indicator values for light to a large extent can be explained in terms of the strong

1 correlation between seed size and plant size (Moles *et al.*, 2005; Wright *et al.*, 2007).  
2 Furthermore, seedlings from large seeds, with more protein and lipid reserves, may have a  
3 competitive advantage and an increased establishment success under stressful environmental  
4 conditions (Leishman *et al.*, 2000; Soriano *et al.*, 2011). The latter mechanism is confirmed in  
5 this study, since a negative relationship between seed mass and indicator values for light and  
6 moisture exists.

7

8 Several propositions have been made about the ecological advantages of storing reserves in  
9 the embryo rather than in the endosperm or some other reserve tissue. It has been suggested  
10 that the evolution towards storage of food reserves in the embryo resulted in an increased  
11 germination rate (Stebbins, 1974). Such a positive relationship between germination rate and  
12 E:S has indeed been observed for Mediterranean climate species (Vivrette, 1995). In seeds  
13 with an underdeveloped embryo at dispersal, germination is delayed because of a period of  
14 embryo growth inside the seed prior to germination, a mechanism termed morphological  
15 dormancy (Baskin and Baskin, 1998). An additional cause for slower germination might be a  
16 reduced mobilization rate when reserves are stored outside the embryo. A reduction in the  
17 number of days required for germination (Verdú, 2006) together with a tendency towards  
18 increasing E:S throughout Angiosperm evolution (Martin, 1946; Forbis *et al.*, 2002)  
19 seemingly suggests a correlated evolution between both traits. However, the lack of a  
20 correlation between E:S and the rate of germination when phylogenetic relatedness is taken  
21 into account discards this scenario and suggests that both traits are associated by common  
22 ancestry and strong phylogenetic conservatism at profound levels of the angiosperm  
23 phylogeny (Verdú, 2006). Interestingly, at lower taxonomic levels, as those occurring within  
24 the Apiaceae (a family known to require embryo growth inside the seed prior to germination),  
25 a significant positive phylogenetic correlation between embryo length and germination rate

1 exists (Vandelook *et al.*, 2012). A higher germination rate combined with the fact that embryo  
2 growth inside the seeds only occurs when seeds are imbibed, makes that a high E:S would be  
3 particularly beneficial in dry habitats (Fenner and Thompson, 2005). In grasslands with a  
4 dense above-ground plant cover and consequently high rates of light competition (Price *et al.*,  
5 2012), such as the calcareous grasslands we examined, more rapid germination also provides  
6 a competitive advantage. Hence species with a low E:S ratio would not reach the adult stage,  
7 but will be outcompeted. Experiments dealing with functional ecology of seed traits such as  
8 seed mass and E:S ratio are particularly scarce and certainly deserve further investigation.

9  
10 A quadratic relationship between phylogenetic structure and a trait expressed as a ratio, such  
11 as E:S, is expected when the trait considered is highly phylogenetically conserved. Plots  
12 showing significant phylogenetic clustering had both extreme E:S values and low variation. In  
13 contrast, plots with intermediate mean E:S values had high trait variation and a significantly  
14 clustered pattern was never observed. As mentioned before, the selective exclusion of low E:S  
15 species in dry plots is likely an important process resulting in clustering of species with a high  
16 E:S ratio. Since no reduction in the range of E:S was observed in moist plots, phylogenetic  
17 clustering in plots composed of low E:S species may also result from secondary correlations  
18 between E:S and some other traits. These plots consist mainly of graminoid species and are all  
19 situated in the dune slack habitats. Like all graminoids, these species typically have copious  
20 endosperm, but they also evolved specific adaptations to survive regular flooding in these  
21 dune slacks (Grootjans *et al.*, 1998).

22  
23 In summary, we obtained good indications that habitat filtering acting on traits related to seed  
24 germination is one important mechanism in determining the phenotypic and phylogenetic  
25 structure of temperate climate plant communities. Genera with low E:S seeds are



1 underrepresented in dry habitats and the E:S is a strongly conserved trait resulting in a  
2 phylogenetically clustered pattern of plant community structure. Nonetheless, actual  
3 measurements of environmental conditions as well as the assessment of biotic interactions,  
4 such as dispersal limitation and competition, are certainly needed to expose other possible  
5 mechanisms affecting the structure of temperate climate habitats.

6

### 7 **Supplementary data**

8 Supplementary data are available online and contain (1) the number of genera and species in  
9 each of the families sampled, (2) the embryo to seed size ratio and seed weight of the genera  
10 sampled, and (3) a list of references used to manually resolve sub-family level taxonomic  
11 relatedness.

12

### 13 **Funding information**

14 This work was funded by the Research Foundation – Flanders (FWO) and was conducted at  
15 the Centro de Investigaciones sobre Desertificación (CIDE) in Valencia.

16

### 17 **Acknowledgements**

18 We thank Tara Forbis for providing the data on the E:S and Olivier Hardy for providing  
19 access to the Spacodi software. We are thankful to Juli Pausas for helping with the statistical  
20 analyses. H. Jacquemyn, and J. Pausas improved the manuscript by providing useful  
21 comments.

22

### 23 **References**

24 **Baskin CC, Baskin JM. 1998.** *Seeds. Ecology, biogeography, and evolution of dormancy*  
25 *and germination.* San Diego: Academic Press.

1 **Beltrán E, Valiente-Banuet A, Verdú M. 2012.** Trait divergence and indirect interactions  
2 allow facilitation of congeneric species. *Annals of Botany*, In press.

3 **Blomberg SP, Garland T Jr, Ives AR. 2003.** Testing for phylogenetic signal in comparative  
4 data: Behavioral traits are more labile. *Evolution* **57**: 717-745.

5 **Bossuyt B, Honnay O, Hermy M. 2003.** An island biogeographical view of the successional  
6 pathway in wet dune slacks. *Journal of Vegetation Science* **14**: 781-788.

7 **Butaye J, Honnay O, Adriaens D, Delescaille L-M, Hermy M. 2005.** Phytosociology and  
8 phytogeography of the calcareous grasslands on Devonian limestone in Southwest Belgium.  
9 *Belgian Journal of Botany* **138**: 24-38.

10 **Cavender-Bares J, Ackerly DD, Baum DA, Bazzaz FA. 2004.** Phylogenetic overdispersion  
11 in Floridian oak communities. *American Naturalist* **163**: 823-843.

12 **Cornwell WK, Schilke DW, Ackerly DD. 2006.** A trait-based test for habitat filtering:  
13 convex hull volume. *Ecology* **87**: 1465-1471.

14 **Dzwonko Z. 2001.** Assessment of light and soil conditions in ancient and recent woodlands  
15 by Ellenberg indicator values. *Journal of Applied Ecology* **38**: 942-951.

16 **Elias M, Gompert Z, Jiggins C, Willmott K. 2008.** Mutualistic interactions drive ecological  
17 niche convergence in a diverse butterfly community. *Plos Biology* **6**: 2642-2649.

18 **Ellenberg H, Weber HE, Düll R, Wirth V, Werner W. 1992.** Zeigerwerte von Pflanzen in  
19 Mitteleuropa. *Scripta Geobotanica* **18**: 3-258.

20 **Fenner M, Thompson K. 2005.** *The ecology of seeds*. Cambridge: Cambridge University  
21 Press.

22 **Firth D. 2008.** *relimp: Relative contribution of effects in a regression model*. R package  
23 version 1.0-1. <http://go.warwick.ac.uk/relimp>, <http://go.warwick.ac.uk/dfirth>

1 **Forbis TA, Floyd SK, De Queiroz A. 2002.** The evolution of embryo size in angiosperms  
2 and other seed plants: Implications for the evolution of seed dormancy. *Evolution* **56**: 2112-  
3 2125.

4 **Grime JP. 2006.** Trait convergence and trait divergence in herbaceous plant communities:  
5 Mechanisms and consequences. *Journal of Vegetation Science* **17**: 255-260.

6 **Grootjans AP, Ernst WHO, Stuyfzand PJ. 1998.** European dune slacks: strong interactions  
7 of biology, pedogenesis and hydrology. *Trends in Ecology and Evolution* **13**: 96-100.

8 **Hanley ME. 1998.** Seedling herbivory, community composition and plant life history traits.  
9 *Perspectives in Plant Ecology, Evolution and Systematics* **1**: 191–205.

10 **Hanley ME, Fenner M, Whibley H, Darvill B. 2004.** Early plant growth: identifying the  
11 end point of the seedling phase. *New Phytologist* **163**: 61-66.

12 **Hanley ME, Cordier PK, May O, Kelly CK. 2007.** Seed size and seedling growth:  
13 differential response of Australian and British Fabaceae to nutrient limitation. *New*  
14 *Phytologist* **174**: 381-388.

15 **Hardy OJ, Jos L. 2008.** Interpreting and estimating measures of community phylogenetic  
16 structuring. *Journal of Ecology* **96**: 849-852.

17 **Hardy OJ, Senterre B. 2007.** Characterizing the phylogenetic structure of communities by  
18 an additive partitioning of phylogenetic diversity. *Journal of Ecology* **95**: 493-506.

19 **Hodgson JG, Mackey JML. 1986.** The ecological specialization of dicotyledonous families  
20 within a local flora - Some factors constraining optimization of seed size and their possible  
21 evolutionary significance. *New Phytologist* **104**: 497-515.

22 **Hodkinson DJ, Askew AP, Thompson K, Hodgson JG, Bakker JP, Bekker RM. 1998.**  
23 Ecological correlates of seed size in the British flora. *Functional Ecology* **12**: 762-766.

24 **Kraft NJB, Ackerly DD. 2010.** Functional trait and phylogenetic tests of community  
25 assembly across spatial scales in an Amazonian forest. *Ecological Monographs* **80**: 401-422.

1 **Kraft NJB, Valencia R, Ackerly DD.** 2008. Functional traits and niche-based tree  
2 community assembly in an Amazonian forest. *Science* **322**: 580-582.

3 **Lameire S, Hermy M, Honnay O.** 2000. Two decades of change in the ground vegetation of  
4 a mixed deciduous forest in an agricultural landscape. *Journal of Vegetation Science* **11**: 695-  
5 704.

6 **Legendre P, Lapointe FJ, Casgrain P.** 1994. Modeling brain evolution from behavior - A  
7 permutational regression approach. *Evolution* **48**: 1487-1499.

8 **Leishman MR, Wright IJ, Moles AT, Westoby M.** 2000. The evolutionary ecology of seed  
9 size. In: Fenner M, ed. *Seeds – the ecology of regeneration in plant communities*.  
10 Wallingford, UK: CAB International, 31–57.

11 **Leishman MR, Westoby M.** 1994. The role of seed size in seedling establishment in dry soil  
12 conditions -- Experimental evidence from semi-arid species. *Journal of Ecology* **82**: 249-258.

13 **Lord J, Westoby M, Leishman M.** 1995. Seed size and phylogeny in six temperate floras -  
14 Constraints, niche conservatism, and adaptation. *American Naturalist* **146**: 349-364.

15 **Martin AC.** 1946. The comparative internal morphology of seeds. *American Midland*  
16 *Naturalist* **36**: 513-660.

17 **Mayfield MM, Levine JM.** 2010. Opposing effects of competitive exclusion on the  
18 phylogenetic structure of communities. *Ecology Letters* **13**: 1085-1093.

19 **Mazer SJ.** 1989. Ecological, taxonomic, and life history correlates of seed mass among  
20 Indiana cune angiosperms. *Ecological Monographs* **59**: 153–175.

21 **Moles AT, Ackerly DD, Webb CO, Tweddle JC, Dickie JB, Pitman AJ, Westoby M.**  
22 **2005.** Factors that shape seed mass evolution. *Proceedings of the National Academy of*  
23 *Sciences* **102**: 10540-10544.

24 **Ojeda F, Pausas JG, Verdú M.** 2010. Soil shapes community structure through fire.  
25 *Oecologia* **163**: 729-735.

1 **Pausas JG, Verdú M. 2010.** The jungle of methods for evaluating phenotypic and  
2 phylogenetic structure of communities. *BioScience* **60**: 614-625.

3 **Price JN, Hiiasalu I, Gerhold P, Pärtel M. 2012.** Small-scale grassland assembly patterns  
4 differ above and below the soil surface. *Ecology*, In press.

5 **Prinzing A, Reiffers R, Braakhekke WG, Hennekens SM, Tackenberg O, Ozinga WA,**  
6 **Schaminee JHJ, van Groenendael JM. 2008.** Less lineages – more trait variation:  
7 phylogenetically clustered plant communities are functionally more diverse. *Ecology Letters*  
8 **11**: 809-819.

9 **R Development Core Team. 2008.** *R: A language and environment for statistical computing.*  
10 R Foundation for Statistical Computing, Vienna, Austria. – <http://www.R-project.org>.

11 **Soriano D, Orozco-Segovia A, Márquez-Guzmán J, Kitajima K, Gamboa-de Buen A,**  
12 **Huante P. 2011.** Seed reserve composition in 19 tree species of a tropical deciduous forest in  
13 Mexico and its relationship to seed germination and seedling growth. *Annals of Botany* **107**:  
14 939-951.

15 **Stebbins GL. 1974.** *Flowering plants: Evolution above the species level.* London: Edward  
16 Arnold (Publishers) Ltd.

17 **Stevens PF. 2001.** *Angiosperm Phylogeny Website.* Version 9, June 2008 [and more or less  
18 continuously updated since]. <http://www.mobot.org/MOBOT/research/APweb/>

19 **Tautenhahn S, Heilmeier H, Götzenberger L, Kotz S, Wirth C, Kühn I. 2008.** On the  
20 biogeography of seed mass in Germany – distribution patterns and environmental correlates.  
21 *Ecography* **31**: 457–468.

22 **Thioulouse JD, Chessel, S, Dolédec S, Olivier JM. 1996.** ADE-4: a multivariate analysis and  
23 graphical display software. *Statistics and Computing* **7**: 75-83.

24 **Thomposon K, Hodkinson DJ. 1998.** Seed mass, habitat and life history: a re-analysis of  
25 Salisbury (1942, 1974). *New Phytologist* **138**: 163-167.

1 **Valiente-Banuet A., Verdú M. 2007.** Facilitation can increase the phylogenetic diversity of  
2 plant communities. *Ecology Letters* **10**: 1029-1036.

3 **Vamosi SM, Heard SB, Vamosi C, Webb CO. 2009.** Emerging patterns in the comparative  
4 analysis of phylogenetic community structure. *Molecular Ecology* **18**: 572-592.

5 **Vamosi JC, Vamosi SM. 2007.** Body size, rarity, and phylogenetic community structure:  
6 insights from diving beetle assemblages of Alberta. *Diversity and Distributions* **13**: 1–10.

7 **Vandelook F, Janssens SB, Probert RJ. 2012.** Relative embryo length as an adaptation to  
8 habitat and life cycle in Apiaceae. *New Phytologist*, In press.

9 **Vandervalk AG. 1981.** Succession in wetlands - A Gleasonian approach. *Ecology* **62**: 688-  
10 696.

11 **Verdú M. 2006.** Tempo, mode and phylogenetic associations of relative embryo size  
12 evolution in angiosperms. *Journal of Evolutionary Biology* **19**: 625-634.

13 **Verdú M., Gómez-Aparicio L., Valiente-Banuet A. (2012).** Phylogenetic relatedness as a  
14 tool in restoration ecology: A meta-analysis. *Proceedings of the Royal Society of London B*  
15 **279**: 1761-1767.

16 **Verdú M, Pausas JG. 2007.** Fire drives phylogenetic clustering in Mediterranean Basin  
17 woody plant communities. *Journal of Ecology* **95**: 1316-1323.

18 **Vivrette NJ. 1995.** Distribution and ecological significance of seed-embryo types in  
19 Mediterranean climates in California, Chile, and Australia. In: Arroyo MKT, Zedler PH, Fox  
20 MD, eds. *Ecology and biogeography of Mediterranean ecosystems in Chile, California and*  
21 *Australia*. New York: Springer Verlag, 274-288.

22 **Webb CO, Ackerly DD, Kembell SW. 2007.** *Phylocom: software for the analysis of*  
23 *community phylogenetic structure and character evolution*. Version 4.1 URL:  
24 <http://www.phylodiversity.net/phylocom/>

1 **Webb CO, Ackerly DD, Mcpeck MA, Donoghue MJ. 2002.** Phylogenies and community  
2 ecology. *Annual Review of Ecology and Systematics* **33**: 475-505.

3 **Weihner E, van der Werf A, Thompson K, Roderick M, Garnier E, Eriksson O. 1999.**  
4 Challenging Theophrastus: A common core list of plant traits for functional ecology. *Journal*  
5 *of Vegetation Science* **10**: 609-620.

6 **Westoby M. 1998.** A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*  
7 **199**: 213-227.

8 **Wikström N, Savolainen V, Chase MW. 2001.** Evolution of the angiosperms: calibrating  
9 the family tree. *Proceedings of the Royal Society B* **268**: 2211-2220.

10 **Wright IJ, Ackerly DD, Bongers F, Harms KE, Ibarra-Manriquez G, Martinez-Ramos**  
11 **M, Mazer SJ, Muller-Landau HC, Paz H, Pitman NCA, Poorter L, Silman MR,**  
12 **Vriesendorp CF, Webb CO, Westoby M, Wright SJ. 2007.** Relationships among  
13 ecologically important dimensions of plant trait variation in seven neotropical forests. *Annals*  
14 *of Botany* **99**: 1003-1015.

15 **Zanne AE, Chapman CA, Kitajima K. 2005.** Evolutionary and ecological correlates of  
16 early seedling morphology in East African trees and shrubs. *American Journal of Botany* **92**:  
17 972-978.

18  
19  
20  
21  
22  
23  
24  
25

1 **Table 1** Total number of plots sampled in the three habitat types differing profoundly in  
 2 moisture and light conditions. The amount of genera sampled and mean number of genera ( $\pm$   
 3 s.d.) per plot in each habitat. Sampling data were retrieved from the references listed.

	Plots	Genera	Genera per plot	Reference
Dune slacks	484	85	$7.5 \pm 3.1$	Bossuyt <i>et al.</i> (2003)
Calcareous grassland	204	167	$25.5 \pm 6.6$	Butaye <i>et al.</i> (2005)
Forest	85	92	$15.2 \pm 8.7$	
Vorte Bossen	28	67		Lameire <i>et al.</i> (2000)
Langerodebos	57	79		O. Honnay (unpublished data)

4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23



1 **Table 2** Results of linear regression across all plots sampled in three temperate climate  
 2 habitats. The mean log seed mass and embryo to seed size ratio (E:S) in each plot are  
 3 dependent variables, and the mean Ellenberg indicator values for light, soil moisture and  
 4 nitrogen calculated for separate plots are predictor variables.

	Log Seed mass (mg)			E:S		
	t	r	p-value	t	r	p-value
Light	-23.60	0.65	0.001	-8.78	0.30	0.001
Moisture	-24.61	0.66	0.001	-15.85	0.49	0.001
Nitrogen	4.25	0.15	0.001	0.68	0.03	0.495

5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23

1 **Table 3** The phylogenetic signal ( $p\Pi_{ST}$ ) of the community structure in temperate climate  
 2 habitats calculated for separate Angiosperm clades. Results of  $p\Pi_{ST}$  were obtained after  
 3 partial tree randomization of specific clades in comparison to  $\Pi_{ST}$ . Means  $\pm$  95 % of the  
 4 central values obtained after 999 randomizations of the specific clade.

	$p\Pi_{ST}$	p-value
Poales	$0.050 \pm 0.004$	0.59
Asterales	$0.047 \pm 0.001$	0.48
Rosales	$0.046 \pm 0.001$	0.01
Asparagales	$0.046 \pm 0.001$	0.09
Eudicots	$0.046 \pm 0.003$	0.59
Whole phylogeny	$0.001 \pm 0.012$	0.01

5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

1 **Fig. 1** (a) The mean embryo to seed size ratio (E:S) per plot is negatively related to mean  
2 Ellenberg indicator value for moisture per plot (open symbols), for plots sampled in dune  
3 slacks (circles), calcareous grasslands (triangles) and temperate forests (squares). (b) The  
4 mean seed mass per plot is negatively related to mean Ellenberg indicator value for light.  
5 Closed symbols denote mean values of the plots per habitat. Vertical and horizontal error bars  
6 represent standard deviations for mean Ellenberg indicator value and mean trait values (closed  
7 symbols), respectively.

8

9 **Fig. 2** The relationship between (a) mean embryo to seed size (E:S) and (b) mean seed mass  
10 per plot and the net relatedness index (NRI) for each plot reveals that the phylogenetic  
11 structure of temperate climate habitats is related to the phenotypic structure. Plots were  
12 sampled in dune slacks (circles), calcareous grasslands (triangles) and forests (squares).



