

Using dark diversity and plant characteristics to guide conservation and restoration

Jesper Erenskjold Moeslund^{*,1} , Ane Kirstine Brunbjerg^{1,2} , Kevin Kuhlmann Clausen³ , Lars Dalby¹ , Camilla Fløjgaard¹ , Anders Juel⁴ and Jonathan Lenoir⁵ 

¹Section for Biodiversity & Conservation, Department of Bioscience, Aarhus University, Grenåvej 14, 8410 Kalø, Denmark; ²School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK; ³Section for Wildlife Ecology, Department of Bioscience, Aarhus University, Kalø, Denmark; ⁴Section for Nature Conservation, The Nature Agency, Ministry of Environment and Food, Copenhagen, Denmark; and ⁵UR “Ecologie et dynamique des systèmes anthropisés” (EDYSAN, FRE 3498 CNRS-UPJV), Université de Picardie Jules Verne, Amiens, France

Summary

1. Dark diversity is a promising concept for prioritising management efforts as it focuses on species that are present in the regional pool, but locally absent even though environmental requirements are met. Currently, we lack knowledge of what characterises species belonging to the dark diversity more often than others, although this is important knowledge for restoration and conservation actions.

2. We applied the concept to a massive national (Danish) plant diversity data base, containing 236 923 records from 15 160 surveys involving 564 species. This enabled the first geographically comprehensive (43 000 km²) assessment of dark diversity, at a spatial resolution relevant for conservation and restoration planning (78 m²) across multiple terrestrial habitats, thereby maximising the practical applications of this concept. The probability for a given plant species to belong to the dark diversity was computed and logistically regressed against variables representing its ecological preferences (e.g. nutrient availability), strategies (competitor, stress tolerant, ruderal), mycorrhizal relationships, establishment capacities (seed mass) and dispersal abilities.

3. Forty-six percent of the species had a high probability (>95%) of being part of dark diversity, whereas for 7% of the species this probability was less than 60%.

4. Typical dark diversity plant species tended to depend on mycorrhiza, were mostly adapted to low light and low nutrient levels, had poor dispersal abilities and were ruderals and stress intolerant.

5. *Synthesis and applications.* Characterising species that are more often absent from suitable sites than others (dark diversity species) has important implications for the planning and management of natural ecosystems. From our study, practitioners gain insight into the factors triggering the absence of individual plant species in a seemingly suitable habitat. We highlight the need to carefully consider mycorrhizal inoculations with a suitable assemblage of fungi to promote the establishment success of dark diversity plants. Additionally, time-lags in plant species dispersal and establishment as well as spatial connectivity in fragmented habitats are central to consider in nature management although assisted migration might also aid poor dispersers. Finally, nutrient-poor localities are probably important ‘islets’ allowing nitrophobic dark diversity plant species to thrive within agricultural landscapes that are generally nutrient-rich.

Key-words: assisted migration, dispersal distance, Ellenberg indicator values, Grime strategies, mycorrhiza, phylogenetic autocorrelation, plant diversity, plant traits, regional species pool, vegetation ecology

*Correspondence author. E-mail: jesper.moeslund@bios.au.dk

Introduction

Recently, Pärtel, Szava-Kovats & Zobel (2011) presented a new concept coined *dark diversity*, which could prove to be a central idea in the development of effective tools for practical biodiversity management (Lewis *et al.* 2017). Dark diversity encompasses the diversity expressed by all species missing locally, even though biogeographic history and current ecological and environmental conditions suggest their presence (Pärtel, Szava-Kovats & Zobel 2011). In a recent essay, Lewis *et al.* (2017) highlighted the importance – for practical conservation and restoration – of identifying the mechanisms determining why some species are more often part of dark diversity than others. Our study is an attempt to address exactly this, by searching across multiple sites and habitats for common characteristics among typical dark diversity plant species.

To our knowledge, the probability of individual species to be part of dark diversity (referred to as plant species' *dark diversity probability* in the remaining text) has never been studied and little is known about the characteristics of species having a higher dark diversity probability than others (but see Riibak *et al.* 2015). Determining these common characteristics may shed light on the suite of mechanisms that matter for plant recruitment and persistence (seed dispersal, seed germination, seedling emergence, establishment, metapopulation and metacommunity dynamics, plant interactions etc.). Consequently, this could prove to be an important key for successful practical application of dark diversity in restoration and conservation actions (Lewis *et al.* 2017). A number of functional traits and other ecological adaptations might be relevant in this context. For instance, plants adapted to nutrient-poor and well-illuminated conditions are expected to belong to dark diversity more often in open and productive environments, since tall and fast-growing species suppressing nitrophobic and light-demanding species tend to dominate under such conditions (Brauer, Stomp & Huisman 2012). As a consequence, competitive plant species (cf. Grime 1979) are likely to have a lower dark diversity probability in these habitat types. Related to this, species with poor dispersal or establishment abilities suffer from a lower recruitment at suitable sites and would therefore be expected to have a higher dark diversity probability than others (Moles & Westoby 2004; Riibak *et al.*, 2015). Finally, plants' mycorrhizal relationships are likely to impact their dark diversity probability: species being highly or obligately dependent on mycorrhiza could be anticipated to belong to dark diversity more often than others since mycorrhiza requires co-occurrence of a certain plant species/genus and a particular fungus species/genus (Finlay 2008) which is less likely than just one of them being present at a given site.

Hitherto, only a handful of studies have applied the dark diversity concept to restoration and conservation issues (Gijbels, Adriaens & Honnay 2012; Yoshioka *et al.* 2014; Riibak *et al.* 2015; Ronk, Szava-Kovats & Pärtel 2015). While these certainly have been successful, they were limited

in spatial extent, or scope of species or habitats preventing a more general application for management purposes. To apply the dark diversity concept more generally in conservation and restoration, co-occurrence plant data – collected systematically at a fine spatial resolution and covering both multiple habitats and large (regional-to-national) areas – are needed. While such data sets are rare, these criteria are more likely to be fulfilled by national biodiversity inventories (e.g. Fredshavn, Nygaard & Ejrnæs 2009).

Here, we present the first national assessment of the characteristics of typical dark diversity plant species at a spatial resolution relevant to conservation and restoration planning and management, covering multiple open terrestrial habitats. To address the hypotheses described above and summarised in Table 1, we used a large national (Danish) plant dataset with high spatial accuracy to calculate dark diversity probabilities of >550 species, and related these probabilities to traits and characteristics of the individual species. Our specific study questions were: (i) what are the relationships between North-European plant species' dark diversity probability and their functional traits and ecological characteristics and (ii) which traits or characteristics are the best to explain the dark diversity probability? Furthermore, we discuss the causal mechanisms most likely determining the plant's dark diversity probability and how our findings may aid effective planning and management initiatives and promote the practical application of the dark diversity concept in conservation and restoration projects.

Materials and methods

VEGETATION DATA

Data on the distribution of vascular plants in Denmark were obtained from the municipalities' vegetation inventory of natural habitat types (Fredshavn, Nygaard & Ejrnæs 2009). We used observations from circular plots (5 m radius) laid out to capture the typical flora of a particular site (typically one plot per site). The sites are 5 ha on average (range: 0.003–900 ha), distributed throughout most of Denmark (Fig. 1) and cover freshwater meadows, salt meadows, heathlands, bogs, moors, fens, grasslands and vegetated dunes (i.e. open habitats). The dataset was extracted on October 6 2014. We used data from 2004 to 2014. We only considered observations at the species level and excluded all shrubs, trees, submersed aquatic species and neophytes (i.e. species that are not considered a natural part of the vegetation given their history and dispersal ability, see appendix tables 6–8 in Buchwald *et al.* 2013). To ensure meaningful calculations of the regional species pool (see below) only plots with more than five plant species records were used. After this initial filtering, our dataset consisted of 236 923 records from 15 160 plots involving 564 plant species.

REGIONAL SPECIES POOL AND DARK DIVERSITY – RESPONSE VARIABLE

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

Table 1. An overview of the explanatory variables included in the study, with a brief description of each variable, the range of values, our hypotheses and relevant literature

Var.	Short description	Variable nature (range of values)	Plant species with a higher dark diversity probability are hypothesised to...	Relevant literature
L	Ellenberg Indicator Value (EIV) for light; plants' preference for light level	Ordinal (1 to 9)	...prefer well-illuminated conditions since tall and competitive species tend to overgrow small light demanding species in productive environments	Ellenberg <i>et al.</i> (2001)
F	EIV for soil moisture; plants' preference for soil moisture	Ordinal (1 to 12)	...prefer high soil moisture because many areas have been drained to enable extensive farming	Ellenberg <i>et al.</i> (2001)
N/R	Ratio between EIV for nitrogen and reaction; termed <i>nutrient ratio</i> in this study. Soil nutrients and pH are usually positively correlated. To take this into account this ratio is used to represent nutrient enrichment	Continuous (0.13 to 2.00)	...prefer nutrient-poor conditions since the regions generally suffer from high nutrient loads and it is consequently likely that nitrophobic species are out-competed	Ellenberg <i>et al.</i> (2001); Andersen <i>et al.</i> (2013)
GC	Grime C; plants' competitive ability	Ordinal (0 to 12)	...be less adapted to a competitive strategy. The competitive strategy is a clear advantage in the nutrient-rich conditions that prevail in anthropogenic regions	Grime (1979)
GS	Grime S; plants' stress tolerance (e.g. tolerance towards grazing, cutting, drought, soil salinity etc.)	Ordinal (0 to 12)	...be more adapted to a stress-tolerance strategy because stress factors like grazing is uncommon in modern landscapes	Grime (1979)
GR	Grime R; plants' ruderality (tendency to occur in ruderal areas)	Ordinal (0 to 12)	...be more ruderal as the disturbances allowing ruderal species to thrive is often absent in modern nature	Grime (1979)
MI	The mean infection percentage of the plants' roots by mycorrhiza	Continuous (1 to 100%)	...be more infected since species with high infection levels are more dependent on fungi so in case the right fungus is not present the species does not thrive	Akhmetzhanova <i>et al.</i> (2012)
MD	Whether or not the plants are obligately mycorrhizal	Binary (0 vs. 1)	...depend on mycorrhiza as in many cases the fungi may not be present at a site and therefore prevent dependent species to thrive	Hempel <i>et al.</i> (2013)
SM	The natural logarithm to the mean seed mass in milligrams	Continuous (-2.52 to 2.13)	...have heavy seeds, implying poor dispersal ability (see explanatory variable D below) or alternatively have light seeds implying poor establishment ability	Kühn, Durka & Klotz (2004); Riibak <i>et al.</i> (2015)
D	The natural logarithm to the potential maximum dispersal distance in meters based on plant traits and taxonomy	Continuous (-1.13 to 10.72)	...be poorer dispersers. They simply cannot reach as many suitable sites as good dispersers	Tamme <i>et al.</i> (2014); Riibak <i>et al.</i> (2015)

plot as recommended by Lewis, Szava-Kovats & Pärtel (2016). Beals' index represents the probability that a focal species will occur within a given plot based on the assemblage of co-occurring species (McCune 1994; Münzbergová & Herben 2004). Initially, a presence/absence matrix with all combinations of plot and species was constructed. Based on this matrix, we calculated Beals' index for each species in each plot excluding the focal species from the calculations as recommended by Oksanen *et al.* (2015) (Fig. 2). We used the 'beals' function in the 'vegan' package (Oksanen *et al.* 2015). The threshold for including a particular species in the regional species pool is species specific and was defined as the 5th percentile of the Beals' index value for the species following Gijbels, Adriaens & Honnay (2012) and Ronk, Szava-Kovats & Pärtel (2015).

Preceding the calculation of each threshold, the lowest Beals' index value among plots with occurrence of the species in question was identified, and all plots having values below that minimum were not considered. For every plot, the dark diversity was composed of all species in the regional pool excluding those that were actually present (Pärtel, Szava-Kovats & Zobel 2011) (Fig. 2).

As a response variable, we computed the species' probability to belong to dark diversity (P_{DD}) for each of the 564 plant species used in our analyses:

$$P_{DD} = \frac{\text{no. of times in dark diversity}}{\text{no. of times in regional pool}} \quad \text{eqn 1}$$

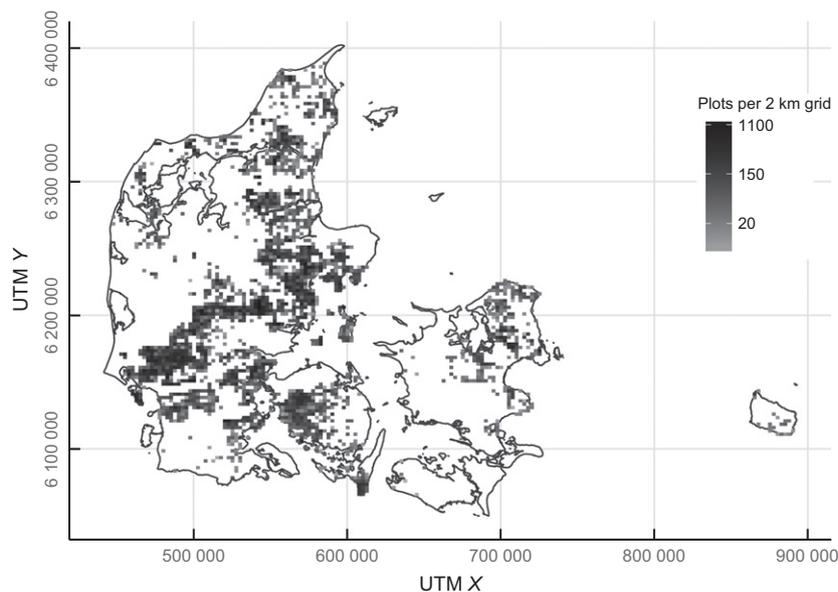


Fig. 1. Map of Denmark showing the density and distribution of the 15 160 vegetation plots used for this study.

SPECIES TRAITS AND CHARACTERISTICS – EXPLANATORY VARIABLES

Ellenberg's indicator values

We used Ellenberg's indicator values (EIVs) (Ellenberg *et al.* 2001) adjusted to British conditions (Hill *et al.* 1999) as these are thought to best match the Danish flora (Moeslund *et al.* 2013). EIVs represent European plant species' preferred position along various environmental gradients and are often used in vegetation studies (Ellenberg *et al.* 2001; Diekmann 2003; Lenoir *et al.* 2010; Moeslund *et al.* 2013). Because salinity is only relevant in coastal environments and variation in temperature and continentality in Denmark is negligible, the EIVs corresponding to these factors were left out. Consequently, we only considered EIVs for: (i) soil moisture (EIV_F), (ii) soil pH (EIV_R), (iii) soil nutrient status (EIV_N) and (iv) ambient light (EIV_L) (Table S1, Supporting Information). EIV_N and EIV_R are typically highly correlated (Diekmann & Falkengren-Grerup 1998; Seidling 2005), also in this study (Figure S1); hence we used a nutrient/pH-ratio based on the two corresponding EIVs to represent nutrient availability. This variable is referred to as the nutrient ratio below (e.g. Andersen *et al.* 2013).

CSR plant strategies

The main plant ecological strategies as proposed by Grime (1979) distinguish plants adapted to competitive (C-species), stressful (S-species) or ruderal (R-species) environments. Although plants can harbour any combination of these three strategies they are in their extreme forms mutually exclusive. For all plant species included here, we obtained the CSR strategy data from the BioFlor data base (Table S1, Kühn, Durka & Klotz 2004). Following Ejrnæs & Bruun (2000), we represented the degree to which a plant is adapted to a given strategy as values ranging from 0 to 12 for each of the three strategies, however restraining their sum to 12.

Mycorrhiza data

We used data on both mycorrhizal infection percentage (1–100%) and dependence (i.e. a factor variable with two levels: obligately

vs. non-obligately mycorrhizal). Data on mycorrhizal infection were retrieved from Akhmetzhanova *et al.* (2012) and data on mycorrhizal dependence were taken from MycoFlor (Hempel *et al.* 2013). These data were available for 33% and 82%, respectively, of the plant species involved in this study (Table S1).

Plant functional traits

Data on plant traits were extracted from LEDA (Kleyer *et al.* 2008) and baseflor (<http://philippe.julve.pagesperso-orange.fr/catminat.htm>): dispersal syndrome (DS), growth form (GF), releasing height (RH), seed mass (SM) and terminal velocity (TV). Where multiple records of a numeric trait were available for a species, the mean value was used. We consulted family, genus and species descriptions in Hansen (1996) to complete DS and GF data availability (5% and 22% of the data on DS and GF respectively came from this source). Unavailable species level DS data for *Equisetum pratense* L. caused us to use the genus level DS from baseflor (all species in this genus were unanimously recorded with the same DS). For SM, we further supplemented by data from BioFlor. Data on DS, GF, RH, SM and TV were available for 100%, 100%, 87%, 90% and 69% of the species, respectively. To calculate potential maximum dispersal distance (MDD) we used the 'dispeRsal' function which calculates MDD using plant traits and taxonomy (Tamme *et al.* 2014). DS and GF were translated to suit the dispeRsal function, however, with the following assumptions: hemerochor = animal, meteorochor = wind.special, hydrochor = wind.none, anemochor = wind.special and barochor = wind.none.

MDD was calculated for all species in the dataset following the hierarchy of best predictive performance given in Tamme *et al.* (2014), preferring simplicity over complexity in models of similar predictive performance. For species with multiple unique entries of a non-numeric trait (e.g. DS) we calculated the mean of the predicted MDDs. Table S1 lists MDDs for each species. SM was also used as an individual explanatory factor.

DATA ANALYSIS

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

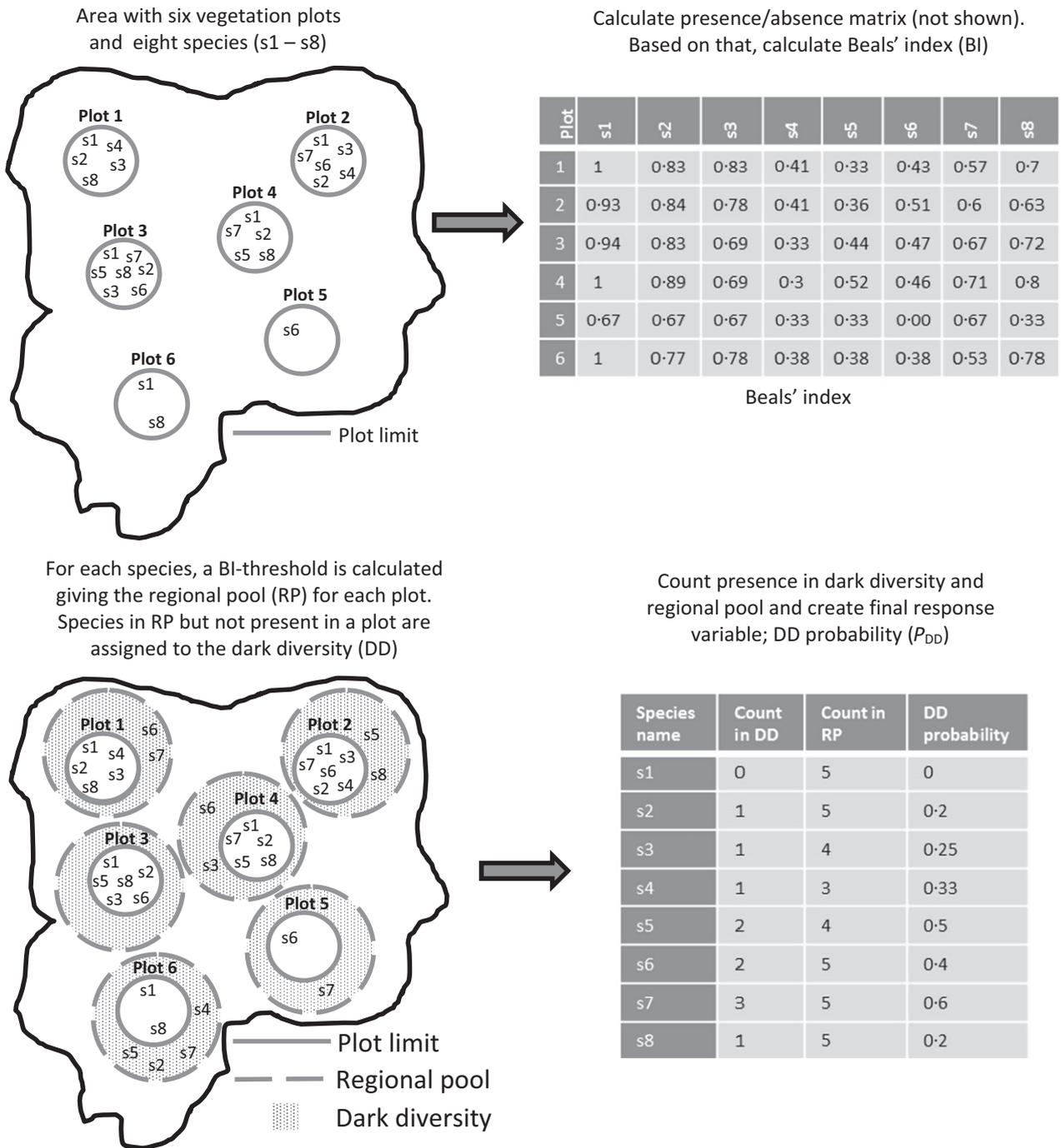


Fig. 2. The steps taken to produce the response variable (DD probability; P_{DD}) from vegetation plot data. A fictive area with six fictive plots and eight fictive species is used for illustrative purposes.

probability of being part of the dark diversity (Eqn 1) and the 10 explanatory variables listed in Table 1. All variables were tested for multicollinearity (see Figure S1). For subsequent analyses, SM and MDD were ln-transformed and all explanatory variables were standardised (subtracting the mean and dividing by the standard deviation).

Not all traits and characteristics were available for every species in the dataset (see method sections above). In order to maximise the sample size we initially used the entire set of study species ($n = 564$), but a limited number of explanatory variables

for statistical modelling (*high sample size models*, Tables 2 and S2). Secondly, we used all explanatory variables listed in Table 1 but a limited number of observations ($n = 167$, *low sample size models*) for statistical modelling. Since the three Grime-based variables were highly dependent on each other, only one Grime variable was included in a model at a time. Also, since the calculations of MDD involved SM these two variables never occurred simultaneously in any of the models to avoid redundancy. Following these constraints, we selected the best model within each sample size group using Akaike's Information Criteria (AIC). We

Table 2. Standardised coefficients of each of the 9 candidate models used to relate the plants' dark diversity probability (P_{DD} ; see eqn 1) to the explanatory variables listed in Table 1. Note that the coefficients shown originate from the models including phylogenetic filters (see methods section) where these were needed to account for phylogenetic autocorrelation. Results are presented for the *high* and *low sample-size models* (see details in Table S2). Grey cells mark explanatory variables that were not included in the focal model. For each model the adjusted explained deviance (D^2_{adj}) is shown together (if necessary) with the D^2_{adj} after phylogenetic filter addition. Int: model intercept, M: mycorrhiza. See Table 1 for remaining abbreviations

Model group	Model name	Predictor variables											D^2_{adj} (with filters)	No. of filters		
		L	F	N/R	GC	GS	GR	D	SM	MI	MD	Int			D^2_{adj}	
High sample size models ($n = 564$)	Grime C + D	-0.21***	0.03***	-0.03***	-0.19***	-0.00		-0.10***						1.92***	0.05	1
	Grime S + D	-0.22***	-0.03***	-0.11***				-0.14***						1.88***	0.02	
	Grime R + D	-0.27***	0.03***	-0.18***			0.21***	-0.12***						1.87***	0.05	
Low sample size models ($n = 167$)	Grime C + D + M	-0.36***	0.02***	-0.33***	-0.05***			-0.00						1.53***	0.09	3
	Grime S + D + M	-0.41***	0.02***	-0.44***				-0.06***						1.46***	0.10	
	Grime R + D + M	-0.45***	0.06***	-0.40***			0.19***	-0.01***						1.45***	0.12	
	Grime C + SM + M	-0.31***	0.03***	-0.35***	-0.07***									1.60***	0.10	
	Grime S + SM + M	-0.43***	0.01	-0.45***										1.48***	0.12	1
	Grime R + SM + M	-0.47***	0.06***	-0.42***			0.20***							1.47***	0.14	

*** $P < 0.001$.

followed best practice as recommended in Burnham & Anderson (2002) (for details on model selection and setup see Table S2).

The fact that closely related species tend to be more similar in traits and characteristics (Gittleman & Kot 1990) suggested testing for phylogenetic autocorrelation which we did as described below. We used the Daphne phylogenetic tree for the European flora (Durka & Michalski 2012) and followed Paradis (2015) to calculate Moran's I of each model's residuals using the reciprocal phylogenetic distances between species. This computation was performed using the 'Moran.I' function in the 'ape' package version 3.5 (Paradis, Claude & Strimmer 2004). Three of our models showed significant phylogenetic autocorrelation. To account for this we constructed phylogenetic eigenvector filters following best practice (Borcard & Legendre 2002; Diniz-Filho & Bini 2005). Significance and effect sizes of explanatory variables in the models were reassessed but no further model selection was performed.

All analyses were conducted using the R statistical software (R Core Team, 2016). D^2_{adj} was calculated using the following formula:

$$D^2_{adj} = 1 - \left(\frac{n-1}{n-p} \times (1 - D^2) \right) \quad \text{eqn 2}$$

where n and p are the numbers of observations and of adjusted coefficients in the model respectively, and D^2 is given by:

$$D^2 = \frac{\text{null deviance} - \text{residual deviance}}{\text{null deviance}} \quad \text{eqn 3}$$

Results

On average, plant species had a dark diversity probability of 88.8% (standard deviation: 12.8%). Forty-six percent of the species had a dark diversity probability greater than 95%, whereas for 7% of the species the probability of being part of dark diversity was less than 60% (Table S1).

The goodness-of-fit for our models was up to 0.15 (D^2_{adj}), with the low sample size models – including mycorrhizal information and either potential maximum dispersal distance or average seed mass – having the best fits (Table 2). Although most explanatory factors included in our models were statistically significant, the factors generally having the strongest relationships with dark diversity probability were: mycorrhizal dependence, preferences for light and nutrient availability, ruderality and seed mass (Table 2, Fig. 3). Maximum dispersal distance, mycorrhizal infection percentage, stress-tolerance and competitive ability were also consistent albeit less strong predictors of dark diversity probability (Table 2). We found strong indications that obligate mycorrhizal plants are more often part of the dark diversity than plants not depending on mycorrhiza (Fig. 4). This finding was supported by the fact that species with higher dark diversity probability had a higher degree of infection by mycorrhiza (Table 2, Fig. 4). In addition, plant species more likely to belong to dark diversity were adapted to thrive under low nutrient availability and low-light conditions (Fig. 4), had heavier seeds and were more often ruderal (Table 2, Fig. 4). Finally, we found some evidence that

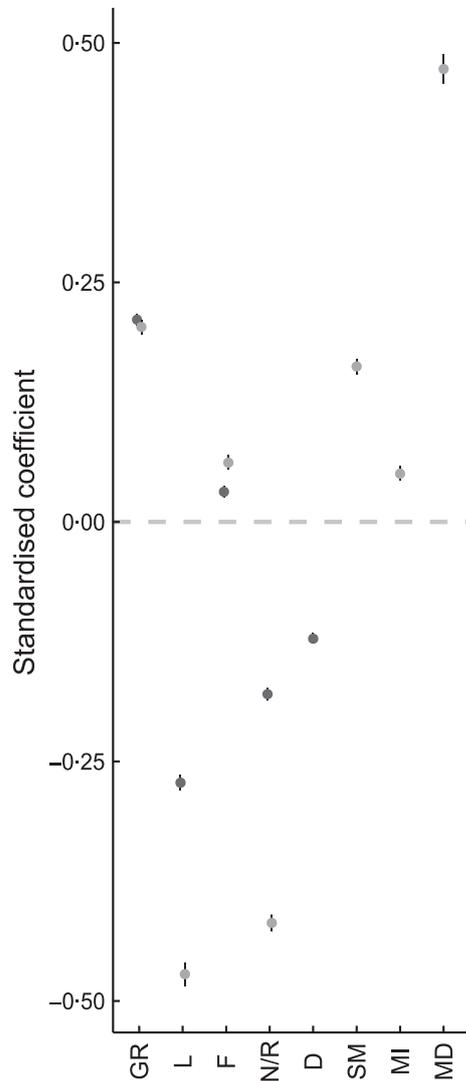


Fig. 3. Coefficient plots (made with the ‘multiplot’ function in the ‘coefplot’ package for R, Lander 2016) for the models named *Grime R + D* (dark grey) and *Grime R + SM + M* (light grey) in Table 2. Respectively, these were the best models in the two different model groups employed in this study (Table S2). Horizontal lines through each point represent the 95% confidence intervals. Consult Table 1 for abbreviations and Table 2 for model intercept values.

plants with a higher dark diversity probability are more often poor dispersers, poor competitors and stress-intolerant (cf. Grime 1979) (Table 2).

We found phylogenetic autocorrelation in the residuals of three models (Moran’s I test, $P < 0.05$). The addition of phylogenetic filters successfully removed autocorrelation and caused no notable shifts in effect sizes or significance.

Discussion

POTENTIAL DETERMINANTS AND THEIR LIKELY CAUSAL MECHANISMS

Apart from stochasticity (Hubbell 2001) and landscape characteristics (e.g. habitat heterogeneity, Krauss *et al.*

2004), the three general stages of successful plant colonisation are important for species’ dark diversity probability: (i) dispersal ability, (ii) establishment success and (iii) persistence in a given habitat (see introduction). The explanatory factors tested in this study all fall within these three categories. Factors primarily involved in establishment and persistence were the most important ones in our study overall: mycorrhizal dependence and the plants’ preference for available nutrients and light. Dispersal related factors such as the plants’ maximum dispersal distance and seed mass were also important (seed mass may be important in the establishment phase also, see below).

Establishment and persistence

A key step in plant establishment is the development of mycorrhizae (Akhmetzhanova *et al.* 2012; Gijbels, Adriaens & Honnay 2012; Hempel *et al.* 2013). Mycorrhizae are also important for the persistence of plant species and consequently for the local plant community composition (Hartnett & Wilson 1999). Conforming to this knowledge and to our hypothesis, we demonstrated that plants depending on mycorrhiza and plants requiring a high degree of mycorrhizal infection had higher dark diversity probability. On the other hand, recent evidence indicates that for some mycorrhizal species this may not be vital (e.g. Gijbels, Adriaens & Honnay 2012).

Competition among plants is another important phenomenon shaping the local structure, composition and richness of plant communities (Tilman 1994, 1997; McKane *et al.* 2002; Moeslund *et al.* 2013). In nutrient-rich landscapes such as Denmark, plant species like *Urtica dioica* L., *Epilobium hirsutum* L. and *Cirsium arvense* (L.) Scop. that benefit from high nutrient availability will be strong competitors (Grime 1979; Hill *et al.* 1999; Ellenberg *et al.* 2001) and consequently have lower dark diversity probability. This is probably the explanation for why we found that plant species with high dark diversity probability preferred low nutrient availability and in some cases were also less adapted to compete for resources (cf. Grime 1979). To broaden our understanding of this relationship, future studies could consider studying functional traits relevant for plants’ competitive ability; for example, specific leaf area and plant height (Cornelissen *et al.* 2003).

Contrary to our hypothesis, we demonstrated that shade tolerant plants had a higher dark diversity probability. The explanation for this finding is probably two-sided. Firstly, shade tolerant plants may have a competitive disadvantage in open landscapes (recall that this study concerns only open habitats). Secondly, the fact that the landscape around 5000 BC was almost completely forested (Fritzboøger & Odgaard 2010) can explain this observation. Indeed, shade tolerant species may have been present more frequently across Denmark in the past, when forests were more dominant. These species are likely to occur dormant in the soil seed bank

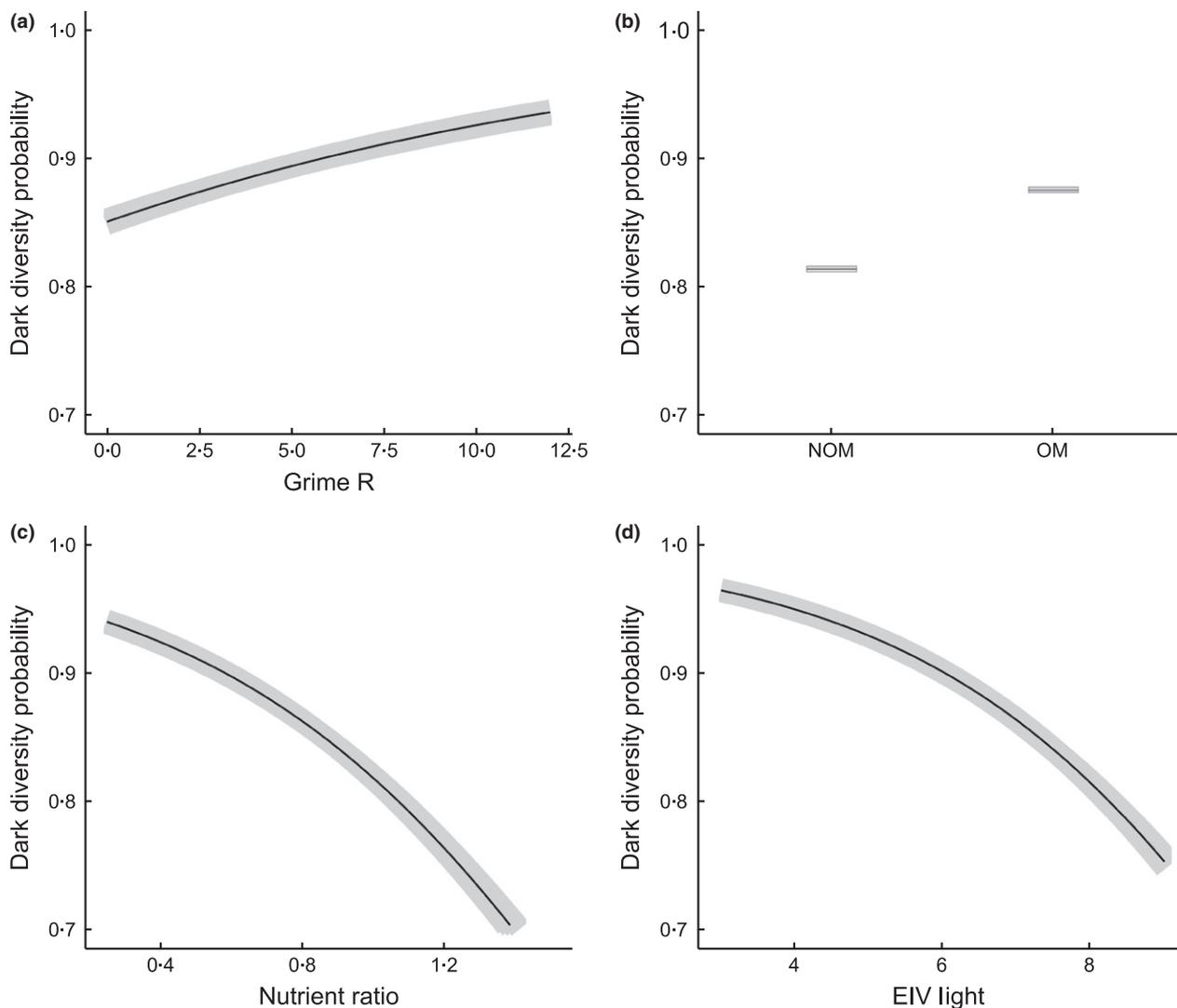


Fig. 4. Predictions (black line) and uncertainty (grey surface) of the relationship between the plant species' dark diversity probability and (a) Grime R, (b) mycorrhizal dependence, (c) nutrient ratio based on Ellenberg Indicator Values (EIVs) for soil nutrient availability and pH and (d) EIV for light. The predictions shown are based on the *Grime R + SM + M* model having the highest goodness-of-fit (Table 2, $n = 167$ species). The other predictor variables (covariates) were held constant (mean) for predictions. Using the 'sim' function in the 'arm' package (Gelman & Su 2015) for R, 1000 random simulations of the model were run. The regression lines of these simulations (grey surface) represent an estimation of uncertainty for each predictor variable. NOM: not obligately mycorrhizal, OM: obligately mycorrhizal.

(Telewski & Zeevaart 2002) and may return with afforestation (cf. memory effect due to land-use legacy, Plue *et al.* 2008).

In many environments, stress tolerance is a key factor shaping local plant diversity (Maun 1994; Ejrnæs & Bruun 2000; Moeslund *et al.* 2011). Recently, researchers showed that stress-tolerance was among the most important determinants of plant dark diversity in North-Eastern European dry calcareous grasslands (Riibak *et al.* 2015). They suggested that in the driest grasslands stress-tolerant species are more likely to thrive and therefore less likely to be part of the dark diversity. For our study this explanation could also be true. Many of the habitats we included are stressful environments either because they are relatively dry (grasslands, heathlands), waterlogged

(fens, bogs) or saline (salt meadows). On the other hand, plants' stress tolerance was negatively related to their ruderality and therefore it follows – from the fact that highly ruderal species cannot be simultaneously tolerant to high stress (Grime 1979) – that the probability of ending up in the dark diversity could actually be related to ruderality, not stress-tolerance. Ruderal species are mostly annual opportunistic species with low capability to persist. They tend to germinate quickly, set seeds and then wilt (Grime 1979); i.e. they depend on disturbances to create the opportunities (open or bare areas) needed for rapid germination. In most of Europe, agricultural landscapes that were formerly disturbed by extensive grazing and trampling have been increasingly abandoned over the latest decades (Henle *et al.* 2008). Encroachment following these

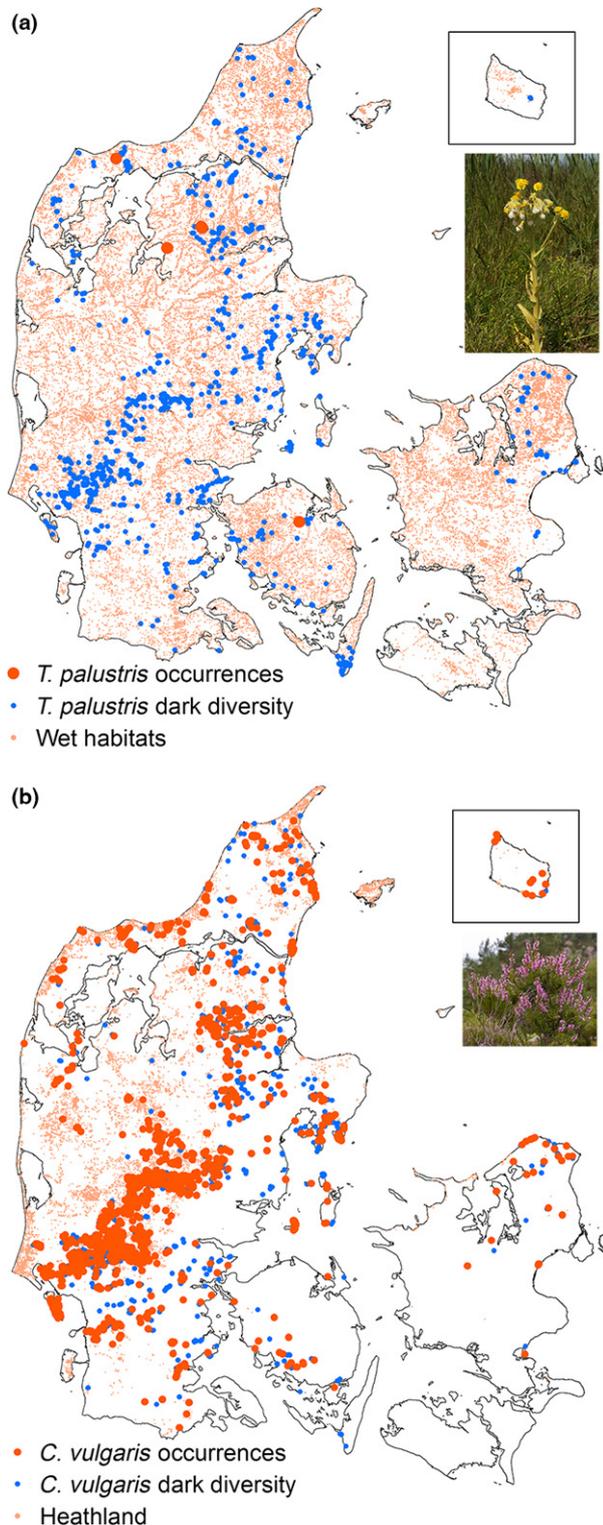


Fig. 5. The distribution of occurrences and dark diversity locations for (a) a species with a high dark diversity probability (*Tephroseris palustris*: 99.6%) and (b) a species with a relatively low dark diversity probability (*Calluna vulgaris* (L.) Hull: 37.3%). For reference, known wet habitats (a) and heathlands (b) are shown. Map inserts: island of Bornholm (see Fig. 1).

land-use changes may explain why ruderal species have a higher dark diversity probability.

Dispersal

Models accounting for the effect of dispersal distance but not for the effects of seed mass and mycorrhizal factors (*high sample size models*) suggested that species with higher dark diversity probability were also poorer dispersers. This supports the results from Riibak *et al.* (2015) who found that in dry calcareous grasslands small-scale dark diversity species are more likely to be dispersal limited. This is also consistent with earlier studies showing that species with poor dispersal abilities have a lower probability to recolonise sites where they formerly occurred (Tilman 1997; Cain, Milligan & Strand 2000; Myers & Harms 2009; Torrez *et al.* 2016). However, seed mass seemed to be a better predictor than dispersal distance in the *low sample size models*, perhaps because it captures both establishment (Moles & Westoby 2004) and dispersal processes: heavy seeds store more reserves for seedlings to establish but there are fewer of them and they are less likely to travel long distances and eventually reach suitable habitats (Cornelissen *et al.* 2003; Marteinsdóttir & Eriksson 2013; Riibak *et al.* 2015; but see Thomson *et al.* 2011).

UNCERTAINTIES AND POSSIBLE MODEL IMPROVEMENTS

As hypothesised – given their importance for plant species establishment, dispersal or persistence – most of the factors included here were important for plant species' dark diversity probability. However, the goodness-of-fit (D^2_{adj}) we obtained implies that other factors not tested here are also involved. Site conditions such as habitat fragmentation and reduced habitat patch sizes may be important issues affecting the metapopulation dynamics of plants and thus explain why so many species are missing in suitable habitats (Fahrig 2003). Therefore, variables capturing plant metapopulation dynamics such as species-specific immigration and extinction rates within a focal landscape are likely to prove useful in explaining plants' dark diversity probability. However, such variables are neither trivial nor easy to measure and attempts to directly relate plant traits to metapopulation dynamics are consequently rare (Dupré & Ehrlén 2002; May *et al.* 2013). Yet another site-related factor likely to affect a species' dark diversity probability is history; temporal and spatial continuity affects the probability of a species' successful establishment and persistence (e.g. Lindborg & Eriksson 2004). While plant ecological strategies, dispersal distance and mycorrhizal dependency relate to this, other factors mirroring plant species' response to management history could improve the explanatory power of models similar to ours in future studies. Furthermore, phenotypic

plasticity and general susceptibility to pathogens are likely to play a role for the persistence of plants (Augsburger & Kelly 1984; Burdon, Thrall & Ericson 2006; Reed *et al.* 2010).

One drawback of basing a study on data from a large monitoring programme is that issues concerning overlooked species (*hidden diversity*, Milberg *et al.* 2008) and misidentifications are almost unavoidable. If such errors are biased towards species of certain traits it could influence our results. Also, using relatively small plots (c. 78 m²) to derive the dark diversity could possibly introduce uncertainty since they are likely not to cover all species actually present in a site if laid out randomly. Notably, rare but typical species could be overlooked in this way. However, the use of trained botanists for these surveys and the representative way to lay out the plots we used (see methods) significantly elevates the probability that the majority of the plots actually do capture the rare and 'hidden' species. Indeed, a large sampling effort such as the one we used here (15 160 plots scattered throughout Denmark, Fig. 1) can considerably overcome the potential issue of overlooking very rare but typical species. For example, *Tephrosia palustris* (L.) Rchb. was often found in the dark diversity (Fig. 5a). This species has a wide distribution in Northern Europe but is relatively rare in Denmark, extinct in Great Britain, the Czech Republic, Hungary, Slovakia and Romania (Kochjarová 2006) and critically endangered in Sweden (Olsson & Tyler 2001). However, it tends to colonise riparian bare mud and is even occasionally found in recycling depots (Frederiksen, Rasmussen & Seberg 2006) – habitats that are common in Denmark. This was also the case for *Leontodon hispidus* L., *Campanula persicifolia* L., *Vicia tetrasperma* (L.) Schreb. and several other rare but typical species in Denmark (Table S1) despite their habitats being common. Hence, habitat availability alone cannot explain rarity. We are therefore confident that there is no confounding effect between a species' dark diversity probability and its rarity. However, for future surveys, we recommend using spatially or temporally repeated censuses at a subset of sites in order to estimate species-occupancy models, thus accounting for the probability that a species is present but not detected. This should limit the potential bias due to the low detectability of rare but typical species in relatively small plots and consequently improve the accuracy of the actual dark diversity value. We thus note that the actual dark diversity probability of a plant species is likely to differ when applying this recommendation, but this is unlikely to affect our findings on the main determinants of plant species absence.

In general, rare but typical species might be overlooked at fine spatial resolutions, but applying the dark diversity concept at relatively coarse resolutions (landscape-to-regional scale, 10–100 km) is likely to underestimate dark diversity at the spatial scale relevant for conservation and restoration issues (local-to-landscape scale, 0.1–10 km)

simply because larger areas support more species, all else being equal (McArthur & Wilson 1967).

MANAGEMENT IMPLICATIONS

To halt biodiversity loss by 2020 (European Commission, 2011) it is imperative to understand the local species pools and identify the mechanisms determining why species are sometimes missing from areas that seem suitable. This is just one aspect through which dark diversity can help enlighten conservation and management issues (Lewis *et al.* 2017). Although not intuitive, we can learn a lot from absent species, not only in a theoretical perspective but also from a more practical point of view (e.g. for conservation and restoration actions). The concept of dark diversity is still in its infancy and this study just represents an example of its application to nature management. Dark diversity could also help us learn more about red-listed species and the susceptibility of communities to invasive species. Lewis *et al.* (2017) provide more inspiration on how dark diversity can be used in conservation and restoration actions. Ground truthing the presence/absence of predicted dark diversity species along a range of environmental gradients, management practices or successional stages is likely to take the application of dark diversity to the next level.

Our findings underpin the importance of assessing the mycorrhizal fungi's ability to sustain the flora at a given site, indicating that inoculation with certain mycorrhizal fungi could be important for successful restoration of plant communities (Torrez *et al.* 2016). However, great care needs to be taken during this process with the best results probably obtained by adding a diverse and locally adapted mycorrhizal community (Klironomos 2003).

The importance of dispersal limitation for species' dark diversity probability strongly suggests that space and time are imperative factors in the planning and management of nature. Given enough time and suitable corridors for dispersal, even poor dispersers will eventually reach suitable but distant habitats. To accommodate this, conservation actions might need to consider assisted migration (Seddon 2010).

Also, our study suggests focusing on creating opportunities for ruderal species in restoration and conservation projects; e.g. by ensuring bare soil for seed germination through disturbance processes such as erosion, flooding, fire and trampling by large herbivores. However, to avoid unintended introductions, care needs to be taken to prevent establishment of exotic ruderal weeds.

Finally, our results strongly emphasise the importance of focusing on nutrients and light availability in conservation and restoration actions. Failing to ensure nutrient-poor sites in regions with heavy nutrient loads will render a number of species unable to thrive in otherwise suitable areas. To accommodate low-light-adapted species, a final recommendation based on this study is to increase vegetation heterogeneity by promoting single-standing trees and bushes in otherwise homogenous swards.

Authors' contributions

J.E.M., A.K.B., K.K.C., L.D., C.F., A.J. and J.L. conceived the ideas and designed methodology; J.E.M., A.K.B., L.D. and C.F. collected the data; J.E.M., A.K.B., K.K.C., L.D., C.F. and J.L. analysed the data; J.E.M. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Acknowledgements

We thank 15. Juni Fonden for economic support (2016-B-34) and three anonymous reviewers who provided invaluable suggestions for improving our paper.

Data accessibility

Species geography data is available from the Danish environment portal (<http://naturdata.miljoeportal.dk/searchnew.aspx>). Species trait data is available from Akhmetzhanova *et al.* (2012) and the following data bases: LEDA (Kleyer *et al.* 2008), baseflor (philippe.julve.pagesperso-orange.fr/baseflor.xlsx, data owned by Philippe Julve, Catholic University of Lille, France), BiolFlor (Kühn, Durka & Klotz 2004) and MycoFlor (Hempel *et al.* 2013). The processed data used for analyses are available from Table S1.

References

- Akhmetzhanova, A.A., Soudzilovskaia, N.A., Onipchenko, V.G., Cornwell, W.K., Agafonov, V.A., Selivanov, I.A. & Cornelissen, J.H.C. (2012) A rediscovered treasure: mycorrhizal intensity database for 3000 vascular plant species across the former Soviet Union. *Ecology*, **93**, 689–690.
- Andersen, D.K., Nygaard, B., Fredshavn, J.R. & Ejrnæs, R. (2013) Cost-effective assessment of conservation status of fens. *Applied Vegetation Science*, **16**, 491–501.
- Augsburger, C.K. & Kelly, C.K. (1984) Pathogen mortality of tropical tree seedlings: experimental studies of the effects of dispersal distance, seedling density, and light conditions. *Oecologia*, **61**, 211–217.
- Beals, E.W. (1984) Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. *Advances in Ecological Research*, **14**, 1–55.
- Borcard, D. & Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*, **153**, 51–68.
- Brauer, V.S., Stomp, M. & Huisman, J. (2012) The nutrient-load hypothesis: patterns of resource limitation and community structure driven by competition for nutrients and light. *The American Naturalist*, **179**, 721–740.
- Buchwald, E., Wind, P., Bruun, H.H., Møller, P.F., Ejrnæs, R. & Svart, H.E. (2013) Hvilke planter er hjemmehørende i Danmark? *Flora & Fauna*, **118**, 73–96.
- Burdon, J.J., Thrall, P.H. & Ericson, A.L. (2006) The current and future dynamics of disease in plant communities. *Annual Review of Phytopathology*, **44**, 19–39.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, New York, NY, USA.
- Cain, M.L., Milligan, B.G. & Strand, A.E. (2000) Long-distance seed dispersal in plant populations. *American Journal of Botany*, **87**, 1217–1227.
- Cornelissen, J.H.C., Lavorel, S., Garnier, E. *et al.* (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, **51**, 335–380.
- Diekmann, M. (2003) Species indicator values as an important tool in applied plant ecology – a review. *Basic and Applied Ecology*, **4**, 493–506.
- Diekmann, M. & Falkengren-Grerup, U. (1998) A new species index for forest vascular plants: development of functional indices based on mineralization rates of various forms of soil nitrogen. *Journal of Ecology*, **86**, 269–283.
- Diniz-Filho, J.A.F. & Bini, L.M. (2005) Modelling geographical patterns in species richness using eigenvector-based spatial filters. *Global Ecology and Biogeography*, **14**, 177–185.
- Dupré, C. & Ehrlén, J. (2002) Habitat configuration, species traits and plant distributions. *Journal of Ecology*, **90**, 796–805.
- Durka, W. & Michalski, S.G. (2012) Daphne: a dated phylogeny of a large European flora for phylogenetically informed ecological analyses. *Ecology*, **93**, 2297.
- Ejrnæs, R. & Bruun, H.H. (2000) Gradient analysis of dry grassland vegetation in Denmark. *Journal of Vegetation Science*, **11**, 573–584.
- Ellenberg, H., Weber, H.E., Düll, R., Wirth, V. & Werner, W. (2001) *Zeigerwerte von planzen in Mitteleuropa*, 3rd edn. Erich Goltze GmbH & Co KG, Göttingen, Germany.
- European Commission (2011) *The EU Biodiversity Strategy to 2020*. Publications Office of the European Union, Luxembourg.
- Fahrig, L. (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 487–515.
- Finlay, R.D. (2008) Ecological aspects of mycorrhizal symbiosis: with special emphasis on the functional diversity of interactions involving the extraradical mycelium. *Journal of Experimental Botany*, **59**, 1115–1126.
- Frederiksen, S., Rasmussen, F.N. & Seberg, O. (2006) *Dansk Flora*, 1st edn. Gyldendal, Copenhagen, Denmark.
- Fredshavn, J.R., Nygaard, B. & Ejrnæs, R. (2009) *Naturtilstand på terrestriske naturarealer – besigtigelser af §3-arealer*. Danmarks Miljøundersøgelser, Aarhus University, Aarhus, Denmark.
- Fritzboeger, B. & Odgaard, B. (2010) Skovenes historie. *Naturen i Danmark – Skovene* (ed P.F. Møller), pp. 55–88. Gyldendal, Copenhagen, Denmark.
- Gelman, A. & Su, Y. (2015) *Arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R Package Version 1.8-6*.
- Gijbels, P., Adriaens, D. & Honnay, O. (2012) An orchid colonization credit in restored calcareous grasslands. *Ecoscience*, **19**, 21–28.
- Gittleman, J.L. & Kot, M. (1990) Adaptation: statistics and a null model for estimating phylogenetic effects. *Systematic Zoology*, **39**, 227–241.
- Grime, J.P. (1979) *Plant Strategies and Vegetation Processes*, 1st edn. John Wiley & Sons, Chichester, New York, Brisbane, Toronto.
- Hansen, K. (1996) *Dansk Feltflora*, 1st edn. Gyldendal, Copenhagen, Denmark.
- Hartnett, D.C. & Wilson, G.W.T. (1999) Mycorrhizae influence plant community structure and diversity in tallgrass prairie. *Ecology*, **80**, 1187–1195.
- Hempel, S., Götzenberger, L., Kühn, I., Michalski, S.G., Rillig, M.C., Zobel, M. & Moora, M. (2013) Mycorrhizas in the Central European flora: relationships with plant life history traits and ecology. *Ecology*, **94**, 1389–1399.
- Henle, K., Alard, D., Clitherow, J. *et al.* (2008) Identifying and managing the conflicts between agriculture and biodiversity conservation in Europe – a review. *Agriculture, Ecosystems & Environment*, **124**, 60–71.
- Hill, M.O., Mountford, J.O., Roy, D.B. & Bunce, R.G.H. (1999) *Ellenberg's Indicator Values for British Plants. Ecofact Volume 2 Technical Annex*. Institute of Terrestrial Ecology, Huntingdon, UK.
- Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*, 1st edn. Princeton University Press, Princeton, NJ, USA and Oxford, UK.
- Kleyer, M., Bekker, R.M., Knevel, I.C. *et al.* (2008) The LEDA Traitbase: a database of life-history traits of the Northwest European flora. *Journal of Ecology*, **96**, 1266–1274.
- Klironomos, J.N. (2003) Variation in plant response to native and exotic arbuscular mycorrhizal fungi. *Ecology*, **84**, 2292–2301.
- Kochjarová, J. (2006) Contribution to the occurrence and former distribution of *Tephrosia palustris* (Compositae) in the Central Europe. *Biologia*, **61**, 361–364.
- Krauss, J., Klein, A., Steffan-Dewenter, I. & Tschamntke, T. (2004) Effects of habitat area, isolation, and landscape diversity on plant species richness of calcareous grasslands. *Biodiversity & Conservation*, **13**, 1427–1439.
- Kühn, I., Durka, W. & Klotz, S. (2004) BiolFlor – a new plant-trait database as a tool for plant invasion ecology. *Diversity and Distributions*, **10**, 363–365.
- Lander, J.P. (2016) *Coeffplot: Plots Coefficients from Fitted Models. R Package Version 1.2.4*.
- Lenoir, J., Gégout, J., Guisan, A. *et al.* (2010) Cross-scale analysis of the region effect on vascular plant species diversity in Southern and Northern European mountain ranges. *PLoS ONE*, **5**, e15734.
- Lewis, R.J., Szava-Kovats, R. & Pärtel, M. (2016) Estimating dark diversity and species pools: an empirical assessment of two methods. *Methods in Ecology and Evolution*, **7**, 104–113.
- Lewis, R.J., de Bello, F., Bennett, J.A. *et al.* (2017) Applying the dark diversity concept to nature conservation. *Conservation Biology*, **31**, 40–47.

- Lindborg, R. & Eriksson, O. (2004) Historical landscape connectivity affects present plant species diversity. *Ecology*, **85**, 1840–1845.
- Marteinsdóttir, B. & Eriksson, O. (2013) Trait-based filtering from the regional species pool into local grassland communities. *Journal of Plant Ecology*, **7**, 347–355.
- Maun, M.A. (1994) Adaptations enhancing survival and establishment of seedlings on coastal dune systems. *Plant Ecology*, **111**, 59–70.
- May, F., Giladi, I., Ristow, M., Ziv, Y. & Jeltsch, F. (2013) Plant functional traits and community assembly along interacting gradients of productivity and fragmentation. *Perspectives in Plant Ecology, Evolution and Systematics*, **15**, 304–318.
- McArthur, R.H. & Wilson, E.O. (1967) *The Theory of Island Biogeography*. Princeton University Press, Princeton, NJ, USA.
- McCune, B. (1994) Improving community analysis with the Beals smoothing function. *Ecoscience*, **1**, 82–86.
- McKane, R.B., Johnson, L.C., Shaver, G.R. *et al.* (2002) Resource-based niches provide a basis for plant species diversity and dominance in arctic tundra. *Nature*, **415**, 68–71.
- Milberg, P., Bergstedt, J., Fridman, J., Odell, G. & Westerberg, L. (2008) Observer bias and random variation in vegetation monitoring data. *Journal of Vegetation Science*, **19**, 633–644.
- Moeslund, J.E., Arge, L., Bøcher, P., Nygaard, B. & Svenning, J. (2011) Geographically comprehensive assessment of salt-meadow vegetation-elevation relations using LiDAR. *Wetlands*, **31**, 471–482.
- Moeslund, J.E., Arge, L., Bøcher, P.K., Dalgaard, T., Odgaard, M.V., Nygaard, B. & Svenning, J. (2013) Topographically controlled soil moisture is the primary driver of local vegetation patterns across a lowland region. *Ecosphere*, **4**, art91.
- Moles, A.T. & Westoby, M. (2004) Seedling survival and seed size: a synthesis of the literature. *Journal of Ecology*, **92**, 372–383.
- Münzbergová, Z. & Herben, T. (2004) Identification of suitable unoccupied habitats in metapopulation studies using co-occurrence of species. *Oikos*, **105**, 408–414.
- Myers, J.A. & Harms, K.E. (2009) Seed arrival, ecological filters, and plant species richness: a meta-analysis. *Ecology Letters*, **12**, 1250–1260.
- Oksanen, J., Blanchet, F.G., Kindt, R. *et al.* (2015) *Vegan: Community Ecology Package. R Package Version 2.3-2*.
- Olsson, K. & Tyler, T. (2001) A historic perspective on the red-listed species of Scania, S. Sweden. *Botaniska Notiser*, **134**, 1–5.
- Paradis, E. (2015) *Moran's Autocorrelation Coefficient in Comparative Methods. R Vignette*.
- Paradis, E., Claude, J. & Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.
- Pärtel, M., Szava-Kovats, R. & Zobel, M. (2011) Dark diversity: shedding light on absent species. *Trends in Ecology & Evolution*, **26**, 124–128.
- Plue, J., Hermy, M., Verheyen, K., Thuillier, P., Saguez, R. & Decocq, G. (2008) Persistent changes in forest vegetation and seed bank 1,600 years after human occupation. *Landscape Ecology*, **23**, 673–688.
- R Core Team. (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Reed, T.E., Waples, R.S., Schindler, D.E., Hard, J.J. & Kinnison, M.T. (2010) Phenotypic plasticity and population viability: the importance of environmental predictability. *Proceedings of the Royal Society of London B: Biological Sciences*, **277**, 3391–3400.
- Riibak, K., Reitalu, T., Tamme, R. *et al.* (2015) Dark diversity in dry calcareous grasslands is determined by dispersal ability and stress-tolerance. *Ecography*, **38**, 713–721.
- Ronk, A., Szava-Kovats, R. & Pärtel, M. (2015) Applying the dark diversity concept to plants at the European scale. *Ecography*, **38**, 1015–1025.
- Seddon, P.J. (2010) From reintroduction to assisted colonization: moving along the conservation translocation spectrum. *Restoration Ecology*, **18**, 796–802.
- Seidling, W. (2005) Ground floor vegetation assessment within the intensive (Level II) monitoring of forest ecosystems in Germany: chances and challenges. *European Journal of Forest Research*, **124**, 301–312.
- Tamme, R., Götzenberger, L., Zobel, M., Bullock, J.M., Hooftman, D.A.P., Kaasik, A. & Pärtel, M. (2014) Predicting species' maximum dispersal distances from simple plant traits. *Ecology*, **95**, 505–513.
- Telewski, F.W. & Zeevaart, J.A.D. (2002) The 120-yr period for Dr. Beal's seed viability experiment. *American Journal of Botany*, **89**, 1285–1288.
- Thomson, F.J., Moles, A.T., Auld, T.D. & Kingsford, R.T. (2011) Seed dispersal distance is more strongly correlated with plant height than with seed mass. *Journal of Ecology*, **99**, 1299–1307.
- Tilman, D. (1994) Competition and biodiversity in spatially structured habitats. *Ecology*, **75**, 2–16.
- Tilman, D. (1997) Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology*, **78**, 81–92.
- Torrez, V., Ceulemans, T., Mergeay, J., de Meester, L. & Honnay, O. (2016) Effects of adding an arbuscular mycorrhizal fungi inoculum and of distance to donor sites on plant species recolonization following topsoil removal. *Applied Vegetation Science*, **19**, 7–19.
- Yoshioka, A., Miyazaki, Y., Sekizaki, Y., Suda, S., Kadoya, T. & Washitani, I. (2014) A “lost biodiversity” approach to revealing major anthropogenic threats to regional freshwater ecosystems. *Ecological Indicators*, **36**, 348–355.

Received 3 October 2016; accepted 5 January 2017

Handling Editor: Luke Flory

Supporting Information

Details of electronic Supporting Information are provided below.

Fig. S1. Correlation plot of all explanatory variables.

Table S1. Processed data used for analyses.

Table S2. Model selection and setup.