1   Dark diversity illuminates the dim side of conservation and restoration

2   Jesper Erenskjold Moeslund[1*], Ane Kirstine Brunbjerg[1,2], Kevin Kuhlmann Clausen[3], Lars Dalby[1], Camilla

3   Fløjgaard[1], Anders Juel[4], Jonathan Lenoir[5]

4   1: Section for Biodiversity & Conservation, Department of Bioscience, Aarhus University, Kalø, Denmark

5   2: School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

6   3: Section for Wildlife Ecology, Department of Bioscience, Aarhus University, Kalø, Denmark

7   4: Section for Nature Conservation, The Nature Agency, Ministry of Environment and Food of Denmark,

8   Copenhagen, Denmark

9   5: UR "Ecologie et dynamique des systèmes anthropisés" (EDYSAN, FRE 3498 CNRS-UPJV), Université de

10   Picardie Jules Verne, France

11   * Corresponding author. E-mail: jesper.moeslund@bios.au.dk, mail-address: Section for Biodiversity &

12   Conservation, Department of Bioscience, Aarhus University, Grenåvej 14, 8410 Rønde, Kalø, Denmark,

13   phone: +45 87158986, fax: N/A.

14

Summary

1 *Dark diversity* is a promising concept for prioritizing management efforts as it focuses on missing species, i.e., species present in the regional pool, but locally absent despite suitable environmental conditions. **2** We applied the concept to a massive national plant diversity database (236,923 records from 15,160 surveys involving 564 species) to provide the first geographically comprehensive assessment of dark diversity across a large area (43,000 km^2), at a spatial scale (~75 m^2) relevant for conservation and restoration planning and across multiple terrestrial habitats, thus maximising its practical application potential. The likelihood for a given plant species to belong to the dark diversity pool was computed and logistically regressed against its ecological preferences (nutrient availability, pH etc.), strategies (competitor, stress tolerance, ruderal), mycorrhizal dependence and infection percentage, seed mass and maximum dispersal distance. **3** Forty-six percent of the species were absent in >95 % of the suitable sites, whereas 7 % of the species were absent in less than 60 % of sites that were deemed suitable. **4** Species that were more likely to belong to the dark diversity tended to depend on mycorrhiza, were mostly adapted to low light and nutrient levels, had poor dispersal abilities, were ruderals and had a low stress tolerance. **Synthesis and applications** Our findings have important implications for the planning and management of natural ecosystems requiring detailed knowledge of what triggers the presence/absence of individual plant species in a seemingly suitable habitat. We conclude that practitioners may need to carefully consider mycorrhizal inoculations with a suitable assemblage of fungi for certain plant species to become established. Also assisted migration might be necessary to help poor dispersers although spatial and temporal processes are also important to have in mind. Finally, it is important to vary nutrient loads making room for plant species to colonise both nutrient-poor and nutrient-rich localities.

Introduction

Recently, Pärtel and co-workers (2011) presented a new concept coined *dark diversity*, which could prove

to be a central idea for the development of effective tools for practical biodiversity management and

conservation prioritization at relevant spatial scales. Dark diversity encompasses the diversity articulated by

all species missing locally, even though biogeographic history and current ecological and environmental

conditions suggest their presence (Pärtel, Szava-Kovats & Zobel, 2011). In other words, dark diversity is the

set of species belonging to the regional species pool of a particular habitat but that are missing locally

within a given site of that habitat (Pärtel, Szava-Kovats & Zobel, 2011). Note that the dark diversity concept

differs from the so-called *hidden diversity* which counts species that are overlooked due to observation bias

(Milberg *et al.*, 2008).

Possible causes for species belonging to the dark diversity are manifold and include, but are not

limited to, lower-level ecological filters involving metapopulation and metacommunity dynamics (cf.

dispersal limitations and habitat fragmentation, Tilman, 1997; Fahrig, 2003) or complex biotic interactions

(e.g. competition, parasitism, mutualism and symbiotic phenomena like mycorrhiza, Grime, 1979; Torrez *et*

al., 2016). To our knowledge, the likelihood of individual species to belong to the dark diversity has never

been studied and little is known about the characteristics of species having a higher chance to belong to the

dark diversity pool than others (but see Riibak *et al.*, 2015).

Recently, the dark diversity concept has been proposed as a new facet of biodiversity for gaining

useful knowledge for restoration and conservation issues (Pärtel, Szava-Kovats & Zobel, 2011). However, till

now only few studies have been inspired by this (Gijbels, Adriaens & Honnay, 2012; Yoshioka *et al.*, 2014;

Riibak *et al.*, 2015; Ronk, Szava-Kovats & Pärtel, 2015). Indeed, these have been successful albeit suffered

from various weaknesses preventing a more general application: (1) spatial extent or resolution was not

directly relevant for practical nature conservation and restoration planning (Riibak *et al.*, 2015; Ronk, Szava-

Kovats & Pärtel, 2015), (2) scope was only a single habitat or ecosystem (Yoshioka *et al.*, 2014; Riibak *et al.*,

62   2015) or (3) only a few species were involved (Gijbels, Adriaens & Honnay, 2012). Although the restoration

63   potential of a given site may depend on a number of site-specific factors, like fragmentation and biotic

64   interactions within the focal community, searching across multiple sites and habitats for common traits

65   among typical dark diversity species may clarify potential drivers of the observed site-specific species

66   distribution patterns. Hence, studying the features of species that belong to the dark diversity more often

67   than others could prove to be an important key for successful practical application of dark diversity in

68   restoration and conservation.

69   Clearly, the delimitation of the regional species pool can impact the assessment of dark diversity

70   massively. Consequently, this issue has been one of the major concerns of the original dark diversity

71   approach: if habitats are delimited rather categorically as suggested in the original concept paper (Pärtel,

72   Szava-Kovats & Zobel, 2011) the actual natural environmental gradients are ignored (Mokany & Paini,

73   2011). For plants, regional species pools have been successfully identified using the species indicator values

74   presented by Ellenberg *et al.* (2001) for Central European plant species (Pärtel *et al.*, 1996). Ellenberg's

75   indicator values (EIVs) represent European plant species' preferred position along various environmental

76   gradients and are often used in vegetation studies (Ellenberg *et al.*, 2001; Diekmann, 2003; Lenoir *et al.*,

77   2010; Moeslund *et al.*, 2013). Recently, Ewald (2002) proposed a probabilistic procedure (Beals' index) to

78   estimate regional species pools based on co-occurrence patterns among species. This approach is currently

79   known to yield the most realistic estimates of the regional species pool (Lewis, Szava-Kovats & Pärtel,

80   2016).

81   Restoration, conservation and nature management typically take place at relatively fine spatial

82   resolutions. Obviously, considering dark diversity at a resolution relevant for conservation and restoration

83   management is likely to inflate the dark diversity simply because smaller areas support fewer species all

84   else being equal (McArthur & Wilson, 1967). For this reason, comparing dark diversity from areas of

85   different sizes where co-occurrence data were collected in a different manner is not meaningful without

4

86   accounting for these effects. To ensure a reliable assessment of dark diversity, co-occurrence data collected

87   at fine spatial resolution across large spatial extent and in a systematic manner – like in national

88   biodiversity inventories (e.g. Fredshavn, Nygaard & Ejrnæs, 2009) – is needed.

89   Here we present the first national assessment of the characteristics of typical dark diversity plants at

90   a spatial resolution relevant to conservation and restoration management, covering multiple open

91   terrestrial habitats. Using a large national plant dataset with high spatial accuracy and a combination of

92   EIVs, Grime's plant strategies, mycorrhizal information and dispersal distance calculations, we address the

93   following specific study questions: (1) does North-European plant species differ in how often they occur in

94   the dark diversity pool (species' likelihood to belong to dark diversity)? (2) If so, which plant traits or

95   ecological characteristics explain this pattern the best? Finally, we discuss the causal mechanisms most

96   likely involved and how our findings may aid effective planning and management initiatives and promote

97   the practical application of the dark diversity concept within conservation and restoration.

Methods

98

Vegetation data

99

100   Data on the distribution of vascular plants in Denmark was obtained from municipalities' vegetation

101   inventory of natural habitat types (Fredshavn, Nygaard & Ejrnæs, 2009). We used observations from 5-m

102   radius circular plots laid out to capture the typical flora of a particular site in question (typically one plot per

103   site). The sites are 5 ha on average (ranging between 0.003 – 900 ha), distributed throughout most of

104   Denmark (Fig. 1) and cover freshwater meadows, salt meadows, heathlands, bogs, moors, fens, grasslands

105   and vegetated dunes (i.e. open habitats). The dataset was extracted 6 October 2014. We used data from

106   2004 – 2014 encompassing 236,923 records from 15,160 plots involving 564 plant species after application

107   of the filters described in the following. We only considered observations at the species level and excluded

108   all neophytes (Appendix tables 6–8 in Buchwald *et al.*, 2013), i.e. species that are not considered a natural

109   part of the vegetation given their history and dispersal ability), shrubs, trees and submersed aquatic

110   species. To ensure meaningful calculations of the regional species pool (see below) only plots with more

111   than five plant species records were used.

112   **Regional species pool and dark diversity**

113   To yield the best estimates of dark diversity, we used Beals' index (Beals, 1984) to assess the regional

114   species pool for each plot as recommended by Lewis, Szava-Kovats & Pärtel (2016). For each plot, Beals'

115   index represents the probability that a focal species will occur within that plot given the assemblage of

116   species co-occurring there (McCune, 1994). See details in Münzbergova & Herben (2004). Initially, a

117   presence/absence matrix with all combinations of plot and species was calculated. Based on this matrix, we

118   calculated Beals' index for each species in each plot excluding the focal species from the calculations (as

119   recommended by Oksanen *et al.*, 2015) (Fig. 2). We used the "beals()" function in the "vegan" package

120   (Oksanen *et al.*, 2015). The threshold for including a species in the regional species pool was defined as the

121   5th percentile of the Beals' index value for the species following Gijbels, Adriaens & Honnay (2012) as well

122   as Ronk, Szava-Kovats & Pärtel (2015). Additionally, we only considered data for plots having Beals' index

123   values above that of the lowest value where the species was indeed present. For every plot the dark

124   diversity was composed by all species in the regional pool excluding those that were actually present

125   (Pärtel, Szava-Kovats & Zobel, 2011) (Fig. 2).

126   **Response variable**

127   As a response variable, we computed the species' likelihood to belong to dark diversity (cf. the ratio of the

128   number of occurrences in the dark diversity pool divided by the number of occurrences in the regional

129   species pool) (Figs 2 & 3) for each of the 564 plant species used in our analyses.

130   **Species traits and characteristics (explanatory variables)**

131   <u>Ellenberg's indicator values</u>

6

132   We used Ellenberg's indicator values (EIVs) (Ellenberg *et al.*, 2001) adjusted to British conditions (Hill *et al.*,

133   1999) as these adjusted values are thought to best match the Danish flora (Moeslund *et al.*, 2013).

134   Variation in temperature (EIV_T) and continentality (EIV_K) in Denmark is negligible and salinity (EIV_S) is only

135   relevant in coastal environments. Consequently, we only considered EIVs for: (1) soil moisture (EIV_F), (2) soil

136   pH (EIV_R), (3) soil nutrient status (EIV_N) and (4) ambient light (EIV_L) (Table S1). EIVs for soil nutrient status

137   and pH are typically highly correlated (Diekmann & Falkengren-Grerup, 1998; Seidling, 2005); hence we

138   calculated a nutrient/pH-ratio ($EIV_{N/R}$) as an alternative variable based on the two corresponding EIVs. This

139   variable represents the species' preference for nutrient availability (see for example Andersen *et al.*, 2013).

140   <u>Grime plant strategies</u>

141   The main plant strategies presented by Grime (1979) enables scientists to distinguish between plants

142   adapted to competitive (C-species), stressful (S-species) or ruderal (R-species) environments. Although

143   plants can harbour any combination of these three strategies they are in their extreme forms mutually

144   exclusive (Grime, 1979). For this study, we obtained the strategies for each plant species from the BiolFlor

145   database (Kühn, Durka & Klotz, 2004). Grime plant strategy data was available for all species included here

146   (Table S1). Following Ejrnæs & Bruun (2000), we represented the degree to which a plant is adapted to a

147   given strategy as values ranging from 1–12 for each of the three strategies, however restraining their sum

148   to 12.

149   <u>Mycorrhiza data</u>

150   We used data on both mycorrhizal infection percentage (0 to 100%) and dependence (i.e., a factor variable

151   with two levels: obligately vs. not obligately mycorrhizal). Data on mycorrhizal infection was retrieved from

152   Akhmetzhanova (2012) and data on mycorrhizal dependence was taken from MycoFlor (Hempel *et al.*,

153   2013). These data were available for 33 % and 82 %, respectively, of the plant species involved in this study

154   (Table S1).

155

156 Plant functional traits

157 The trait data has two purposes: (1) it is used for modelling dispersal distance (see below) and (2) seed

158 mass is used as an explanatory variable representing an alternative measure of dispersal distance as well as

159 the plants ability to establish at new sites. We obtained seed mass (SM), dispersal syndrome (DS), releasing

160 height (RH), terminal velocity (TV) and growth form (GF) data from the LEDA and BiolFlor databases (Kühn,

161 Durka & Klotz, 2004; Kleyer *et al.*, 2008). Where multiple records of SM, RH and TH were available for the

162 same species, the mean value was used. Missing data on DS and GF was taken from Hansen (1996) making

163 these two traits available for all species (94 % and 77 % of the species respectively were covered by the

164 aforementioned databases). Data on SM, RH and TV was available for 89 %, 91 % and 68 % of our study

165 species, respectively.

166 Subsequently, we calculated maximum dispersal distance (MDD) using the "dispeRsal()" function (Tamme

167 *et al.*, 2014). This function calculates MDD using plant traits and taxonomy. For 70% of the species, MDD

168 was calculated with a model including DS, GF and TV. For the remaining 30% of the species MDD was

169 calculated using simpler models following the hierarchy of best predictive performance given in Tamme *et*

170 *al.* (2014). For species with multiple different entries of a trait (e.g. DS) we calculated the mean of the

171 predicted MDDs. Table S1 lists the MDDs calculated for each species.

172 **Data analysis**

173 We used binomial generalised linear models (GLMs) for proportion data to explore the relationship

174 between the plants' likelihood of being part of the dark diversity (binomial response variable; no. of times

175 in dark diversity/no. of times in the regional species pool) and the 13 explanatory variables listed in Table 1.

176 All variables were tested for multicollinearity (see Figure S1). The EIVs for nutrients and pH both showed a

177 high degree of multicollinearity (tolerances below 0.15, Quinn & Keough, 2002) and therefore we decided

178 to use the nutrient/pH-ratio (described above) instead to represent the plants' nutrient preferences (Table

179 1).

8

180   Not all traits and characteristics were available for every species in the dataset (see method sections

181   above). In order to maximize the sample size we initially ran model selection (see section below) on models

182   including only the explanatory variables which were available for the majority of the species (n = 564, Table

183   S2). These models are referred to as *high sample size models* in the following. Being available for a lower

184   number of species (n = 457) dispersal distance was subsequently added to the models subject to model

185   selection (*medium sample size* models). Finally, we ran model selection on models with all explanatory

186   variables listed in Table 1 (n = 151, *low sample size* models). Since the three Grime-based variables were

187   highly dependent on each other, only one Grime variable was included in a model at a time. Also, since the

188   calculations of maximum dispersal distance involved seed mass these two variables never occurred

189   simultaneously in any of the models to avoid redundancy. Following the above-mentioned constraints, we

190   built three sets of candidate models and used Akaike's Information Criteria (AIC, Aikaike, 1974) to select the

191   best model within each sample size group following best practice as recommended in Burnham & Anderson

192   (2002) (for details on model selection and setup, see Table S2).

193   A typical issue in the investigation of plant species traits and characteristics is phylogenetic

194   autocorrelation, i.e. the fact that closely related species also tend to be more similar in traits and

195   characteristics (Gittleman & Kot, 1990). Accordingly, we checked all models' residuals for phylogenetic

196   autocorrelation as described below. We used the Daphne phylogenetic tree for the European flora (Durka

197   & Michalski, 2012) and followed the vignette by Paradis (2015) to calculate Moran's I of each model's

198   residuals using the reciprocal phylogenetic distances between species. This computation was performed

199   using the "Moran.I()" function in the "ape" package (Paradis, Claude & Strimmer, 2004). Six models showed

200   significant phylogenetic autocorrelation. To account for this we constructed phylogenetic eigenvector

201   filters following best practise (Borcard & Legendre, 2002; Diniz-Filho & Bini, 2005). Filters that explained

202   significant variation in the best models' residuals were added successively to these best models in addition

203   to the explanatory variables. After each addition of a filter, the model residuals were checked for

204   autocorrelation. When Moran's I showed no significant autocorrelation the process stopped. Significance

205   and effect sizes of explanatory variables in the models were reassessed but no further model selection was

206   performed.

207   All analyses were conducted using R (R Core Team, 2015).

Results

209   On average, species were part of the dark diversity in 88.6 % of the plots for which the species was indeed

210   in the regional pool (Table S1).

211   We found phylogenetic autocorrelation in the residuals of six models (Moran's I test, $P < 0.05$). The

212   addition of 1 to 8 phylogenetic filters successfully removed autocorrelation and caused no notable shifts in

213   effects sizes or significance (Table 2, Table S3).

214   The goodness of fit for our models was up to 0.14 (D^2_{adj}), with the more complex low sample size

215   models – including potential maximum dispersal rate, average seed mass and mycorrhizal information –

216   having the best fits (Table 2). Our models suggest that the factors best explaining the plants' likelihood of

217   being in the dark diversity are (mentioned in the order of importance): mycorrhizal dependence,

218   preference for light and nutrients, ruderality, maximum dispersal distance, seed mass, mycorrhizal infection

219   percentage, stress-tolerance, competitive ability and preference for soil moisture (Table 2). We found

220   strong indications that obligate mycorrhizal plants are more often part of the dark diversity than plants not

221   depending on mycorrhiza (Fig. 4). This finding was supported by the fact that species with higher dark

222   diversity likelihood had a higher degree of infection by mycorrhiza (Table 2, Fig. 4). In addition, plants more

223   frequently in the dark diversity were adapted to thrive under low nutrient availability and low-light (Fig. 4),

224   had poor dispersal abilities and heavier seeds. In terms of Grime strategies, the high dark-diversity-

225   likelihood species were more ruderal, less competitive and less stress-tolerant (Table 2).

Discussion

227 **Differences in species' dark diversity likelihood**

228 Intuitively, typical dark diversity species are rarer than their habitat would suggest, for example rare species

229 with a fairly common habitat (Fig. 3). A species such as *Tephroseris palustris* (L.) Rchb. was often found in

230 the dark diversity (Fig. 3a). This species is rare in Denmark although it has a wide distribution in Northern

231 Europe including Scandinavia, the Baltic region and all the way to arctic Russia (Kochjarová, 2006). It is

232 extinct in Great Britain, the Czech Republic, Hungary, Slovakia and Romania (Kochjarová, 2006) and

233 critically endangered in Sweden (Olsson & Tyler, 2001). However, it tends to colonize bare mud along pond

234 margins and is even known from recycling depots (Frederiksen, Rasmussen & Seberg, 2006) – habitats that

235 are far from rare in Denmark. Hence, habitat availability alone cannot explain its rarity. This was also the

236 case for *Leontodon hispidus* L., *Campanula persicifolia* L., *Vicia tetrasperma* (L.) Schreb. and several others

237 (Table S1). On the other hand, species that occurred less often in the dark diversity were in many cases

238 species that are common within their habitats, such as *Calluna vulgaris* (L.) Hull (common in heathlands,

239 Fig. 3) and *Tripolium vulgare* Nees (common in salt meadows) or alternatively they were common

240 generalists like *Achillea millefolium* L., *Festuca rubra* L., *Urtica dioica* L. etc. (Frederiksen, Rasmussen &

241 Seberg, 2006).

242 **The relative importance of explanatory factors and their likely causal mechanisms**

243 Apart from stochasticity (Hubbell, 2001), there are generally three factors important for species' dark

244 diversity likelihood: (1) dispersal ability, (2) establishment success and (3) persistence in a given habitat.

245 The explanatory factors tested in this study all fall within these three categories. Factors primarily involved

246 in establishment and persistence were the most important ones in our study overall: mycorrhizal

247 dependence and the plants' preference for available nutrients and light. Dispersal related factors such as

248 the plants' maximum dispersal distance and the seed mass were also important, with the latter playing a

249 role for both dispersal and establishment (see below).

250 Establishment and persistence

251   A key ability in plants' life cycle is their establishment once the seed has settled at a given site (Muller-

252   Landau *et al.*, 2002). Many factors such as seed herbivory, seed resistance to pathogens, stress-events like

253   drought or flooding and the seeds' endosperm resources can affect this ability (Maun, 1994; Muller-Landau

254   *et al.*, 2002; Moles & Westoby, 2004). Another imperative aptitude for establishment of some plants is the

255   development of mycorrhizae (Akhmetzhanova *et al.*, 2012; Gijbels, Adriaens & Honnay, 2012; Hempel *et*

256   *al.*, 2013). Additionally, mycorrhizae have been shown to be important not only in the establishment phase

257   but also for the persistence of plant species and consequently for the local plant community composition

258   (Hartnett & Wilson, 1999). Here we demonstrated that plants depending on mycorrhiza and plants that

259   require a high degree of mycorrhizal infection had higher dark diversity likelihood. This result is not

260   surprising given the importance of mycorrhiza for plant establishment and persistence.

261   Competition for resources is an important phenomenon shaping the local structure, composition and

262   richness of plant communities throughout the world (Tilman, 1994; Tilman, 1997; McKane *et al.*, 2002;

263   Moeslund *et al.*, 2013). In the current study, we showed that plants with high dark diversity likelihood

264   preferred low levels of nutrients. In most of Northern Europe and particularly in Denmark the landscape is

265   relatively nutrient rich. In such a setting, plant species like *Urtica dioica* L., *Epilobium hirsutum* L. and

266   *Cirsium arvense* (L.) Scop. adapted to benefit from high nutrient availability will be strong competitors

267   (Grime, 1979; Hill *et al.*, 1999; Ellenberg *et al.*, 2001) and therefore have lower dark diversity likelihood.

268   Plants with competitive advantages (e.g., regarding resources) will tend to appear more often in suitable

269   habitats than species being less strong competitors in these habitats.

270   In our study demonstrated that shade tolerant plants were more likely to be part of the dark

271   diversity. The possible explanation for this finding is probably multifaceted. Plants that are more or less

272   shade tolerant may have competitive disadvantages in open landscapes (recall that this study concerns only

273   open habitats). Also, the fact that the landscape was prehistorically (around 5000 BC) almost completely

274   forested (Fritzbøger & Odgaard, 2010) can explain this observation. Indeed, shade tolerant species likely to

12

275   occur dormantly in the soil seed bank may have been present and more frequent across Denmark during

276   the past, when forests were more dominant, and may return (cf. memory effect due to land-use legacy) if

277   land-use returns to forest habitats. A third explanation could be that since forests are not included in this

278   study some shade tolerant species may have higher dark diversity likelihood simply because they prefer

279   forest environments. While there are indeed examples of such "forest-species" that occasionally occur in

280   open habitats in our dataset (e.g. *Galium odoratum* (L.) Scop. and *Allium ursinum* L.), there are also several

281   examples of shade tolerant species often belonging to the dark diversity albeit in reality observed almost as

282   often in open landscapes as in forests: e.g. *Maianthemum bifolium* (L.) F.W.Schmidt and *Trientalis europaea*

283   L. (EIV_L = 3 and 5 respectively, both widespread in grasslands, meadows, heathlands, forests etc.) (Mossberg

284   & Stenberg, 2005).

285   In many environments, stress tolerance is a key factor shaping local plant diversity (Osmond *et al.*,

286   1987; Maun, 1994; Ejrnæs & Bruun, 2000; Moeslund *et al.*, 2011). Recently, researchers showed that

287   stress-tolerance was among the most important determinants of plant dark diversity in North-Eastern

288   European dry calcareous grasslands (Riibak *et al.*, 2015). They suggested that in the driest grasslands stress-

289   tolerant species are more likely to thrive and therefore less likely to be part of the dark diversity. For our

290   study this explanation could also be true. Many of the habitats we included are stressful environments

291   either because they are relatively dry (grasslands, heathlands), waterlogged (fens, bogs, meadows) or saline

292   (salt meadows). However, another equally plausible explanation could be that ruderal species occurred

293   more often in the dark diversity than non-ruderals and the plants' ruderality was significantly negatively

294   related to the plants' stress-tolerance. From the fact that highly ruderal species cannot be highly stress-

295   tolerant at the same time (Grime, 1979) it follows that the likelihood of ending up in the dark diversity

296   could actually be related to ruderality, not stress-tolerance.

297   Dispersal

13

298 We showed that species with higher dark diversity likelihood were also generally poorer dispersers. This

299 result was supported by Riibak (2015) for dry calcareous grasslands. It is intuitively meaningful and a well-

300 known fact that species with poor dispersal abilities also have a lower probability of dispersal to new

301 suitable sites and to recolonize sites where these species have earlier gone extinct (Tilman, 1997; Cain,

302 Milligan & Strand, 2000; Myers & Harms, 2009; Torrez *et al.*, 2016). This explanation is corroborated by the

303 result that species with a higher seed mass was also more often part of the dark diversity. Heavy seeds are

304 more unlikely to spread long distances and thereby reach suitable habitats (Marteinsdóttir & Eriksson,

305 2013). Also, heavy-seed species typically produce less seeds lowering the probability that one of them will

306 eventually arrive at a suitable site (Cornelissen *et al.*, 2003; Riibak *et al.*, 2015).

307 **Explaining dark diversity likelihood**

308 As expected – given their importance for plant species establishment, dispersal or persistence – the factors

309 included here were able to significantly explain plant species' dark diversity likelihood. The goodness of fit

310 (D^2_{adj}) we obtained implies that even though most of the explanatory factors were important for plants'

311 dark diversity likelihood other factors not tested here may also be involved. Site conditions such as habitat

312 fragmentation and reduced habitat patch sizes may be important issues that can explain why many species

313 are missing in suitable places (Fahrig, 2003). However, the plants' ability to survive in a fragmented small-

314 habitat-patch landscape is not trivial to measure and attempts to directly relate plant traits to this ability

315 are therefore rare (Dupré & Ehrlén, 2002; May *et al.*, 2013). We attempted to account for this by including

316 dispersal distance and seed mass, as suggested by May *et al.* (2013) and Grime's plant strategies, which

317 relate to seed number, life span, seed bank strategy and pollination strategy (Grime, 1979; Rees, 1994; Šerá

318 & Šerý, 2004; Pierce *et al.*, 2014), which again relate to species' probability to go extinct in a suitable

319 habitat patch (Eriksson, 1996; Dupré & Ehrlén, 2002; Gijbels, Adriaens & Honnay, 2012). On the other hand,

320 researchers have recently found that while seed production traits are integral to Grime's plant strategies

321 this relationship is not straight-forward (Pierce *et al.*, 2014). This might also be true for relationships

322   between Grime's plant strategies and factors such as life span and pollination strategy and therefore future

323   studies may obtain a higher goodness of fit by analysing the individual traits instead of proxies for these.

324   Furthermore, phenotypic plasticity and general susceptibility to pathogens are likely to play a role for the

325   persistence of plants (Augspurger & Kelly, 1984; Burdon, Thrall & Ericson, 2006; Reed *et al.*, 2010).

326   **Management implications**

327   In the EU, part of the biodiversity strategy is to halt biodiversity loss by 2020. This goal requires both

328   researchers and managers to make the most of the limited funds available. Restoring or conserving

329   ecosystems or habitats is often focused on re-establishing or improving environmental conditions and then

330   hoping that biodiversity will respond positively. Tools that improve our understanding of the causal

331   mechanisms behind which species remain absent and why despite seemingly suitable conditions is

332   extremely useful and could be highly beneficial for restoration and conservation efforts.

333   For conservation and restoration our findings underpin the importance of assessing the funga's

334   ability to sustain the flora of concern at a given site. Inoculation with certain fungi could aid to restore plant

335   communities (Torrez *et al.*, 2016). However, great care needs to be taken during this proces with the best

336   results probably obtained by adding a diverse and locally adapted mycorrhizal community (Klironomos,

337   2003).

338   The fact that dispersal limitation is an important factor for species' dark diversity likelihood strongly

339   suggests that space and time are imperative factors in the planning and management of nature. Given

340   enough time and suitable corridors for dispersal, even poor dispersers will eventually reach suitable but

341   distant habitats. This finding also highlights the need to accommodate this by considering assisted

342   migration (Seddon, 2010).

15

343   Also, our study suggest focusing on creating opportunities for ruderal species in restoration and

344   conservation projects; e.g. by making sure that bare soil for seed germination is made available through

345   disturbances such as erosion, fire, trampling by large herbivores etc.

346   Finally, our results strongly emphasize the importance of focusing on nutrients and light availability in

347   conservation and restoration. Failing to ensure nutrient poor sites in regions with heavy nutrient loads will

348   render a number of species unable to thrive in otherwise suitable areas. Creating more shade in open

349   habitats could potentially aid low-light-adapted species in reappearing from the dark diversity pool.

Acknowledgements

Data accessibility

353   - Species geography: www.naturdata.dk

354   - Species traits: LEDA, BiolFlor and MycoFlor databases, Akhmetzhanova *et al.* (2012), TableS1

References

356   Akhmetzhanova, Soudzilovskaia, Onipchenko, Cornwell, Agafonov, Selivanov & Cornelissen. (2012) A
357   rediscovered treasure: mycorrhizal intensity database for 3000 vascular plant species across the former
358   Soviet Union. Ecology, **93**, 689–690.

359   Andersen, D.K., Nygaard, B., Fredshavn, J.R. & Ejrnæs, R. (2013) Cost-effective assessment of conservation
360   status of fens. Applied Vegetation Science, **16**, 491–501.

361   Augspurger, C.K. & Kelly, C.K. (1984) Pathogen mortality of tropical tree seedlings: experimental studies of
362   the effects of dispersal distance, seedling density, and light conditions. Oecologia, **61**, 211–217.

363   Beals, E.W. (1984) Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data.
364   Advances in Ecological Research, **14**, 1–55.

365   Borcard, D. & Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal
366   coordinates of neighbour matrices. Ecological Modelling, **153**, 51–68.

367 Buchwald, E., Wind, P., Bruun, H.H., Møller, P.F., Ejrnæs, R. & Svart, H.E. (2013) Hvilke planter er
368 hjemmehørende i Danmark? Flora & Fauna, **118**, 73–96.

369 Burdon, J.J., Thrall, P.H. & Ericson, a.L. (2006) The current and future dynamics of disease in plant
370 communities. Annual Review of Phytopathology, **44**, 19–39.

371 Burnham, K.P. & Anderson, D.R. (2002) Model selection and multimodel inference: a practical information-
372 theoretic approach, Springer, New York.

373 Cain, M.L., Milligan, B.G. & Strand, A.E. (2000) Long-distance seed dispersal in plant populations. American
374 Journal of Botany, **87**, 1217–1227.

375 Cornelissen, J.H.C., Lavorel, S., Garnier, E., Diaz, S., Buchmann, N., Gurvich, D.E., Reich, P.B., ter Steege, H.,
376 Morgan, H.D., van der Heijden, M G A, Pausas, J.G. & Poorter, H. (2003) A handbook of protocols for
377 standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany, **51**,
378 335–380.

379 Diekmann, M. (2003) Species indicator values as an important tool in applied plant ecology – a review. Basic
380 and Applied Ecology, **4**, 493–506.

381 Diekmann, M. & Falkengren-Grerup, U. (1998) A new species index for forest vascular plants: development
382 of functional indices based on mineralization rates of various forms of soil nitrogen. Journal of Ecology, **86**,
383 269–283.

384 Diniz-Filho, J.A.F. & Bini, L.M. (2005) Modelling geographical patterns in species richness using eigenvector-
385 based spatial filters. Global Ecology and Biogeography, **14**, 177–185.

386 Dupré, C. & Ehrlén, J. (2002) Habitat configuration, species traits and plant distributions. Journal of Ecology,
387 **90**, 796–805.

388 Durka & Michalski. (2012) Daphne: a dated phylogeny of a large European flora for phylogenetically
389 informed ecological analyses. Ecology, **93**, 2297–2297.

390 Ejrnæs, R. & Bruun, H.H. (2000) Gradient analysis of dry grassland vegetation in Denmark. Journal of
391 Vegetation Science, **11**, 573–584.

392 Ellenberg, H., Weber, H.E., Düll, R., Wirth, V. & Werner, W. (2001) Zeigerwerte von planzen in Mitteleuropa,
393 3rd edn. Erich Goltze GmbH & Co KG, Göttingen.

394 Eriksson, O. (1996) Regional dynamics of plants: a review of evidence for remnant, source-sink and
395 metapopulations. Oikos, **77**, 248–258.

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

398 Fahrig, L. (2003) Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and
399 Systematics, **34**, 487–515.

400 Frederiksen, S., Rasmussen, F.N. & Seberg, O. (2006) Dansk Flora, 1st edn. Gyldendal, Copenhagen.

401 Fredshavn, J.R., Nygaard, B. & Ejrnæs, R. (2009) Naturtilstand på terrestriske naturarealer - besigtigelser af
402 §3-arealer.

403 Fritzbøger, B. & Odgaard, B. (2010) Skovenes historie. Naturen i Danmark - Skovene (ed P.F. Møller),
404 Gyldendal, Copenhagen.

405 Gijbels, P., Adriaens, D. & Honnay, O. (2012) An orchid colonization credit in restored calcareous grasslands.
406 Ecoscience, **19**, 21–28.

407 Gittleman, J.L. & Kot, M. (1990) Adaptation: statistics and a null model for estimating phylogenetic effects.
408 Systematic Zoology, **39**, 227–241.

409 Grime, J.P. (1979) Plant strategies and vegetation processes, 1st edn. John Wiley & Sons, Chichester, New
410 York, Brisbane, Toronto.

411 Hansen, K. (1996) Dansk Feltflora, 1st edn. Gyldendal, Copenhagen.

412 Hartnett, D.C. & Wilson, G.W.T. (1999) Mycorrhizae Influence Plant Community Structure and Diversity in
413 Tallgrass Prairie. Ecology, **80**, 1187–1195.

414 Hempel, Götzenberger, Kühn, Michalski, Rillig, Zobel & Moora. (2013) Mycorrhizas in the Central European
415 flora: relationships with plant life history traits and ecology. Ecology, **94**, 1389–1399.

416 Hill, M.O., Mountford, J.O., Roy, D.B. & Bunce, R.G.H. (1999) Ellenberg's indicator values for British plants,
417 Natural Environment Research Council, UK.

418 Hubbell, S.P. (2001) The unified neutral theory of biodiversity and biogeography, 1st edn. Princeton
419 University Press, Princeton and Oxford.

420 Kleyer, M., Bekker, R.M., Knevel, I.C., Bakker, J.P., Thompson, K., Sonnenschein, M., Poschlod, P., Van
421 Groenendael, J.M., Klimeš, L., Klimešová, J., Klotz, S., Rusch, G.M., Hermy, M., Adriaens, D., Boedeltje, G.,
422 Bossuyt, B., Dannemann, A., Endels, P., Götzenberger, L., Hodgson, J.G., Jackel, A., Kühn, I., Kunzmann, D.,
423 Ozinga, W.A., Römermann, C., Stadler, M., Schlegelmilch, J., Steendam, H.J., Tackenberg, O., Wilmann, B.,
424 Cornelissen, J.H.C., Eriksson, O., Garnier, E. & Peco, B. (2008) The LEDA Traitbase: a database of life-history
425 traits of the Northwest European flora. Journal of Ecology, **96**, 1266–1274.

426 Klironomos, J.N. (2003) Variation in plant response to native and exotic arbuscular mycorrhizal fungi.
427 Ecology, **84**, 2292–2301.

428 Kochjarová, J. (2006) Contribution to the occurrence and former distribution of *Tephroseris palustris*
429 (Compositae) in the Central Europe. Biologia, **61**, 361–364.

430 Kühn, I., Durka, W. & Klotz, S. (2004) BiolFlor — a new plant-trait database as a tool for plant invasion
431 ecology. Diversity and Distributions, **10**, 363–365.

432 Lenoir, J., Gégout, J., Guisan, A., Vittoz, P., Wohlgemuth, T., Zimmermann, N.E., Dullinger, S., Pauli, H.,
433 Willner, W., Grytnes, J., Virtanen, R. & Svenning, J. (2010) Cross-scale analysis of the region effect on
434 vascular plant species diversity in Southern and Northern European mountain ranges. PLoS ONE, **5**, e15734.

435 Lewis, R.J., Szava-Kovats, R. & Pärtel, M. (2016) Estimating dark diversity and species pools: an empirical
436 assessment of two methods. Methods in Ecology and Evolution, **7**, 104–113.

437 Marteinsdóttir, B. & Eriksson, O. (2013) Trait-based filtering from the regional species pool into local
438 grassland communities. Journal of Plant Ecology, **7**, 347–355.

439 Maun, M.A. (1994) Adaptations enhancing survival and establishment of seedlings on coastal dune systems.
440 Plant Ecology, **111**, 59–70.

441 May, F., Giladi, I., Ristow, M., Ziv, Y. & Jeltsch, F. (2013) Plant functional traits and community assembly
442 along interacting gradients of productivity and fragmentation. Perspectives in Plant Ecology, Evolution and
443 Systematics, **15**, 304–318.

444 McArthur, R.H. & Wilson, E.O. (1967) The theory of island biogeography, Princeton University Press,
445 Princeton, NJ.

446 McCune, B. (1994) Improving community analysis with the Beals smoothing function. Ecoscience, **1**, 82–86.

447 McKane, R.B., Johnson, L.C., Shaver, G.R., Nadelhoffer, K.J., Rastetter, E.B., Fry, B., Giblin, A.E., Kielland, K.,
448 Kwiatkowski, B.L., Laundre, J.A. & Murray, G. (2002) Resource-based niches provide a basis for plant species
449 diversity and dominance in arctic tundra. Nature, **415**, 68–71.

450 Milberg, P., Bergstedt, J., Fridman, J., Odell, G. & Westerberg, L. (2008) Observer bias and random variation
451 in vegetation monitoring data. Journal of Vegetation Science, **19**, 633–644.

452 Moeslund, J.E., Arge, L., Bøcher, P., Nygaard, B. & Svenning, J. (2011) Geographically comprehensive
453 assessment of salt-meadow vegetation-elevation relations using LiDAR. Wetlands, **31**, 471–482.

454 Moeslund, J.E., Arge, L., Bøcher, P.,K., Dalgaard, T., Odgaard, M.V., Nygaard, B. & Svenning, J. (2013)
455 Topographically controlled soil moisture is the primary driver of local vegetation patterns across a lowland
456 region. Ecosphere, **4**, art91.

457 Mokany, K. & Paini, D.R. (2011) Dark diversity: adding the grey. Trends in Ecology & Evolution, **26**, 264–265.

458 Moles, A.T. & Westoby, M. (2004) Seedling survival and seed size: a synthesis of the literature. Journal of
459 Ecology, **92**, 372–383.

460 Mossberg, B. & Stenberg, L. (2005) Den nye nordiske flora, 1st edn. Gyldendal, Copenhagen, Denmark.

461 Muller-Landau, H.C., Wright, S.J., Calderón, O., Hubbell, S.P. & Foster, R.B. (2002) Assessing recruitment
462 limitation: concepts, methods and case-studies from a tropical forest. Seed dispersal and frugivory: ecology,
463 evolution and conservation (eds D.J. Levey, W.R. Silva & M. Galetti), pp. 35–53. CABI Publishing,
464 Wallingford, UK.

465 Münzbergová, Z. & Herben, T. (2004) Identification of suitable unoccupied habitats in metapopulation
466 studies using co-occurrence of species. Oikos, **105**, 408–414.

467 Myers, J.A. & Harms, K.E. (2009) Seed arrival, ecological filters, and plant species richness: a meta-analysis.
468 Ecology Letters, **12**, 1250–1260.

469 Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P.,
470 Stevens, M.H.H. & Wagner, H. (2015) Vegan: Community ecology package. R package version 2.3-2.

471 Olsson, K. & Tyler, T. (2001) A historic perspective on the red-listed species of Scania, S. Sweden. Botaniska
472 Notiser, **134**, 1–5.

473 Osmond, C.B., Austin, M.P., Berry, J.A., Billings, W.D., Boyer, J.S., Dacey, J.W.H., Nobel, P.S., Smith, S.D. &
474 Winner, W.E. (1987) Stress physiology and the distribution of plants. Bioscience, **37**, 48.

475 Paradis, Claude & Strimmer. (2004) APE: Analyses of phylogenetics and evolution in R language.
476 Bioinformatics, **20**, 289–290.

477 Paradis, E. (2015) Moran's autocorrelation coefficient in comparative methods. .

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

480 Pärtel, M., Zobel, M., Zobel, K. & van der Maarel, E. (1996) The species pool and its relation to species
481 richness: evidence from Estonian plant communities. Oikos, **75**, 111–117.

482 Pierce, S., Bottinelli, A., Bassani, I., Ceriani, R.M. & Cerabolini, B.E.L. (2014) How well do seed production
483 traits correlate with leaf traits, whole-plant traits and plant ecological strategies? Plant Ecology, **215**, 1351–
484 1359.

485 Quinn, G.P. & Keough, M.J. (2002) Experimental design and data analysis for biologists, 1st edn. Cambridge
486 University Press, Cambridge.

487 R Core Team. (2015) R: A language and environment for statistical computing.

488 Reed, T.E., Waples, R.S., Schindler, D.E., Hard, J.J. & Kinnison, M.T. (2010) Phenotypic plasticity and
489 population viability: the importance of environmental predictability. Proceedings of the Royal Society of
490 London B: Biological Sciences, **277**, 3391–3400.

491 Rees, M. (1994) Delayed germination of seeds: a look at the effects of adult longevity, the timing of
492 reproduction, and population age/stage structure. The American Naturalist, **144**, 43–64.

493 Riibak, K., Reitalu, T., Tamme, R., Helm, A., Gerhold, P., Znamenskiy, S., Bengtsson, K., Rosén, E., Prentice,
494 H.C. & Pärtel, M. (2015) Dark diversity in dry calcareous grasslands is determined by dispersal ability and
495 stress-tolerance. Ecography, **38**, 713–721.

496 Ronk, A., Szava-Kovats, R. & Pärtel, M. (2015) Applying the dark diversity concept to plants at the European
497 scale. Ecography, **38**, 1015–1025.

498 Seddon, P.J. (2010) From reintroduction to assisted colonization: moving along the conservation
499 translocation spectrum. Restoration Ecology, **18**, 796–802.

500 Seidling. (2005) Ground floor vegetation assessment within the intensive (Level II) monitoring of forest
501 ecosystems in Germany: chances and challenges. European Journal of Forest Research, **124**, 301–312.

502 Šerá, B. & Šerý, M. (2004) Number and weight of seeds and reproductive strategies of herbaceous plants.
503 Folia Geobotanica, **39**, 27–40.

504 Tamme, R., Götzenberger, L., Zobel, M., Bullock, J.M., Hooftman, D.A.P., Kaasik, A. & Pärtel, M. (2014)
505 Predicting species' maximum dispersal distances from simple plant traits. Ecology, **95**, 505–513.

506 Tilman, D. (1997) Community invasibility, recruitment limitation, and grassland biodiversity. Ecology, **78**,
507 81-92.

508 Tilman, D. (1994) Competition and biodiversity in spatially structured habitats. Ecology, **75**, 2–16.

509 Torrez, V., Ceulemans, T., Mergeay, J., de Meester, L. & Honnay, O. (2016) Effects of adding an arbuscular
510 mycorrhizal fungi inoculum and of distance to donor sites on plant species recolonization following topsoil
511 removal. Applied Vegetation Science, **19**, 7–19.

512 Yoshioka, A., Miyazaki, Y., Sekizaki, Y., Suda, S., Kadoya, T. & Washitani, I. (2014) A "lost biodiversity"
513 approach to revealing major anthropogenic threats to regional freshwater ecosystems. Ecological
514 Indicators, **36**, 348-355.

515

516

517 **Table 1.** An overview of the included explanatory variables with a brief description of each variable. For

518 abbreviations see Table 2

| Explanatory variable	Short description	Variable nature	Relevant literature
EIV_L	Plants' preference for light level	Ordinal (1–9)	(Ellenberg et al. 2001)
EIV_F	Plants' preference for soil moisture	Ordinal (1–12)	(Ellenberg et al. 2001)
$EIV_{N/R}$	Soil nutrients and pH is usually positively correlated. To take this into account this ratio is used to represent nutrient enrichment	Continuous (0.13 – 2.00)	(Ellenberg et al. 2001, Andersen et al. 2013)
Grime C	Plants' competitive ability	Ordinal (1–12)	(Grime 1979)
Grime S	Plants' stress tolerance (e.g., tolerance towards grazing, cutting, drought, soil salinity etc.)	Ordinal (1–12)	(Grime 1979)
Grime R	Plants' ruderality (tendency to occur in ruderal areas)	Ordinal (1–12)	(Grime 1979)
Mean mycorrhiza infection %	The mean infection percentage of the plants' roots by mycorrhiza	Continuous (1– 100 %)	(Akhmetzhanova et al. 2012)
Mycorrhizal dependence	Whether (1) or not (0) the plants are obligately mycorrhizal. Binary variable	Nominal (0 vs. 1)	(Hempel et al. 2013)
Log(Seed mass)	The logarithm to the mean seed mass	Continuous (-2.52 – 2.13)	(Kühn et al. 2004, Riibak et al. 2014)
Log(Max. dispersal distance)	The logarithm to the maximum dispersal distance based on plant traits and taxonomy	Continuous (-1.13 – 10.72)	(Tamme et al. 2014, Riibak et al. 2014)

519

520

22

521 **Table 2.** Standardised coefficients (their rank based on the numeric size of the standard coefficient is given

522 in parentheses) of each of the 12 candidate models relating the probability for plant species of being part of

523 the dark diversity to the explanatory factors used in this study. Results are presented both for the high-

524 sample size models, medium sample size models and the low sample-size models (See details in Table S2).

525 Grey cells mark explanatory variables that were not included in the model. For each model the deviance

526 explained – $D^2_{adjusted}$ (D^2_{adj}) – is shown together with the D^2_{adj} for models after adding the necessary number

527 of filters (also shown) to remove phylogenetic autocorrelation (see methods). For each explanatory variable

528 the absolute (abs.) mean standardised coefficient (coef.) is given to represent its overall relative

529 importance. Abbreviations: C: competition, S: stress, R: ruderality, Disp.: dispersal, Dist.: distance, SM: seed

530 mass, EIV: Ellenberg Indicator for light (L), moisture (F), nutrients (N) and reaction (R), M.: mycorrhiza

531

Explanatory variable	High sample size models (n = 564)			Medium sample size models (n = 457)			Low sample size models (n = 151)						Mean abs. coef.
	Grime C	Grime S	Grime R	Grime C + D	Grime S + D	Grime R + D	Grime C + D + M	Grime S + D + M	Grime R + D + M	Grime C + SM + M	Grime S + SM + M	Grime R + SM + M	
EIV_L	-.23***	-.24***	-.29***	-.23***	-.24***	-.28***	-.45***	-.51***	-.53***	-.45***	-.54***	-.54***	.38
EIV_F	.05***	-.01***	.05***	.06***	.01*	.07***	-.01	.01***	.05***	-.00	-.00	.06***	.03
$EIV_{N/R}$	-.03***	-.11***	-.20***	-.19***	-.26***	-.35***	-.33***	-.46***	-.39***	-.34***	-.49***	-.41***	.30
Grime C	-.22***			-.20***			-.02***			-.02***			.12
Grime S		.02***			.01***			-.19***			-.21***		.11
Grime R			.22***			.21***			.17***			.18***	.20
Log(MDD)				-.21***	-.25***	-.23***	.02***	-.05***	-.00				.13
Log(SM)										.11***	.12***	.14***	.12
Mean M inf. %							.15***	.16***	.14***	.12***	.12***	.10***	.13
M dependence							.34***	.41***	.42***	.35***	.40***	.42***	.39
D^2_{adj}	.04	.01	.04	.08	.06	.09	.11	.12	.13	.11	.13	.14	
D^2_{adj} (w/ filters)				.10	.08	.10	.11			.12	.14		

| No. of filters | | 8 | 8 | 7 | 1 | | 1 | 1 | | |

532   P: < 0.05: *, <0.01: **, <0.001***

533

24
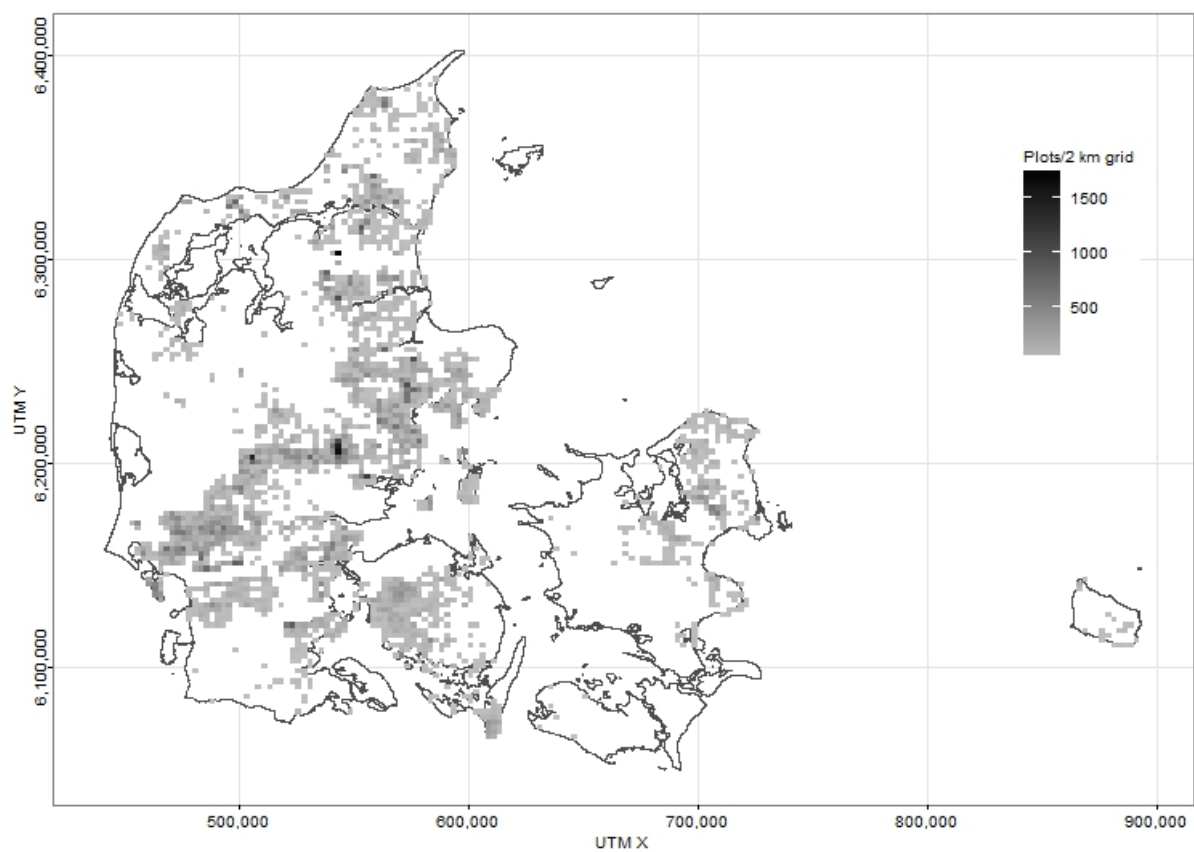
534

535 **Figure 1.** The density of the 15.160 vegetation plots used for this study
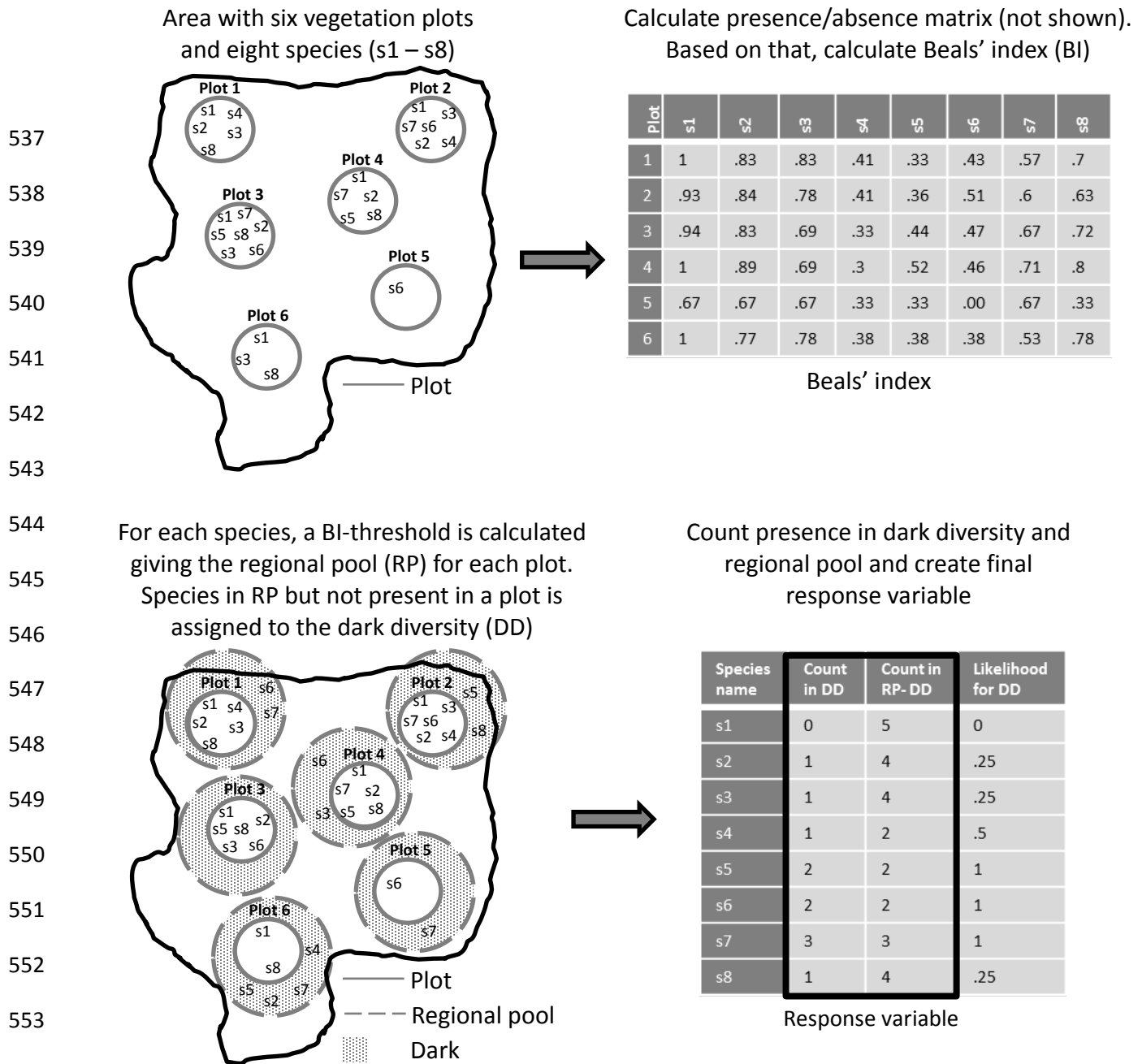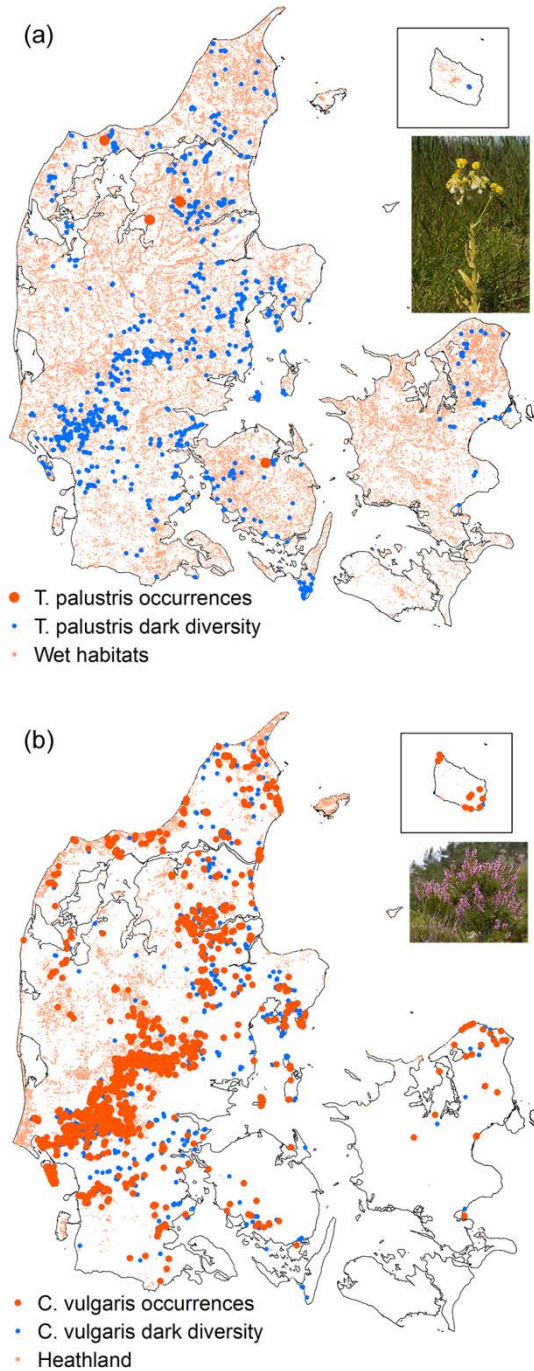
536

Figure 2. The steps taken from vegetation plot data to response variable. A fictive area with eight fictive

species is used for illustration purpose

Figure 3. The distribution of occurrences and dark diversity locations for (a) a species often found in dark diversity (*Tephroseris palustris*) and (b) a species relatively rarely found in the dark diversity in this study (*Calluna vulgaris*). For reference, known wet habitats and heathlands are shown. Map inserts: island of Bornholm (Fig. 1)

563
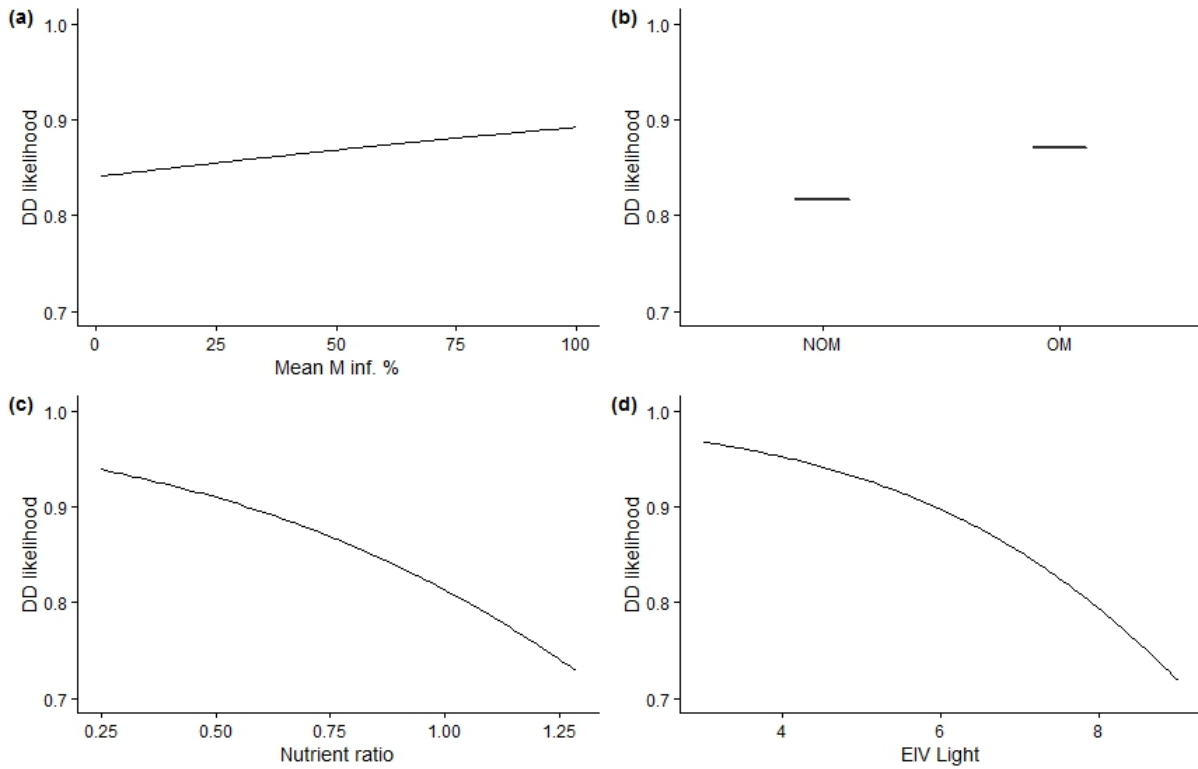
Figure 4. Predictions of the relationship between the dark diversity likelihood of a plant and (a) mean

mycorrhizal infection percentage, (b) mycorrhizal dependence, (c) nutrient/reaction-ratio (nutrient ratio)

based on Ellenberg Indicator Values (EIV) and (d) EIV for light. The predictions shown are based on the

Grime R + SM + M model having the highest goodness of fit (Table 2). Predictive variables not in question

were held constant during the predictions shown. DD: dark diversity, NOM: not obligately mycorrhizal, OM:

obligately mycorrhizal, M: mycorrhizal, inf.: infection