

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Mönkkönen, Mikko; Devictor, Vincent; Forsman, Jukka T.; Lehtikoinen, Aleks; Elo, Merja

Title: Linking species interactions with phylogenetic and functional distance in European bird assemblages at broad spatial scales

Year: 2017

Version:

Please cite the original version:

Mönkkönen, M., Devictor, V., Forsman, J. T., Lehtikoinen, A., & Elo, M. (2017). Linking species interactions with phylogenetic and functional distance in European bird assemblages at broad spatial scales. *Global Ecology and Biogeography*, 26 (8), 952-962. doi:10.1111/geb.12605

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

1 Linking species interactions with phylogenetic and functional distance in European bird assemblages
2 at broad spatial scales

3 Mikko Mönkkönen, Vincent Devictor, Jukka T Forsman, Alekski Lehtikoinen & Merja Elo

4

5 Mikko Mönkkönen, University of Jyväskylä, Department of Biological and Environmental Sciences,
6 POB 35, FI-40014 University of Jyväskylä, Finland. e-mail: mikko.monkkonen@jyu.fi

7 Vincent Devictor, Institut des Sciences de l'Évolution de Montpellier, Place Eugène Bataillon, 34095
8 Montpellier Cedex 05, France. e-mail: vincent.devictor@univ-montp2.fr

9 Jukka T. Forsman, Department of Ecology and Genetics, POB 3000, FI-90014 University of Oulu,
10 Finland. e-mail: jukka.forsman@oulu.fi

11 Alekski Lehtikoinen, The Helsinki Lab of Ornithology, Finnish Museum of Natural History, POB 17, FI-
12 00014 University of Helsinki, Finland. e-mail: aleksi.lehtikoinen@helsinki.fi

13 Merja Elo, University of Jyväskylä, Department of Biological and Environmental Sciences, POB 35, FI-
14 40014 University of Jyväskylä, Finland. email: merja.t.elo@jyu.fi

15 Keywords: aggregated distribution, functional similarity, heterospecific attraction, interspecific
16 competition, pairwise interactions in communities, phylogenetic distance, segregated
17 distribution

18 Running title: Diverse pairwise interactions

19 Corresponding author: Mikko Mönkkönen

20 Number of words in the abstract: 279

21 Number of words in the main text: 5,103

22 Number of references: 62

23 **Abstract**

24 **Aim** Understanding the relative contribution of different species interactions in shaping community
25 assembly has been a pivotal aim in community ecology. Biotic interactions are acknowledged to be
26 important at local scales although their signal is assumed to weaken over longer distances. We
27 examine the relationship between positive, neutral, and negative pairwise bird abundance
28 distributions and the phylogenetic and functional distance between these pairs after first controlling
29 for habitat associations.

30 **Location** France and Finland.

31 **Methods** We used results from French and Finnish land bird monitoring programs from which we
32 created three independent data sets (French forests, French farmlands, Finnish forests). Separately
33 for the three data, we fitted linear mixed-effects models for pairwise abundance values across years
34 per point count station to infer the association between all common species pairs while controlling
35 for geographic distribution and habitat associations, and saved pairwise regression coefficients for
36 further analyses. We used a null model approach to infer whether the observed associations (effect
37 sizes) differ from random. Finally, using quantile regression we analyzed the relationships between
38 functional dissimilarity/phylogenetic distance and effect sizes.

39 **Results** Our results show both negative and positive species interactions although negative
40 interactions were twice as common as positive interactions. Closely-related species were more likely
41 to show strong associations, both negative and positive, than more distant species across broad
42 spatial scales. For functional dissimilarity the results varied across data sets.

43 **Main conclusions** Our results emphasize the potential of functional and phylogenetic proximity in
44 generating both negative and positive species associations, which can produce pervasive patterns
45 from local to geographical scales. Future assembly studies should refrain from strict dichotomies
46 such as compensatory dynamics versus environmental forcing and instead consider the possibility of
47 positive interactions.

48 **Introduction**

49 A fundamental and long-standing goal in community ecology has been to understand the complexity
50 of dependencies among species, and thereby the mechanisms by which communities are assembled.
51 Since the 1950's (MacArthur, 1958; Hutchinson, 1959) interspecific competition and niche
52 partitioning have been the focus of explanations for species coexistence and community assembly
53 patterns (Diamond, 1975; Connor & Simberloff, 1983; Cornell, 1985; Cornell & Lawton, 1992).
54 Experimental evidence also shows that competition is undeniably an important factor for community
55 assembly (Connell, 1983; Schoener, 1983; Goldberg *et al.*, 1992; Gurevitch *et al.*, 2000). However, it
56 is not the only one. A growing body of literature demonstrates the importance of facilitative or
57 positive non-trophic interactions (Bertness & Callaway, 1994; Cardinale *et al.*, 2002; Bruno *et al.*,
58 2003). Moreover, interactions often result from combinations of positive, negative and neutral
59 relationships (Seppänen *et al.*, 2007) potentially generating asymmetric interactions (Mönkkönen *et al.*,
60 1999). Indeed, Gross (2008) concluded that the joint effects of different interactions may be the
61 most important factor for community assembly, the key-question being the relative contribution of
62 each interaction.

63 Species interactions may affect species abundances in communities leading to patterns in space,
64 which are independent of habitat characteristics. Negative interactions, such as competition, are
65 predicted to cause segregated distributions (Gotelli *et al.*, 2010). Positive interactions, predicted to
66 result in aggregated distributions, are well known among plants and sessile animals (Bertness &
67 Callaway, 1994; Bruno *et al.*, 2003), but in mobile animals, the prevalence and mechanisms of
68 positive interactions are poorly known. Recent findings about information use in animals imply a
69 likely mechanism. While assessing the best site or resources for reproduction, individuals may use
70 the presence, behavior or success of, not only conspecifics (Danchin *et al.*, 2004), but also competing
71 heterospecific individuals (Seppänen *et al.*, 2007) as cues to decide where to settle. Heterospecific
72 information use and attraction to heterospecifics and potential competitors has been demonstrated

73 from ants to apes (Seppänen *et al.*, 2007), including in breeding site choices of birds (Mönkkönen *et*
74 *al.*, 1990; Thomson *et al.*, 2003; Seppänen & Forsman, 2007; Loukola *et al.*, 2013; Kivelä *et al.*, 2014).

75 Another source producing positive species associations is the benefits of certain keystone species,
76 such as beavers and woodpeckers, which free or create new resources for other species and result in
77 positive species spatial associations (Heikkinen *et al.*, 2007; Belmaker *et al.*, 2015).

78 Although biotic interactions, and particularly competition, have a long history in the study of
79 community patterns their importance is often neglected at broad spatial scales where speciation,
80 extinction and geographic dispersal are expected to be the main driving processes (Gaston &
81 Blackburn, 2000). Recently, Araújo and Rozenfeld (2014) modelled spatial consequences of all types
82 of species interactions (continuum from positive to negative) and concluded that, in particular,
83 positive species interactions can be manifested from local to larger scales. Indeed, a few empirical
84 studies have revealed both aggregated and segregated distribution patterns among ecologically
85 similar species, likely resulting from both competitive and positive interactions, independent of
86 habitat characteristics, at broad spatial scales (Gotelli *et al.*, 2010; Ricklefs, 2012).

87 The strength of species associations may depend on functional and phylogenetic similarity of species
88 involved. Both heterospecific information use and the limiting similarity principle predict that
89 species associations are most intense, but in opposite directions, among functionally similar species
90 (Fig. 1). Heterospecific information use predicts positive associations between species that use
91 similar resources, i.e. are functionally similar, because the information value decreases with
92 increasing ecological distance (Seppänen *et al.*, 2007), while limiting similarity predicts mutual
93 avoidance between similar species due to costs of competition (MacArthur & Levins, 1967).

94 However, it is likely that both competition and heterospecific information use are context
95 dependent and the net outcome of an association between similar species depends on the costs of
96 competition and benefits of information use (Mönkkönen *et al.*, 1999; Seppänen *et al.*, 2007).

97 Therefore, associations between functionally similar species may be strong, either positive or
98 negative and the strength of the association may decrease with decreased species similarity. The

99 likelihood of finding such a relationship depends, however, on the traits considered (Trisos *et al.*,
100 2014). Using phylogenetic information can be a relevant approach to avoid a priori trait selection.
101 Similar patterns are indeed also likely for phylogenetic distance providing that trait conservatism
102 renders closely related species also ecologically more similar (Mouquet *et al.*, 2012). In this case,
103 phylogenetically closely related species should also show both the strongest positive and negative
104 associations, and phylogenetically distant species should show neutral associations. So far, however,
105 empirical tests of this hypothesis have yielded mixed results (Violle *et al.*, 2011; Godoy *et al.*, 2014).

106 Earlier research on general mechanisms of community assembly commonly separated the effects of
107 competitive interactions from environmental forcing causing positive covariation among species
108 abundances and provided support for environmental variation rather than competition driving the
109 variation in species abundances (Houlahan *et al.*, 2007; Mutshinda *et al.*, 2009; Ricklefs, 2012).

110 However, some of the positive covariation may be due to positive interactions between species,
111 over and above habitat filtering and productivity. Moreover, even if environmental forcing generally
112 prevails over competitive interactions, negative interactions may not be trivial. Thus both negative
113 and positive interactions can leave a signature on community assembly that affects both historical
114 and ecological distribution patterns. In this article, we study (i) whether there are positive or
115 negative associations among bird species in local communities, independently of habitat
116 characteristics, and consequent community patterns at large geographical scales, and (ii) whether
117 these positive or negative associations are related to functional similarity and/or phylogenetic
118 distance of the species. We predict that functional dissimilarity/phylogenetic distance and the
119 strength of the associations form a 'funnel plot' where the strongest associations, either positive or
120 negative, are between functionally similar or phylogenetically close species whereas the associations
121 grow weaker with increasing dissimilarity/distance (Fig. 1). We use comprehensive bird census data
122 from French forests and farmlands as well as Finnish forests, and analyze the three datasets
123 separately to test for consistency of results across habitat types and geographic areas. Using these
124 three datasets separately offers a possibility to test the same question on the same group of species

125 originating from quite different kinds of landscapes and climatic conditions. Consistent patterns
126 would imply generality and call for further scrutiny of underlying mechanisms.

127

128 **Material and methods**

129 *Data*

130 Data were extracted from two independent datasets from France and Finland. They are very well
131 suited for the study because together they cover a wide extent (with a relatively small grain size) and
132 consequently also a wide spectrum of climatic and environmental conditions, and they are both of
133 high quality gathered over multiple years and multiple sites. The French Breeding Bird Survey is a
134 large-scale, multi-year and multi-plot monitoring program of the French avifauna. The program
135 followed a standardized protocol from 2001 to 2012 (Jiguet *et al.*, 2012) where 2x2 km sampling
136 plots are randomly selected within 10-km radius areas, which ensures a representative sampling of
137 existing habitats. In each plot, 10 point count stations were evenly distributed. At each station, the
138 observer recorded all birds heard and seen during 5 minutes, in two sessions during the breeding
139 season. For each species in each point count station and each year, the maximum number of
140 individuals recorded during the two sessions is retained as a proxy for the local abundance of that
141 species in that plot and year.

142 In Finland, point counts have been conducted as part of the national common bird monitoring
143 scheme between 1984 and 2011 (Laaksonen & Lehtikoinen, 2013). Each census route included 20
144 point count stations located in the habitat that is uniform within 50-m radius of the station. The
145 habitat of each point was classified into 17 different habitat categories. Stations within a route were
146 at least 250 m apart in forested habitats and 350 m apart in open habitats to avoid pseudo-
147 replication. Due to very short and synchronized breeding season of boreal birds, a route was
148 censused only once per season. At each station, an observer counted, for 5 mins, all the observed
149 land birds during late spring – early summer (May 20 – June 20 in south-central Finland, May 30 -

150 June 30 in northern Finland). In both monitoring projects, surveys were conducted early in the
151 morning (typically between sunrise and 10 am), which is when the birds were most active, and only
152 on days with good weather conditions (no rain or heavy wind).

153

154 *Data handling*

155 Original French data included plots sampled for at least two years, i.e. 1,914 plots and 19,140 point
156 count stations. Finnish data included 286 routes and 5,760 stations. From these data we first created
157 three independent data sets by selecting the point count stations situated in forests or in farmlands
158 only (French farmlands, French forests, Finnish forests; Finnish farmland data were too small for the
159 analysis). The aim was to remove most of the variation in bird abundances resulting from habitat
160 structures (habitat filtering). Indeed, one of the obvious sources of species segregation or
161 aggregation is the main habitat type in which a given plot is monitored. A negative association
162 between farmland birds and forest birds would be interpreted as a signal of competition although
163 those two groups simply do not co-occur. The three datasets contain about 2,900–9,000 point count
164 stations but a rather narrow range of habitats (Table 1). We analyzed the three datasets separately
165 to test for the consistency of results across habitat types and geographic areas.

166 We excluded all waterbirds and birds of prey since point count census at the local point count
167 station level provides reliable information on species abundances only for land birds with relatively
168 small home ranges. We also excluded very rare species that were present in less than 2% of the
169 point count stations. After filtering, our data included 43,000-206,000 observations for 76-83 species
170 (Table 1), for which we calculated average abundance values across years per point count station.
171 The same species occurred in multiple data sets: the number of shared species was 67 in French
172 forest and farmland data, 51 in French and Finnish forest data, and 45 in French farmland and
173 Finnish forest data.

174 The matrix of pairwise functional distances was produced from 22 functional traits (Appendix S1 in
175 Supporting Information) using methods described in Devictor *et al.* (2010). These traits encompassed
176 life-history traits and feeding habits (Petchey *et al.*, 2007) and were identified as being important in
177 determining the response of bird species to environmental change and in determining the
178 contribution of bird species to ecosystem functions (Sekercioglu, 2006). From these traits, we
179 calculated the Gower distance to represent pairwise trait distances estimated from the species trait
180 matrix (Legendre & Legendre, 1998). As we aim to remove habitat-induced distances between
181 species we calculated a separate matrix for each set of species (French farmlands, French forests,
182 Finnish forests). Gower distance accounts for both continuous and qualitative traits (Mouchet *et al.*,
183 2008) and were measured with the function 'daisy' of the R package 'cluster' (Maechler *et al.*, 2016).
184 All pairwise distances were standardized by dividing original distance values with the range of
185 values.

186 We extracted pair-wise phylogenetic distances directly from a dated molecular phylogenetic tree
187 assembled by Thuiller *et al.* (2011) and then used ultrametric distances from this tree representing
188 relative phylogenetic distances among species using the function 'cl_ultrametric' in the R package
189 'clue' (Hornik, 2005). Phylogenetic information was not available for 7 species in Finnish forest
190 dataset, and thus we performed analyses of phylogenetic distance on 69 species.

191 *Statistical analyses*

192 We adopt the pairwise approach to analyze species effects on each other's abundance, i.e., consider
193 a species pair as the fundamental unit in interactions. The procedure has advantages over the matrix
194 method i.e. where a target of interest is the whole community i.e. presence-absence matrix that has
195 been used since the beginning of the studies of co-occurrence patterns (Diamond, 1975; Connor *et al.*,
196 2013). Most communities contain many potential species pairs each of which may exhibit
197 positive, negative or random associations. Therefore, single metrics that summarizes an entire
198 assemblage can mask the type and strength of pairwise interactions and it is therefore instructive to

199 analyze individual pairs of species (Boulangéat *et al.*, 2012; Veech, 2013; Blois *et al.*, 2014). We
200 relate species abundances against each other to reveal signals of positive, neutral and negative
201 associations, after controlling for variation in bird abundances due to geographic distribution and
202 corollary climatic variation as well as finer habitat associations, beyond the main habitat type (i.e.
203 forest and farmland). We fitted a linear mixed-effects model for each species pair to infer the
204 association between species_i and species_j while controlling for geographic distribution and finer
205 habitat associations within the main habitat type (forest or farmland; see Appendix S2 for a detailed
206 flowchart of the analyses). First, the biogeographic zone was entered into the model as a random
207 factor to control for large scale variation in climatic conditions. In each dataset, we assigned the
208 point count stations to biogeographical zones according to information from European
209 Environmental Agency (2015) for France and Järvinen & Väisänen (1980) for Finland (see Table 1 for
210 zones). Second, we included site (a dummy variable for a point count plot or route) nested within
211 the biogeographic zone, to control for small scale variation in environmental conditions among plots
212 or routes. We then added a second-order trend surface ($X + Y + X^2 + Y^2 + XY$; X = longitude and Y =
213 latitude) to control for the geographic (e.g. temperature and precipitation related) variation in
214 species abundances. Although a minimum of third order trend surface is generally recommended
215 (Legendre & Legendre, 1998) random effects (zone and site) already control for spatial
216 autocorrelation at the scale of biogeographic zone and at very small spatial scales. Moreover, the
217 models with a third-order trend surface tended to be over-parameterized leading to model
218 convergence problems. Including second-order trend surfaces ensures that our pairwise regression
219 analysis operates on local scale variation in species abundances. Fourth, we entered habitat type of
220 the point count station (Table 1) to further control for species habitat preferences. Because we
221 analyzed forest and farmland data separately (i.e. narrowed down the variation in habitat structures
222 prior to analysis) and entered farmland type (ploughed meadow, unploughed meadow, mixed
223 farmland, open field, permanent crop) or forest type (e.g. deciduous, coniferous or mixed forests in
224 France; spuce, pine, deciduous or mixed forests in Finland) as a random factor, our analysis

225 effectively controls for fine scale habitat filtering. Fifth, we entered community size (summed
226 abundances of all species, excluding species_i and species_j) as a fixed effect controlling for the
227 possibility that species abundances may covary with total community size because of independent
228 responses to productivity. Finally, we added the abundance of species_j as a fixed effect. We log-
229 transformed ($\log(n+1)$) all abundances prior to analyses and saved pairwise regression coefficients
230 for further analyses. Estimating pairwise associations using Poisson distribution yielded too many
231 convergence problems and was not feasible due to a very high number of models. Note that using
232 Poisson distribution should not change the general conclusions derived from our framework based
233 on log-transformed abundances (Ives, 2015).

234 To infer whether the observed associations differ from what could be observed on the basis of
235 randomly distributed individuals we used a null model approach. First, we defined a regional species
236 pool as all observed species and their abundances, separately for each habitat in each
237 biogeographical zone. Next, we randomly sampled the observed number of individuals from the
238 regional species pool while preserving the abundance of each species and observed total abundance
239 in each point count station. In other words, we kept the size of the regional habitat-specific species
240 pools and the size of the local communities fixed, and within these constraints we randomized the
241 composition of local communities at point count stations (see Crist *et al.*, 2003, for a similar
242 approach). This type of randomization makes a plausible assumption that species habitat
243 associations, their relative abundances in regional pools and local community size are real ecological
244 properties worth retaining while relaxing deterministic pairwise associations. Then we fitted the
245 linear mixed-effect model described above for the randomized datasets. We repeated this
246 procedure 1000 times and calculated the standardized effect size as the difference between
247 observed pairwise regression coefficient and mean expected coefficient, divided by the standard
248 deviation. Thus, the effect size measures the direction (positive or negative) and strength of species
249 interactions in the datasets, independent from what one expects by chance and sampling artefacts.

250 Because the effect sizes within species pairs (effect of species_i on species_j, and *vice versa*) were
251 strongly correlated (Spearman's rho = 0.943, $n = 3,403$, $P < 0.001$; Spearman's rho = 0.956, $n = 2,926$,
252 $P < 0.001$; Spearman's rho = 0.960, $n = 2,850$, $P < 0.001$; in French farmlands, French forest, and
253 Finnish forests, respectively) we used the mean of the effect sizes of each species pair (ES_{mean}) as an
254 observation unit in the analysis. Given that these observations are not independent (every species is
255 represented in multiple species pairs) we used bootstrap method to calculate standard errors
256 (Koenker, 2013).

257 We considered pairwise species effect sizes "strong" when absolute effect size values were >2 (i.e.
258 observed association deviated more than two standard deviations from the expected) and "weak"
259 when effect size values were <2 . We predicted a relationship between the magnitude of the effect
260 sizes, both positive and negative, and functional dissimilarity/phylogenetic distance. The magnitude
261 should decrease with increasing dissimilarity/distance. To test this, we used quantile regression
262 (Cade & Noon, 2003) for all quantiles from 1 to 99% quantiles (τ ranges from 0.01 to 0.99) at
263 intervals of 1%. We predicted that the regression between effect size and dissimilarity/distance
264 would result in a negative coefficient in the upper ($>50\%$) quantiles (reflecting smaller positive effect
265 sizes with increasing dissimilarity/distance) and a positive coefficient in the lower ($<50\%$) quantiles
266 (reflecting smaller negative effect sizes with increasing dissimilarity/distance). If these predictions
267 are verified, the shape of the relationship should be a 'funnel plot' with higher numbers of positive
268 and negative associations for lower values of functional dissimilarity or phylogenetic distance (Fig.
269 1). To infer whether this is truly the case we plotted the coefficient of each of the quantile regression
270 as a function of the quantile (τ) in question, and expect to see a negative relationship. As we do not
271 have a specific hypothesis about the overall relationship in the data we do not concentrate on the
272 general tendency, i.e. 50% quantile, but on upper vs lower quantiles. We performed linear mixed-
273 effect models with package 'lme4' (Bates *et al.*, 2014) and quantile regression with 'quantreg'
274 (Koenker, 2013) in R Version 3.0.3 (R Development Core Team, 2014), and the iterations for the null
275 model approach with Taito supercluster provided by CSC - IT Center for Science Ltd

276 (<https://research.csc.fi/research-home>). The R script for calculating the effect sizes is provided as
277 Appendix S3. Other analyses we performed with IBM SPSS Statistics 22.0 (IBM, Armonk, New York,
278 USA).

279

280 **Results**

281 We observed both positive and negative regression coefficients among species pairs tested in all
282 three datasets. The means of the observed coefficients were positive (0.02) and the distribution of
283 coefficients were highly similar in the three datasets (see Appendix S4). Likewise, the effect sizes
284 (ES_{mean}) were centered close to zero. A large proportion of pairwise species effect sizes can be
285 considered strong as absolute effect size values larger than two (i.e. observed association deviated
286 more than two standard deviations from the expected) constituted 83, 63 and 41% of all pairwise
287 effect sizes in French farmland ($n=6,806$), French forest ($n=5,852$) and Finnish forest dataset
288 ($n=5,700$), respectively. In all three datasets, strong negative associations were approximately twice
289 as common as strong positive associations: 57 vs 26% in French farmland, 39 vs 23% in French forest,
290 and 27 vs 15% in Finnish forest data. Thus, birds are not distributed randomly with respect to each
291 other in local communities, when controlling for habitat filtering and productivity, and we found
292 asymmetry between positive and negative associations with prevalence of the latter.

293 The relationship between functional dissimilarity and the mean of the effect sizes for species pairs
294 (ES_{mean}) showed differences among the three datasets. In French farmlands there was no
295 relationship in the lower quantiles but a negative relationship in the upper quantiles whereas in
296 French forests there was only a trend in both lower and upper quantiles (Fig. 2a,b). By contrast, in
297 Finnish forests there was a positive relationship in the lower quantiles and a negative relationship in
298 the upper quantiles (Fig. 2c). Thus, in French farmlands the level of aggregation increased with
299 decreasing functional similarity, in Finnish forests both the level of segregation and aggregations
300 increased with decreasing functional similarity, while in French forests there were no relationships.

301 Phylogenetic distance and ES_{mean} showed significant relationships in both ends of the quantile
302 spectrum in French farmlands and forests (Fig. 3a,b). Thus, both aggregated and segregated
303 distributions of species abundances increased as a function of decreasing phylogenetic distance. In
304 Finnish forests, there was a positive relationship between phylogenetic distance and ES_{mean} in the
305 lower quantiles but no relationship in the upper quantiles (Fig. 3c).

306

307 **Discussion**

308 Bird species showed both spatial aggregation and segregation in their abundances, independently of
309 habitat structures. Even though overall mean effect sizes of species associations in the local
310 communities were centered on zero, we found the majority of species were relatively strongly
311 associated with each other. This highlights the importance of both negative and positive biotic
312 interactions in affecting community assembly. Our study using natural communities across a wide
313 geographic area yielded results that are consistent with the results of the local manipulative
314 experiments: negative associations are common but also positive associations frequently occur
315 (Bertness & Callaway, 1994; Gurevitch *et al.*, 2000; Forsman *et al.*, 2002; Martorell & Freckleton,
316 2014; but see Gotelli & Ulrich, 2010). Moreover, our results suggest that high functional and
317 phylogenetic similarities can be important determinants increasing the probability of both negative
318 and positive associations. Also, we found that in bird communities strong negative associations were
319 twice as common as strong positive associations. This is in line with empirical results of Gotelli *et al.*
320 (2010) who showed strong predominance of spatially segregated over aggregated distributions
321 within foraging and congeneric guilds in Danish avifauna.

322 Species abundance is affected by a multitude of factors, such as habitat, productivity and
323 geographical position. Our pairwise approach controlled for the effects of habitat, first, by restricting
324 the analysis only to a limited set of habitat classes (Table 1) and, second, by entering habitat class as
325 a factor in the model. Moreover, we entered biogeographic zone, sampling site, and second order

326 trend surface in abundances to control for the effects of geographic factors such as species
327 geographic distributions and climate related variation. Finally, we controlled for the community size,
328 i.e. species independent responses to productivity. Yet our analysis revealed strong signals of
329 positive abundance associations. Thus, species abundances were probably affected also by positive
330 biotic interactions between species, and these interactions were strong enough to show up as
331 aggregated distributions. The recurrent finding that species abundances in natural communities tend
332 to covary positively rather than negatively has commonly been attributed to environmental forcing
333 (Houlihan *et al.*, 2007; Mutshinda *et al.*, 2009; Ricklefs, 2012). Our results suggest that a proportion
334 of this positive covariation is likely due to positive interactions. Therefore, dominance of positive
335 covariation over compensatory dynamics in community dynamics does not necessarily imply low
336 frequency of species interactions. Future studies testing assembly theories should therefore refrain
337 from using strictly dichotomist approaches such as compensatory dynamics versus environmental
338 forcing, but also consider the possibility of positive interactions and underlying processes such as
339 conspecific or heterospecific information use.

340 In all three datasets, closer species pairs with respect to phylogenetic distances are those showing
341 stronger segregation in their abundances. Neutral associations, however, were found throughout
342 the phylogenetic distance spectrum. Thus, a part of phylogenetically close bird species indeed seem
343 to compete more strongly than distantly related species which plausibly leads to avoidance of such
344 species in habitat selection. In Finnish forest bird assemblages, but not in the two French datasets,
345 the signal of spatial segregation increased also with increasing functional similarity. In concordance
346 with the prediction of heterospecific information use (Seppänen *et al.*, 2007), the signal of spatial
347 aggregations was stronger for pairs consisting of similar species and grew weaker with increasing
348 dissimilarity. However, whether the pattern was detected for phylogenetic distance, functional
349 dissimilarity or both showed variation among datasets: in French forests aggregation was related to
350 phylogenetic distance, in Finnish forests to functional dissimilarity, and in French farmlands to both.

351 In summary, segregated abundances were related to phylogenetic distance in all of the three cases,
352 to functional dissimilarity in only one, whereas aggregated abundances were related both to
353 phylogenetic distance and functional dissimilarity in two cases. Thus, the phylogenetic signal in
354 segregated abundances was stronger than that of functional (ecological) similarity. This suggests that
355 the traits we used for functional dissimilarity might not be those that actually affect species
356 competitive environment but more relevant in terms of heterospecific information use, whereas
357 phylogenetic signal encompass traits of direct relevance in both respects. This result also raises the
358 issue of trait selection in trait-based analysis. It is likely that a part of the results is dependent on the
359 particular combination of trait used. A finer examination of pairwise associations i.e. whether
360 segregated versus aggregated associations are influenced by a specific combination of traits,
361 whether there is any specific trait enhancing coexistence or whether aggregations are
362 phylogenetically clustered, would be an interesting extension of our approach.

363 Another reason why species pairs consisting of phylogenetically closely related species showed
364 strong segregation in their abundances could result from allopatric speciation. Phylogenetic
365 overdispersion, which is often attributed to negative biotic interactions, may instead be consistent
366 with a neutral model of allopatric speciation (Pigot & Etienne, 2015). If allospecies rarely co-occurred
367 in the bird assemblages we studied, and differ somewhat in their functional traits, one could see a
368 pattern where phylogenetic distance is more directly driving negative abundance associations than
369 functional similarity. Allopatric speciation does not, however, provide an explanation for
370 phylogenetic and functional patterning of positive abundance associations. It is possible, and even
371 probable, that for some species pairs the costs of competition and benefits of information use adds
372 up to show as a neutral association. Hence, it is important to bear in mind that small effect sizes in
373 our study do not necessarily indicate no or weak interactions.

374 Our approach where we simultaneously addressed both negative and positive interactions, and
375 provided support for both, may help to understand why earlier work has found mixed results
376 concerning the role of phylogenetic and ecological distance in species interactions (Violle *et al.*,

2011; Godoy *et al.*, 2014). Our results support a common expectation in the community assembly literature that due to competition, close relatives and functionally similar species should show segregated patterns in abundance (Webb *et al.*, 2002). But more interestingly, our results suggest that close relatives and functionally similar species also may also aggregate, which is reflected as positive associations among species' abundances. The likely mechanism is the use of social information (Seppänen *et al.*, 2007) and/or facilitative interactions among species (Bruno *et al.*, 2003). It is evident that both clustering and overdispersion of co-occurring species within communities may occur simultaneously. Our results challenge the implicit assumption pertinent to community phylogenetics that assembly through positive associations decreases with increasing assembly by competition (see Gerhold *et al.*, 2015). Thus, the relative balance of positive and negative interactions in community assembly cannot be quantified by a single parameter of phylogenetic (or functional) dispersion.

We acknowledge the fact that only experimental set-ups can truly prove the strength and the sign of the interaction between a species pair. It is clear that the data used here do not capture all the small scale habitat characteristics which may affect species aggregation and segregation patterns. Indeed, a detailed data of environmental conditions might attenuate the coefficients we found. On the other hand, it has been shown that even when modeling forest bird species distributions with a very detailed forest structure data the density of a bird species remains a significant predictor of the density of a close relative (Kosicki *et al.*, 2015).

Social information use and subsequent aggregated distribution in local communities result in variation in local species diversity at a given site that deviates from diversity predicted by environmental factors only, creating both hot and cold spots of species diversity in the landscape (Seppänen *et al.*, 2007). Interspecific competition resulting in segregated distribution may also create similar deviations from predictions. Our results suggest that such diversity anomalies should carry phylogenetic/functional signal. Species' interactions may render a proportion of suitable habitat patches unoccupied by the species also because dispersal among patches in the landscape

403 may be affected by the presence of close relatives and/or functionally similar species. Consequently,
404 colonization and extinction in fragmented landscapes is no longer a sole function of landscape
405 patterns and species dispersal abilities but hinges also on other species' ability to persist in
406 fragmented landscapes. Given that interspecific social information use is widespread (Seppänen *et*
407 *al.*, 2007), from the point of view of species' conservation, it is important to keep in mind that the
408 effect of close relatives and ecologically similar species may also be positive.

409 Predicting species responses to various global changes has become crucial due to the ongoing
410 biodiversity crisis. Our results accompany earlier literature (e.g. Araújo & Rozenfeld, 2014)
411 suggesting that a failure to incorporate species interactions may account for the mixed results of
412 earlier species distribution modeling efforts that ignore interactions.

413

414 **Acknowledgements**

415 We are grateful to the hundreds of volunteers involved in the French and Finnish bird censuses. Data
416 are available upon request from AL and VD. We are also grateful to the Academy of Finland (project#
417 275329 to MM and project# 275606 to AL) and Kone Foundation (to ME and JTF) for funding. We
418 thank M. Puurtinen and A. Kahilainen for useful piece of statistical advice as well as the Editor of
419 *Global Ecology and Biogeography* and two anonymous referees for their comments, which
420 significantly helped to improve the paper. K. Eyvindson kindly edited the text. This paper was initially
421 submitted, reviewed and revised in *Peerage of Science* (<http://www.peerageofscience.org/>), and we
422 are grateful to peers #1093, #1096, #1101, and #1111 for very constructive comments.

423

424 **Supporting Information**

425 Additional supporting information may be found in the online version of this article at the
426 publisher's web-site.

427 **Appendix S1** Functional traits

428 **Appendix S2** Detailed flowchart of the analyses

429 **Appendix S3** R script for calculating the effect sizes

430 **Appendix S4** Frequency distributions for observed coefficients, mean expected coefficients and effect sizes

431

432

433 **Biosketch**

434 Mikko Mönkkönen is a professor in applied ecology at the University of Jyväskylä, Finland. His main
435 research interests are community and landscape ecology, conservation biology and environmental
436 economics.

437

438 **References**

439 Araújo, M.B. & Rozenfeld, A. (2014) The geographic scaling of biotic interactions. *Ecography*, **37**,
440 406–415.

441 Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014) lme4: Linear mixed-effects models using Eigen
442 and S4. R package version 1.1-7, <URL: <http://CRAN.R-project.org/package=lme4>>.

443 Belmaker, J., Zarnetske, P., Tuanmu, M.-N., Zonneveld, S., Record, S., Strecker, A. & Beaudrot, L.
444 (2015) Empirical evidence for the scale dependence of biotic interactions. *Global Ecology and*
445 *Biogeography*, **24**, 750–761.

446 Bertness, M.D. & Callaway, R.M. (1994) Positive interactions in communities. *Trends in Ecology &*
447 *Evolution*, **9**, 191–193.

448 Blois, J.L., Gotelli, N.J., Behrensmeier, A.K., Faith, J.T., Lyons, S.K., Williams, J.W., Amatangelo, K.L.,
449 Bercovici, A., Du, A., Eronen, J.T., Graves, G.R., Jud, N., Labandeira, C., Looy, C. V., Mcgill, B.,
450 Patterson, D., Potts, R., Riddle, B., Terry, R., Tóth, A., Villaseñor, A. & Wing, S. (2014) A
451 framework for evaluating the influence of climate, dispersal limitation, and biotic interactions

- 452 using fossil pollen associations across the late Quaternary. *Ecography*, **37**, 1095–1108.
- 453 Boulangeat, I., Gravel, D. & Thuiller, W. (2012) Accounting for dispersal and biotic interactions to
454 disentangle the drivers of species distributions and their abundances. *Ecology Letters*, **15**, 584–
455 593.
- 456 Bruno, J.F., Stachowicz, J.J. & Bertness, M.D. (2003) Inclusion of facilitation into ecological theory.
457 *Trends in Ecology & Evolution*, **18**, 119–125.
- 458 Cade, B.S. & Noon, B.R. (2003) A gentle introduction to quantile regression for ecologists. *Frontiers in*
459 *Ecology and the Environment*, **1**, 412–420.
- 460 Cardinale, B.J., Palmer, M. a & Collins, S.L. (2002) Species diversity enhances ecosystem functioning
461 through interspecific facilitation. *Nature*, **415**, 426–9.
- 462 Connell, J.H. (1983) On the Prevalence and Relative Importance of Interspecific Competition :
463 Evidence from Field Experiments. *American Naturalist*, **122**, 661–696.
- 464 Connor, E.F., Collins, M.D. & Simberloff, D. (2013) The Checkered History of Checkerboard
465 Distributions : Reply. *Ecology*, **94**, 2403–2414.
- 466 Connor, E.F. & Simberloff, D. (1983) Interspecific competition and species co-occurrence patterns on
467 islands: null models and the evaluation of evidence. *Oikos*, **41**, 455–465.
- 468 Cornell, A.H. V & Lawton, J.H. (1992) Species Interactions , Local and Regional Processes , and Limits
469 to the Richness of Ecological Communities : A Theoretical Perspective. *Journal of Animal*
470 *Ecology*, **61**, 1–12.
- 471 Cornell, H. V (1985) Local and regional richness of cynipine gall wasps on California oaks. *Ecology*, **66**,
472 1247–1260.
- 473 Crist, T.O., Veech, J. a, Gering, J.C. & Summerville, K.S. (2003) Partitioning species diversity across
474 landscapes and regions: a hierarchical analysis of alpha, beta, and gamma diversity. *The*
475 *American naturalist*, **162**, 734–43.
- 476 Danchin, E., Giraldeau, L.-A., Valone, T.J. & Wagner, R.H. (2004) Public information: from nosy
477 neighbors to cultural evolution. *Science*, **305**, 487–91.

- 478 Devictor, V., Moullot, D., Meynard, C., Jiguet, F., Thuiller, W. & Mouquet, N. (2010) Spatial mismatch
479 and congruence between taxonomic, phylogenetic and functional diversity: the need for
480 integrative conservation strategies in a changing world. *Ecology Letters*, **13**, 1030–1040.
- 481 Diamond, J.M. (1975) *Assembly of species communities. Ecology and Evolution of communities* (ed.
482 by M.L. Cody) and J.M. Diamond), pp. 342–444. Belknap Press, Cambridge, Massachusetts,
483 USA.
- 484 European Environmental Agency (2015) *The European Environment: State and Outlook 2015. Data
485 and Maps.*.
- 486 Forsman, J.T., Seppänen, J.-T. & Mönkkönen, M. (2002) Positive fitness consequences of interspecific
487 interaction with a potential competitor. *Proceedings. Biological sciences / The Royal Society*,
488 **269**, 1619–23.
- 489 Gaston, K.J. & Blackburn, T.M. (2000) *Pattern and process in macroecology*, Blackwell, Oxford.
- 490 Gerhold, P., Cahill, J.F., Winter, M., Bartish, I. V. & Prinzing, A. (2015) Phylogenetic patterns are not
491 proxies of community assembly mechanisms (they are far better). *Functional Ecology*, **29**, 600–
492 614.
- 493 Godoy, O., Kraft, N.J.B. & Levine, J.M. (2014) Phylogenetic relatedness and the determinants of
494 competitive outcomes. *Ecology letters*, **17**, 836–44.
- 495 Goldberg, D.E., Barton, A.M., The, S., Naturalist, A. & Apr, N. (1992) Patterns and Consequences of
496 Interspecific Competition in Natural Communities : A Review of Field Experiments with Plants.
497 *American Naturalist*, **139**, 771–801.
- 498 Gotelli, N.J., Graves, G.R. & Rahbek, C. (2010) Macroecological signals of species interactions in the
499 Danish avifauna. *Proceedings of the National Academy of Sciences of the United States of
500 America*, **107**, 5030–5035.
- 501 Gotelli, N.J. & Ulrich, W. (2010) The empirical Bayes approach as a tool to identify non-random
502 species associations. *Oecologia*, **162**, 463–477.
- 503 Gross, K. (2008) Positive interactions among competitors can produce species-rich communities.

- 504 *Ecology letters*, **11**, 929–36.
- 505 Gurevitch, J., Morrison, J.A., Hedges, L. V, The, S., Naturalist, A. & April, N. (2000) The Interaction
506 between Competition and Predation : A Meta-analysis of Field Experiments. *American*
507 *Naturalist*, **155**, 435–453.
- 508 Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Körber, J.H. (2007) Biotic interactions
509 improve prediction of boreal bird distributions at macro-scales. *Global Ecology and*
510 *Biogeography*, **16**, 754–763.
- 511 Hornik, K. (2005) A CLUE for CLUster Ensembles. *Journal of Statistical Software*, **14**.
- 512 Houlahan, J.E., Currie, D.J., Cottenie, K., Cumming, G.S., Ernest, S.K.M., Findlay, C.S., Fuhlendorf, S.D.,
513 Gaedke, U., Legendre, P., Magnuson, J.J., McArdle, B.H., Muldavin, E.H., Noble, D., Russell, R.,
514 Stevens, R.D., Willis, T.J., Woivod, I.P. & Wondzell, S.M. (2007) Compensatory dynamics are
515 rare in natural ecological communities. *Proceedings of the National Academy of Sciences of the*
516 *United States of America*, **104**, 3273–7.
- 517 Hutchinson, G.E. (1959) Homage to Santa-Rosalía Or Why are there so Many Kinds of Animals.
518 *American Naturalist*, **93**, 145–159.
- 519 Jiguet, F., Devictor, V., Julliard, R. & Couvet, D. (2012) French citizens monitoring ordinary birds
520 provide tools for conservation and ecological sciences. *Acta Oecologica*, **44**, 58–66.
- 521 Järvinen, O. & Väisänen, R.A. (1980) Quantitative biogeography of Finnish land birds as compared
522 with regionality in other taxa. *Annales Zoologici Fennici*, **17**, 67–85.
- 523 Kivelä, S.M., Seppänen, J.-T., Ovaskainen, O., Doligez, B., Gustafsson, L., Mönkkönen, M. & Forsman,
524 J.T. (2014) The past and the present in decision-making: the use of con- and heterospecific cues
525 in nest-site selection. *Ecology*, **95**, 3428–3439.
- 526 Koenker, R. (2013) *quantreg: Quantile Regression. R package version 5.05*. [http://CRAN.R-](http://CRAN.R-project.org/package=quantreg)
527 [project.org/package=quantreg](http://CRAN.R-project.org/package=quantreg).
- 528 Kosicki, J.Z., Stachura, K., Ostrowska, M. & Rybska, E. (2015) Complex species distribution models of
529 Goldcrests and Firecrests densities in Poland: are remote sensing-based predictors sufficient?

- 530 *Ecological Research*, **30**, 625–638.
- 531 Laaksonen, T. & Lehikoinen, A. (2013) Population trends in boreal birds: Continuing declines in
532 agricultural, northern, and long-distance migrant species. *Biological Conservation*, **168**, 99–107.
- 533 Legendre, P. & Legendre, L. (1998) *Numerical Ecology*, Elsevier Science, Amsterdam.
- 534 Loukola, O.J., Seppänen, J.-T., Krams, I., Torvinen, S.S. & Forsman, J.T. (2013) Observed fitness may
535 affect niche overlap in competing species via selective social information use. *The American*
536 *naturalist*, **182**, 474–83.
- 537 MacArthur, R. & Levins, R. (1967) The Limiting Similarity, Convergence, and Divergence of
538 Coexisting Species. *The American Naturalist*, **101**, 377–385.
- 539 MacArthur, R.H. (1958) Population Ecology of Some Warblers of Northeastern Coniferous Forests.
540 *Ecology*, **39**, 599–619.
- 541 Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2016) cluster: Cluster Analysis
542 Basics and Extensions. R package version 2.0.5.
- 543 Martorell, C. & Freckleton, R.P. (2014) Testing the roles of competition, facilitation and stochasticity
544 on community structure in a species-rich assemblage. *Journal of Ecology*, **102**, 74–85.
- 545 Mouchet, M., Guilhaumon, F., Villéger, S., Mason, N.W.H., Tomasini, J.-A. & Moullot, D. (2008)
546 Towards a consensus for calculating dendrogram-based functional diversity indices. *Oikos*, **117**,
547 794–800.
- 548 Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.F., Chave, J., Coutron, P., Dalecky,
549 A., Fontaine, C., Gravel, D., Hardy, O.J., Jabot, F., Lavergne, S., Leibold, M., Moullot, D.,
550 Münkemüller, T., Pavoine, S., Prinzing, A., Rodrigues, A.S.L., Rohr, R.P., Thébault, E. & Thuiller,
551 W. (2012) Ecophylogenetics: Advances and perspectives. *Biological Reviews*, **87**, 769–785.
- 552 Mutshinda, C.M., O’Hara, R.B. & Woiwod, I.P. (2009) What drives community dynamics? *Proceedings*
553 *of the Royal Society B*, **276**, 2923–2929.
- 554 Mönkkönen, M., Helle, P. & Soppela, K. (1990) Numerical and behavioural responses of migrant
555 passerines to experimental manipulation of resident tits (*Parus* spp.): heterospecific attraction

- 556 in northern breeding bird communities? *Oecologia*, **85**, 218–225.
- 557 Mönkkönen, M., Härdling, R., Forsman, J.T., Tuomi, J., Härdling, R., Forsman, J.T. & Tuomi, J. (1999)
- 558 Evolution of heterospecific attraction: using other species as cues in habitat selection.
- 559 *Evolutionary Ecology*, **13**, 91–104.
- 560 Petchey, O.L., Evans, K.L., Fishburn, I.S. & Gaston, K.J. (2007) Low functional diversity and no
- 561 redundancy in British avian assemblages. *The Journal of animal ecology*, **76**, 977–85.
- 562 Pigot, A.L. & Etienne, R.S. (2015) A new dynamic null model for phylogenetic community structure.
- 563 *Ecology Letters*, **18**, 153–163.
- 564 R Development Core Team (2014) R: A language and environment for statistical computing.
- 565 Ricklefs, R.E. (2012) Habitat-independent spatial structure in populations of some forest birds in
- 566 eastern North America. *Journal of Animal Ecology*, **82**, 145–154.
- 567 Schoener, T.W. (1983) Field Experiments on Interspecific Competition. *American Naturalist*, **122**,
- 568 240–285.
- 569 Sebastián-González, E., Sánchez-Zapata, J.A., Botella, F. & Ovaskainen, O. (2010) Testing the
- 570 heterospecific attraction hypothesis with time-series data on species co-occurrence.
- 571 *Proceedings. Biological sciences / The Royal Society*, **277**, 2983–90.
- 572 Sekercioglu, C.H. (2006) Increasing awareness of avian ecological function. *Trends in Ecology and*
- 573 *Evolution*, **21**, 464–471.
- 574 Seppänen, J.-T. & Forsman, J.T. (2007) Interspecific social learning: novel preference can be acquired
- 575 from a competing species. *Current biology : CB*, **17**, 1248–52.
- 576 Seppänen, J.-T., Forsman, J.T., Mönkkönen, M. & Thomson, R.L. (2007) Social information use is a
- 577 process across time, space, and ecology, reaching heterospecifics. *Ecology*, **88**, 1622–1633.
- 578 Thomson, R.L., Forsman, J.T. & Mönkkönen, M. (2003) Positive interactions between migrant and
- 579 resident birds: testing the heterospecific attraction hypothesis. *Oecologia*, **134**, 431–8.
- 580 Thuiller, W., Lavergne, S., Roquet, C., Boulangeat, I., Lafourcade, B. & Araujo, M.B. (2011)
- 581 Consequences of climate change on the tree of life in Europe. *Nature*, **470**, 531–4.

- 582 Trisos, C.H., Petchey, O.L. & Tobias, J.A. (2014) Unraveling the Interplay of Community Assembly
583 Processes Acting on Multiple Niche Axes across Spatial Scales. **184**.
- 584 Veech, J.A. (2013) A probabilistic model for analysing species co-occurrence. *Global Ecology and*
585 *Biogeography*, **22**, 252–260.
- 586 Violle, C., Nemergut, D.R., Pu, Z. & Jiang, L. (2011) Phylogenetic limiting similarity and competitive
587 exclusion. *Ecology letters*, **14**, 782–7.
- 588 Webb, C.O., Ackerly, D.D., McPeck, M. a. & Donoghue, M.J. (2002) Phylogenies and Community
589 Ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.
- 590

591 Table 1 Biogeographic zones, habitat types, and number of point count stations, observations and species
 592 included in each of the three datasets (French farmlands, French forests, Finnish forests)

Data set	Biogeographic zone	Habitat type	# point count stations	# observations	# species
French farmlands	Alpine	Ploughed meadow	60	1,291	74
		Unploughed meadow	70	1,372	75
		Mixed farmland	91	1,963	76
		Open field	122	2,485	76
		Permanent crop	380	7,358	78
	Atlantic	Ploughed meadow	496	11,930	81
		Unploughed meadow	595	14,666	82
		Mixed farmland	1,168	28,456	82
		Open field	1,759	34,626	82
		Permanent crop	259	6,306	80
	Continental	Ploughed meadow	506	12,228	82
		Unploughed meadow	1,306	33,795	82
		Mixed farmland	859	21,436	82
		Open field	976	19,763	81
		Permanent crop	128	3,125	79
	Mediterranean	Ploughed meadow	37	726	67
		Unploughed meadow	134	2,637	73
Mixed farmland		49	1,095	70	
Open field		6	177	55	
		Permanent crop	25	538	60
Total			9,026	205,973	83
French forests	Alpine	Deciduous woodland	94	1,543	69
		Coniferous woodland	113	1,806	61
		Mixed woodland	161	2,952	67
	Atlantic	Deciduous woodland	1,136	25,670	77
		Coniferous woodland	216	4,537	74
		Mixed woodland	274	6,715	74
	Continental	Deciduous woodland	1,519	33,458	75
		Coniferous woodland	311	6,144	76
		Mixed woodland	440	9,227	75
	Mediterranean	Deciduous woodland	181	3,054	71
		Coniferous woodland	138	2,490	71
Mixed woodland		187	3,282	75	
Total			4,770	100,878	77
Finnish forests	Hemi- and south boreal	Spruce forest	592	9,288	76
		Pine forest	563	9,322	76
		Deciduous forest	320	5,325	75
		Mixed forest	900	15,289	75
	Mid-boreal	Spruce forest	51	737	56
		Pine forest	134	1,887	68
		Deciduous forest	30	437	55
		Mixed forest	90	1,262	64

North boreal	Spruce forest	30	427	48
	Pine forest	109	1,081	50
	Deciduous forest	21	328	43
	Mixed forest	77	936	53
Total		2,917	46,319	76

593

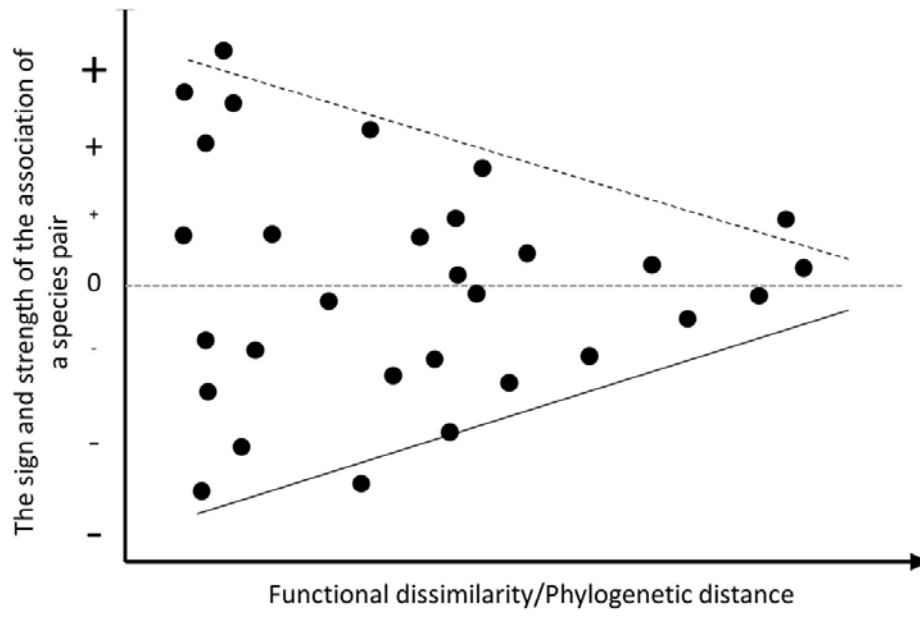
594 **Figure captions**

595 **Figure 1** Schematic figure of the relationship between increasing functional dissimilarity/phylogenetic distance
596 (x-axis), and the sign [positive (+), neutral (0), negative (-)] and the strength (increases with increasing symbol
597 size) of the association of a species pair (y-axis). Every black dot represents an imaginary pair of species. Both
598 heterospecific information use (black dashed line) and the limiting similarity principle (solid line)
599 predict that species associations are most intense among functionally and phylogenetically similar
600 species, but in opposite directions. Heterospecific information use predicts positive association
601 between functionally/phylogenetically similar species. Since the information value decreases with
602 increasing functional similarity/phylogenetic distance so does the strength of the association. The
603 limiting similarity predicts functionally/phylogenetically similar species to have negative association
604 whereas functional dissimilar/phylogenetically distant species show neutral associations. The net
605 outcome of an association between similar species depends on the costs of competition and benefits
606 of information use and may thus result in a neutral association, and altogether they form a 'funnel
607 plot'.

608 **Figure 2** Left-hand panels show the relationship between increasing functional dissimilarity and the strength
609 and the sign of an association of the abundances of each species pair in French farmlands (a), French forests
610 (b), and Finnish forests (c). The strength and the sign of an association of the abundances is measured as the
611 mean of the standardized effect sizes (the difference between observed pairwise regression
612 coefficient and mean expected coefficient, divided by the standard deviation) between a species pair
613 (see text for further information). In a case of statistically significant relationship the regression line
614 for the quantile regressions in lower ($\tau = 0.05$) and/or upper ($\tau = 0.95$) quantiles are shown. Right-
615 hand panels show the relationship between all quantiles (τ) at intervals of 0.01 and the coefficient
616 from the quantile regressions (standard errors are shown in grey). Positive coefficients denote a
617 positive relationship in a given quantile between increasing functional dissimilarity and the strength
618 of an association of the abundances whereas negative coefficients denote a negative relationship.

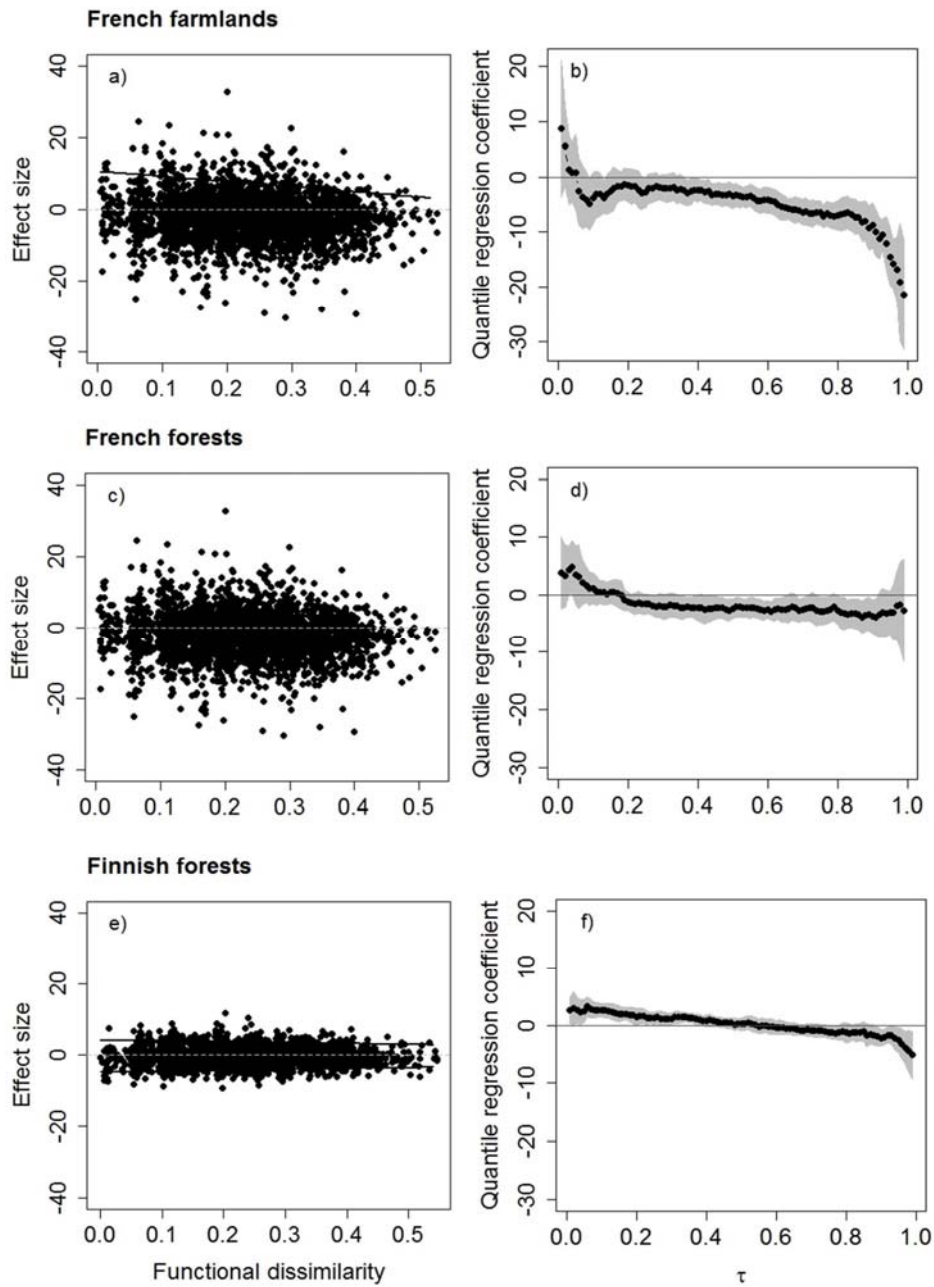
619 **Figure 3** Left-hand panels show the relationship between increasing relative phylogenetic distance and the
620 strength and the sign of an association of the abundances of a species pair in French farmlands (a), French
621 forests (b), and Finnish forests (c). The strength and the sign of an association of the abundances is measured
622 as the mean of the standardized effect sizes (the difference between observed pairwise regression
623 coefficient and mean expected coefficient, divided by the standard deviation) between a species pair
624 (see text for further information). In a case of statistically significant relationship the regression line
625 for the quantile regressions in lower ($\tau = 0.05$) and/or upper ($\tau = 0.95$) quantiles are shown. Right-
626 hand panels show the relationship between all quantiles (τ) at intervals of 0.01 and the coefficient
627 from the quantile regressions (standard errors are shown in grey). Positive coefficients denote a
628 positive relationship in a given quantile between increasing relative phylogenetic distance and the
629 strength of an association of the abundances whereas negative coefficients denote a negative
630 relationship.

631



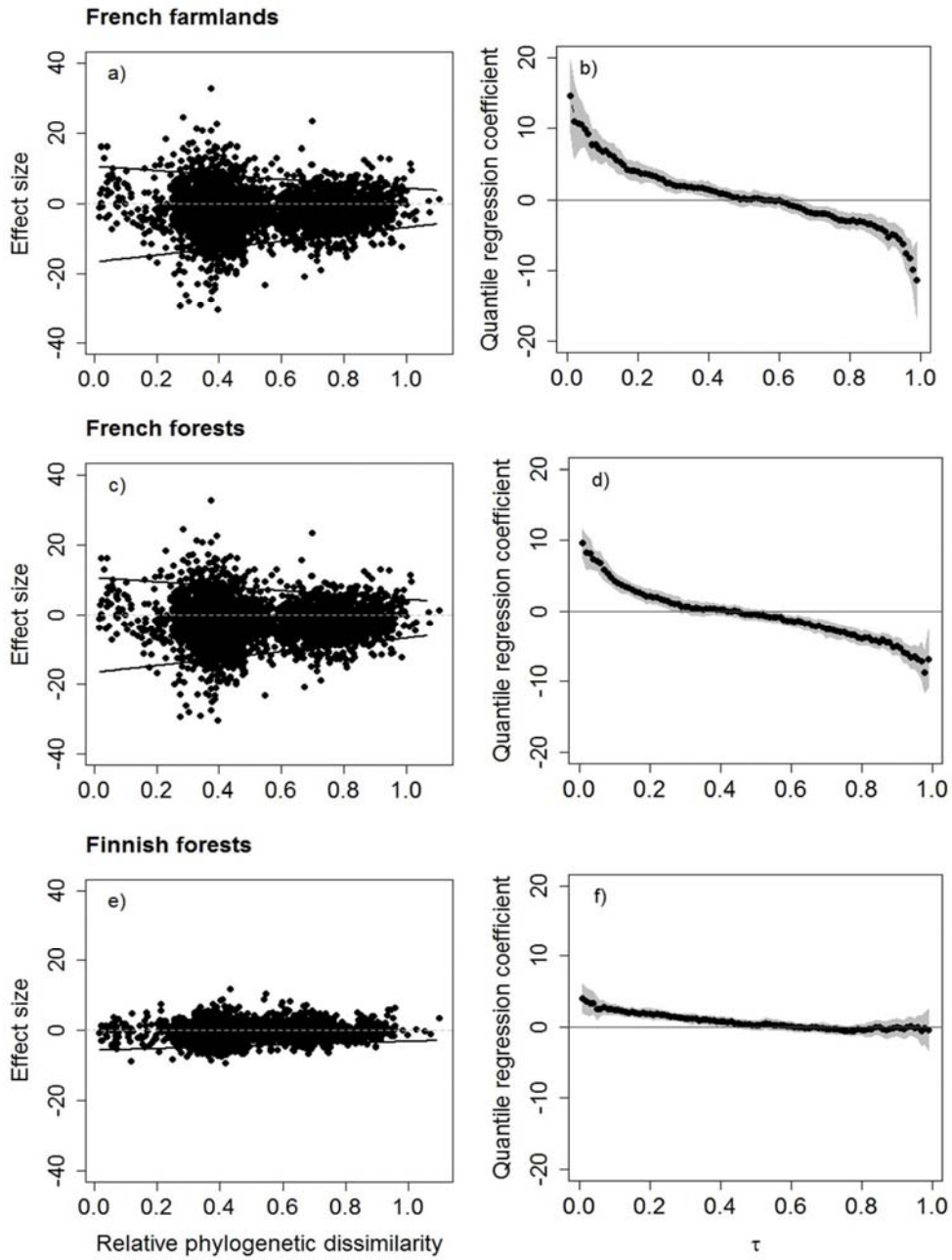
632

633 Fig.1



634

635 Fig. 2



636

637

638 Fig. 3