Journal of Applied Ecology

BRITISH ECOLOGICAL SOCIETY

Journal of Applied Ecology 2017, **54,** 116–126 doi: 10.1111/1365-2664.12636

Genetic differentiation within multiple common grassland plants supports seed transfer zones for ecological restoration

Walter Durka[1,2*†], Stefan G. Michalski[1†], Kenneth W. Berendzen[3], Oliver Bossdorf[4], Anna Bucharova[4], Julia-Maria Hermann[5], Norbert Hölzel[6] and Johannes Kollmann[5]

[1]*Department of Community Ecology (BZF), Helmholtz Centre for Environmental Research – UFZ, Theodor-Lieser-Str. 4, 06120 Halle, Germany;* [2]*German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany;* [3]*ZMBP – Central Facilities, University of Tübingen, Auf der Morgenstelle 32, 72076 Tübingen, Germany;* [4]*Plant Evolutionary Ecology, Institute of Evolution & Ecology, University of Tübingen, Auf der Morgenstelle 5, 72076 Tübingen, Germany;* [5]*Restoration Ecology, Department of Ecology & Ecosystem Management, Technical University Munich, Emil-Ramann-Str. 6, 85354 Freising, Germany; and* [6]*Biodiversity and Ecosystem Research Group, Institute of Landscape Ecology, University of Münster, Heisenbergstr. 2, 48149 Münster, Germany*

Summary

1. Ecological restoration of grasslands is increasingly based on regional seeds derived from predefined seed transfer zones. However, the degree and spatial pattern of genetic differentiation among provenances of different seed transfer zones is largely unknown.

2. We assessed the genetic differentiation among eight out of 22 German seed transfer zones for seven common grassland species (*Arrhenatherum elatius, Centaurea jacea, Daucus carota, Galium album, Hypochaeris radicata, Knautia arvensis* and *Lychnis flos-cuculi*) using AFLP markers. We analysed genetic population structure with AMOVA and Bayesian cluster analysis and tested for isolation by distance and isolation by environment.

3. In all of the investigated species, almost all pairs of provenances were genetically differentiated. Bayesian cluster analysis revealed species-specific numbers and spatial patterns of gene pools, with between two (*Arrhenatherum*) and eight clusters (*Lychnis*). Most investigated seed transfer zones represented a unique gene pool in the majority of the species.

4. We found isolation by distance in four species, isolation by environment, driven by climatic seasonality, in three species, and a lack of both in three species. Thus, the observed genetic differentiation appears to be caused by both neutral and adaptive processes.

5. *Synthesis and applications*. Our study shows that grassland plants are indeed strongly genetically differentiated across Germany supporting the strategy of seed transfer zones for ecological restoration. Although the predefined seed transfer zones are unlikely to match the exact genetic structure of many species, they serve their purpose by capturing a substantial amount of intraspecific genetic variation across species.

Key-words: amplified fragment length polymorphism, ecological restoration, genetic differentiation, genetic diversity, grasslands, isolation by distance, isolation by environment, *Knautia arvensis*, local provenancing, polyploidy, seed transfer zone

Introduction

Semi-natural, extensively used grasslands in Europe are threatened by habitat destruction and fragmentation, land-use intensification, climate change and biological invasions (MEA 2005). Because of their biodiversity, aesthetic value and the ecosystem services they provide, semi-natural grasslands are an important target of conservation and ecological restoration (Bakker *et al*. 2012; Kiehl *et al*. 2014).

The identity of the seed sources is a major issue in practical grassland restoration. The use of local or regional

*Correspondence author. E-mail: walter.durka@ufz.de
†These authors contributed equally to the paper

seeds is often advocated because they are considered to be better adapted to local site conditions which are expected to increase restoration success (Sackville Hamilton 2001; Millar, Byrne & Coates 2008) compared to potentially maladapted non-local plant material (Bischoff, Steinger & Müller-Schärer 2010; but see Sgrò, Lowe & Hoffmann 2011). The use of local seed sources is also an important means to conserve genetic biodiversity (Krauss *et al.* 2013), and it reduces the risks of genetic swamping and of outbreeding depression (Hufford & Mazer 2003).

The use of local seed provenances is usually implemented through a geographic delineation of seed transfer zones within which seeds are to be collected, propagated and used in restoration. Ideally, as the use of local seed sources is motivated by the existence of intraspecific genetic differentiation (McKay *et al.* 2005), the seed transfer zones should reflect the spatial genetic structure of plant species (Hufford & Mazer 2003). In forestry, there is a long history of using seed transfer zones for trees based on either trait divergence or genetic differentiation in molecular markers (FoVHgV 2003; De Kort *et al.* 2014). However, for herbaceous plants, the first attempts to delineate seed transfer zones have been made only recently based on trait variation (e.g. Miller *et al.* 2011; St Clair *et al.* 2013) or molecular markers (Malaval *et al.* 2010; Jorgensen *et al.* 2014).

While seed transfer zones for individual species can be based on their trait or molecular variation, generalized seed transfer zones are sometimes also defined based on climate, geology and other biophysical and biogeographic criteria (Vander Mijnsbrugge, Bischoff & Smith 2010; Bower, Clair & Erickson 2014). Such generalized seed transfer zones for grassland species are implemented, for example, in Germany (ErMiV 2011) and Switzerland (SKEW 2009), implicitly assuming that the criteria for delineation and the patterns of genetic differentiation are largely similar among species. However, while genetic differentiation and local adaptation are common in plants and have been demonstrated in a large number of individual species and case studies (Leimu & Fischer 2008), few attempts have been made to compare these characteristics for multiple species in the same geographic context (but see Malaval *et al.* 2010; Miller *et al.* 2011). Thus, it is largely unknown how similar genetic differentiation patterns are among species across seed transfer zones.

The factors that affect genetic population structure in plants are generally well understood. Genetic differentiation among populations builds up due to the joint influences of dispersal limitation, adaptation and colonization history (Orsini *et al.* 2013). Dispersal limitation reduces gene flow among populations and results in isolation by distance (IBD), that is an increase in genetic differentiation with increasing geographic distance. Adaptation to local or regional environmental conditions results in a pattern of isolation by environment (IBE: Sexton, Hangartner & Hoffmann 2014), where genetic differentiation increases with increasing environmental distance, for example climatic differences. However, the lack of IBD and IBE is also a common pattern and can indicate whether gene flow or genetic drift is more influential (Hutchison & Templeton 1999). Finally, the colonization history of a species can strongly affect its genetic structure, often through founder effects. Particularly in areas affected by previous glacial cycles, pronounced phylogeographic patterns are found (e.g. Fjellheim *et al.* 2006), which to a certain extent are generalizable across species (Taberlet *et al.* 1998).

Another important factor that can affect the genetic population structure of species is polyploidy. Because of breeding barriers between cytotypes (e.g. Köhler, Mittelsten Scheid & Erilova 2010), the existence of multiple cytotypes within species can drastically affect their population structure. However, both the large-scale distribution and the degree of small-scale coexistence of different cytotypes are unknown for many species and geographic ranges (Kolář *et al.* 2009). Therefore, such different cytotypes need to be recognized and taken into account when interpreting the patterns of genetic differentiation.

All of the above-mentioned factors affecting genetic differentiation are to some degree species-specific, and it is therefore an important question whether generalized seed transfer zones make sense across many species. To address this question, we investigated the patterns of genetic variation among seed transfer zones in Germany for seven grassland species. For each species we asked: (i) Are seed provenances genetically differentiated? (ii) If yes, is genetic differentiation related to spatial and/or climatic distance? (iii) How consistent are the patterns of genetic differentiation among seed transfer zones across the seven species?

Materials and methods

STUDY SPECIES AND SEED SOURCES

In Germany, a system of seed sourcing, propagation and marketing for common species used in grassland restoration has recently been established (Prasse, Kunzmann & Schröder 2010; ErMiV 2011). It defines 22 seed transfer zones ('Herkunftsregionen', Fig. 1) using the system of physiographic regions of Germany based on climate, geological substrate and soil types (Meynen & Schmithüsen 1953-1962). For each zone, a specific list of native plant species has been defined that can be collected, propagated and marketed. Seeds must be collected in their typical habitat in Natura 2000 areas protected by EU legislation or in sites of similar quality, that is natural or semi-natural habitats in which no sowing has taken place for at least 40 years. Mixing of several source sites within a seed transfer zone is advocated. Seeds can be propagated for a maximum of five generations within eight larger regions ('Produktionsräume') to which several seed transfer zones have been merged (Fig. 1).

For our study, we selected seven common grassland species: *Arrhenatherum elatius* (L.) P.B. ex J. et C. Presl, *Centaurea jacea* L., *Daucus carota* L., *Galium album* Mill., *Hypochaeris radicata* L., *Knautia arvensis* (L.) Coult. and *Lychnis flos-cuculi* (L.) Greuter & Burdet (genera used as abbreviation hereafter). Because the 22 seed transfer zones have only recently been established and are not fully functional yet and only eight major

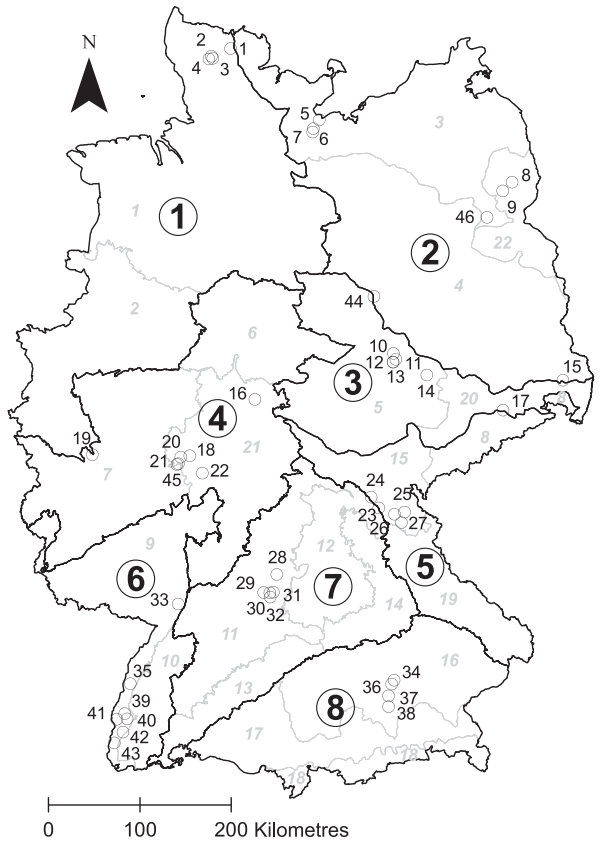

Fig. 1. Map of German seed transfer zones (grey outlines, grey numbers 1–22 in italics; after Prasse, Kunzmann & Schröder (2010)), regions (black outlines, large circled numbers 1–8) and the source sites of analysed provenances (1–46; see Table S1 for information on which species was collected where).

regions were informally distinguished previously, we focused, if possible, on one particular seed transfer zone in each of the eight regions. For each species, we purchased seeds from one provenance within each of the eight regions from Rieger-Hoffmann GmbH (Blaufelden, Germany) and affiliated seed producers (Fig. 1, Table S1 in Supporting Information). Out of the 56 provenances, 44 originated from a single source site and 12 were mixtures from 2 to 5 source sites. Seeds had been propagated for up to four generations. For the molecular analyses, we grew plants in standard soil and collected, when possible, leaf material from 12 plants per provenance and species (see Bucharova *et al.* 2016 for details).

PLOIDY

To identify possible multiple cytotypes within species, we used flow cytometry. The analyses were carried out on leaves from plants grown for this purpose from the same seed material that was used for the DNA analysis. For each provenance, we sampled five random plants. For methodological details, see Appendix S1.

GENOTYPING

For each species, we performed amplified fragment length polymorphism analysis (AFLP) following the protocol of Kloss, Fischer & Durka (2011). We extracted DNA with DNeasy 96 kits (QIAGEN) and performed restriction ligation in 11 μl with 6 μl of DNA (~150 ng DNA) and MseI and EcoRI restriction enzymes at 37 °C for 2 h. After 1 : 5 dilution, we used 4 μl for preselective amplification, which again was diluted between 1 : 5 and 1 : 20, depending on species, for selective amplification. After screening 32 primer combinations, we selected three or four primer combinations per species for genotyping (Table S2). The fragments were separated on an ABI 3130 genetic analyser and binned manually in GeneMapper 5.0. After exporting peak-height data, we calculated the frequency distribution for each band and, if possible, used it to define the genotyping threshold (default value = 100 rfu) to optimize presence–absence calling. To estimate error rates, we analysed between 3 and 30 (mean: 17) duplicate samples per species. AFLP bands with large individual error rates, that is non-reproducible bands, and bands with a unimodal frequency distribution were excluded from the analysis. Eventually, we retained between 153 and 268 AFLP loci per species with a mean genotypic error rate of 2·4% (Table S2).

DATA ANALYSIS

Unless otherwise stated, all data analyses were done with R 3·1·2 (R Core Team 2015). Genetic population structure was assessed and quantified in several steps. First, we used principle coordinates analysis (PCoA) to illustrate Euclidian genetic distances between individuals. Next, we used analysis of molecular variance (AMOVA: Excoffier, Smouse & Quattro 1992) to quantify overall and pairwise genotypic differentiation (F_{ST}) among provenances, as implemented in GenAlex 6.5 (Peakall & Smouse 2012).

To further assess the relationships between individuals and provenances, we applied a Bayesian clustering approach in which we did not use population origin or spatial coordinates as prior. We used structure 2·3·4 (Falush, Stephens & Pritchard 2007) in the recessive allele mode advocated for dominant markers. For each species we ran, for assumed cluster numbers ranging from K = 1–10, ten independent runs of an admixture model with 150 000 MCMC (Markov chain Monte Carlo) iterations, discarding the first 50 000 as burn-in. When model likelihood showed a large variation across runs in the range of particular Ks, we repeated the analysis with 400 000 MCMC iterations (200 000 burn-in). In order to identify the most probable number of genetic clusters, we scrutinized whether maxima were reached for both the model likelihood L (K) and the parameter ΔK (Evanno, Regnaut & Goudet 2005). This is particularly important since the Evanno method is unable to identify a lack of structure (K = 1). As structure may detect only an upper hierarchical level of population structure, we repeated the analyses with subsets of the data. Provenances were assigned to clusters when the mean assignment probability was >0·5. Consensus results across replicate runs were obtained with CLUMPP (Jakobsson & Rosenberg 2007).

To test for IBD and IBE, we calculated a matrix of geographic and climatic distances among the original source sites. For climatic distance, we extracted data from WorldClim (http://www.worldclim.org/, Hijmans *et al.* 2005) in 2·5 arc-min resolution, carried out a principle component analysis on all 19 scaled bioclimatic variables and extracted the first two PCs which together explained 64% of variation. The first climate PC (Clim1) represents a dry/hot-wet/cool cline, that is a temperature

and precipitation gradient, with significant loadings by most minimum, mean and maximum values of temperature (negative loading) and by all precipitation variables. The second climate PC (Clim2) mostly represents climate seasonality, with high loadings by temperature range as well as temperature and precipitation seasonality. Using the provenance scores of Clim1 and Clim2, we constructed two Euclidian climatic distance matrices, one for each of the two climate PCs. We then correlated the matrices of genetic differentiation (AMOVA-derived pairwise F_{ST}) for each species with the matrices of geographic and climatic distances, and tested for significance with (partial) Mantel tests.

Finally, we tested for signatures of selection and regional adaptation at the level of individual AFLP loci in each species through genome scans that tested for departure from a neutral model, while accounting for population structure, using BayeScan v2.1 (Fischer *et al.* 2011). We used default parameter settings except for the F_{IS} beta prior, which was collected from the literature (Table S3).

Results

DIFFERENTIATION AMONG PROVENANCES

The only species in which we detected different ploidy levels was *Knautia*, where three provenances (4, 7, 8) were diploids, while all other provenances were tetraploids. We therefore treated 2×- and 4×-*Knautia* separately in some of the subsequent analyses.

Principal coordinate analysis (PCoA) of AFLP markers indicated strong separation of provenances in *Galium*, *Hypochaeris*, *Knautia* and *Lychnis*, whereas there was much more overlap among provenances in *Arrhenatherum*, *Centaurea* and *Daucus* (Fig. 2). In *Knautia*, there was a strong separation between the diploid and tetraploid provenances.

Overall population differentiation was significant ($P < 0{\cdot}001$) in all species as revealed by AMOVA

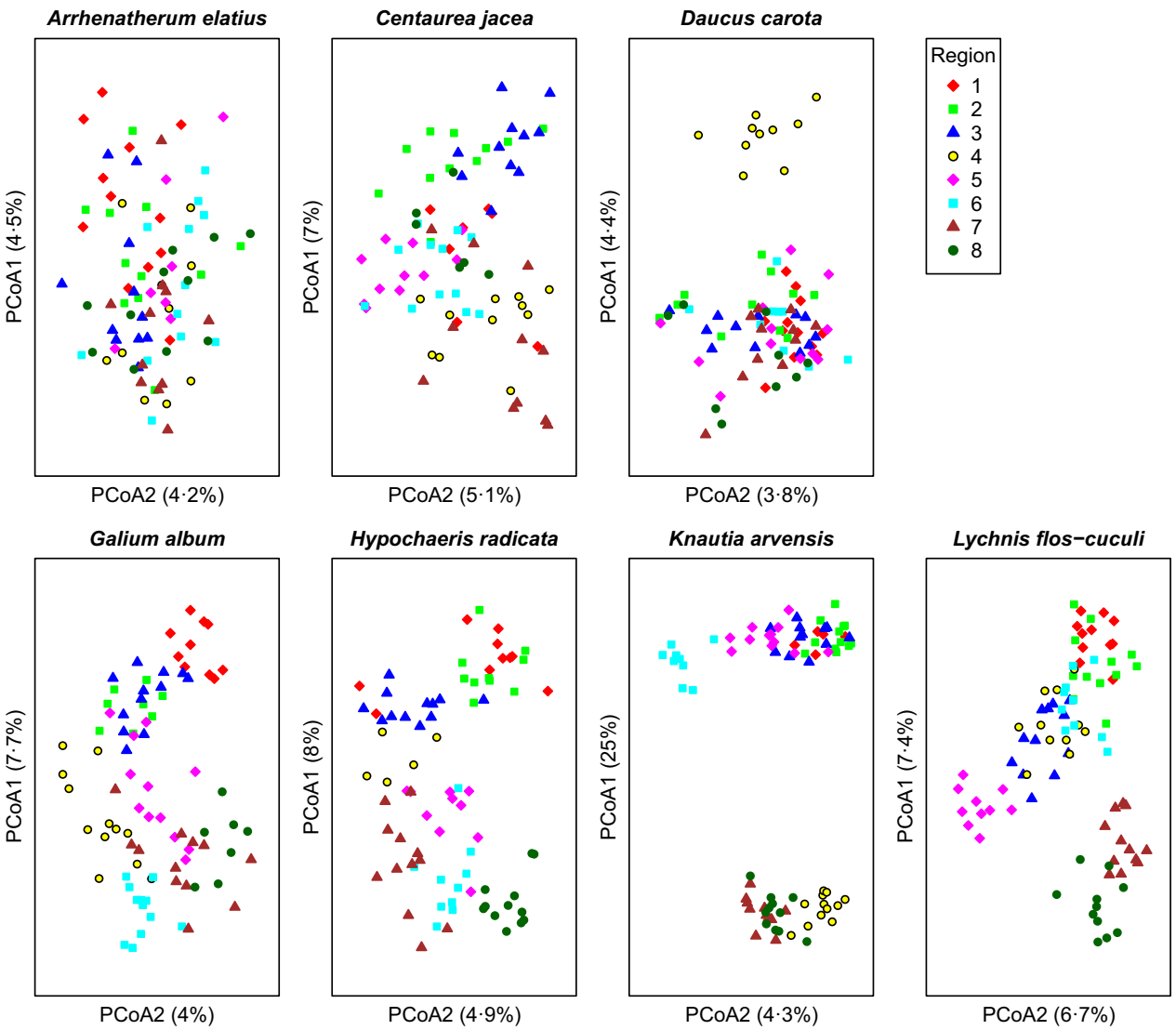

Fig. 2. PCoA plots of all studied individuals, based on their amplified fragment length polymorphism genotypes, with different colours representing provenances from the eight regions.

(Table 1). The percentage of molecular variance among provenances ranged from 4% in *Arrhenatherum* to 25% in *Lychnis*. In *Knautia*, differentiation amounted to 34% in an overall analysis, but decreased to 14% (in each cytotype) when cytotypes were analysed separately. Pairwise differentiation between provenances was significant ($P < 0{\cdot}05$) in almost all cases except for five pairs in *Arrhenatherum* (2/3, 2/5, 3/4, 5/7 and 6/7) and two pairs in *Daucus* (5/6 and 7/8; Table 2).

Bayesian cluster analysis with STRUCTURE revealed species-specific patterns (Table 2, Appendix S2). In *Arrhenatherum* and *Daucus*, we found only two clusters, and in both cases, the second cluster was represented by only one provenance. Four species (*Galium*, *Hypochaeris*, *Knautia* and *Lychnis*) showed a hierarchical genetic structure with two main clusters that contained further subclusters, resulting in a total of six to eight clusters. The two main clusters generally separated northern from southern provenances, albeit with a different line of separation for each species (Table 2, Fig. 3). In *Knautia*, the two main clusters corresponded to the two ploidy levels. In *Centaurea*, we found seven clusters. Although not all regions represented unique clusters for all species, each of the eight regions represented a unique cluster of at least some of the species (Table 2). This uniqueness ranged from two times in region 2, which thus was the least genetically distinct region, to six times in region 4. The separation of provenances into gene pools was not clear-cut in some cases (see Appendix S2), with admixture or mixture occurring mostly between adjacent regions (*Centaurea*: regions 2/5, 4/7, 7/8; *Galium*: 4/5, 6/7; *Hypochaeris*: 1/3, 4/5; *Knautia* (4x): 3/5, *Lychnis*: 3/6), and for *Arrhenatherum* across all regions.

Population differentiation followed an isolation-by-distance pattern in *Arrhenatherum, Galium, Hypochaeris* and *Knautia* (4x), but not so in the other species (all $P > 0{\cdot}3$; Fig. 4). We found no isolation by environment with Clim1, the temperature and precipitation gradient, in any of the species (Table 3). However, there were significant isolation-by-environment patterns with Clim2, that is climate seasonality, in *Arrhenatherum* and *Hypochaeris*, and a marginally significant correlation in *Galium* (Fig. 4). However, Clim2 was significantly correlated with geographic distance, and partial Mantel tests revealed that only in *Arrhenatherum*, a statistically independent effect of Clim2 remained after controlling for geographic distance (Table 3). For *Centaurea, Daucus* and *Lychnis*, we detected neither IBD nor IBE.

The genome scans detected loci putatively under selection in five species (Table S4). No outlier loci were identified in *Arrhenatherum* and *Hypochaeris*, one locus was found in *Centaurea* and *Daucus*, two in *Galium* and three in *Lychnis*, indicating differential selection among regions. While 17 loci were found in a combined analysis of diploid and tetraploid *Knautia*, none were detected in the single cytotypes.

Table 1. Descriptive statistics of population differentiation among seed transfer zones of seven common grassland plants in Central Europe

Species	Overall F_{ST}	Number of significant pairwise F_{ST} values	Range of pairwise F_{ST} values
Arrhenatherum elatius	0·035	23/28	0·003–0·073
Centaurea jacea	0·154	28/28	0·060–0·205
Daucus carota	0·066	26/28	0·010–0·174
Galium album	0·148	28/28	0·056–0·290
Hypochaeris radicata	0·170	28/28	0·078–0·279
Knautia arvensis (2x + 4x)	0·344	28/28	0·076–0·476
2x	0·139	3/3	0·104–0·180
4x	0·142	10/10	0·076–0·269
Lychnis flos-cuculi	0·246	28/28	0·153–0·373

Table 2. Assignment of regions to genetic clusters (1, 2, … 7) or subclusters (1a, 1b, …) through Bayesian cluster analyses with STRUCTURE (see also Appendix S2). The colours correspond to Fig. 3

Region	1	2	3	4	5	6	7	8	$N_{cluster}$*
Arrhenatherum	1	2	2	2	2	2	2	2	2
Centaurea	1	2	2	3	4	5	6	7	7
Daucus	1	1	1	2	1	1	1	1	2
Galium	1a	1b	1c	2a	2b	2c	2b	2b	6
Hypochaeris	1a	1a	1b	1c	2a	2b	2b	2c	6
Knautia	1a	1a	1b	2a	1c	1d	2b	2c	7
Lychnis	1a	1b	1c	1d	1e	1f	2a	2b	8
Region uniqueness†	4	2	4	6	5	4	3	4	

*$N_{cluster}$ total number of clusters and subclusters.
†The number of cases (out of seven species) for which a region is a separate cluster or subcluster and thus represents a unique gene pool.

Discussion

GENETIC DIFFERENTIATION AMONG PROVENANCES

All of the seven investigated species showed a significant genetic differentiation among most provenances. Species-level differentiation ranged between $F_{ST} = 0{\cdot}04$ and 0·25, a range expected based on reviews of dominant marker diversity (Nybom 2004; Reisch & Bernhardt-Römermann 2014), with lowest values in the wind-pollinated grass (*Arrhenatherum*), intermediate levels in insect-pollinated outcrossing herbs and highest levels in the self-compatible and insect-pollinated *Lychnis*.

The study species showed species-specific patterns of provenance clustering. However, in at least four species, there was a separation between northern and southern provenances, resembling phylogeographic patterns found in other species (Balfourier, Imbert & Charmet 2000;

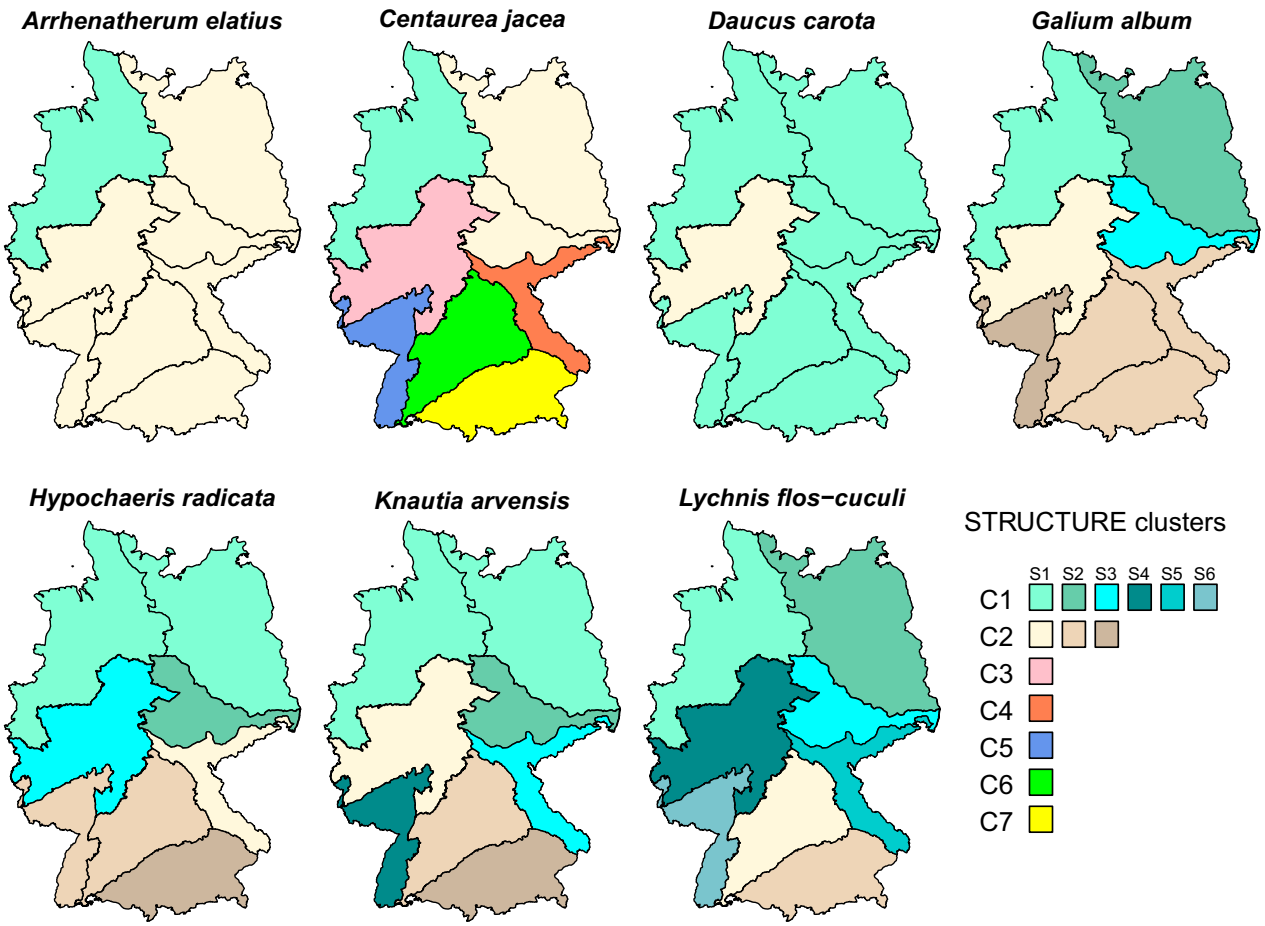

Fig. 3. Clusters and subclusters as identified in the Bayesian cluster analysis mapped onto the eight regions. Note, however, that only one seed transfer zone per region was studied (see Fig. 1). For detailed individual-level results, see Appendix S2.

Harter, Jentsch & Durka 2015). Although such phylogeographic structure is likely related to post-glacial colonization history, it can be paralleled by local adaptation (Frei *et al*. 2012).

Isolation by distance was more common than isolation by environment. We found IBD in four out of seven species, at scales of 200–800 km. At such large scales, IBD likely mirrors dispersal limitation during long periods of time, including post-glacial recolonization (Treier & Müller-Schärer 2011). Interestingly, in a previous study on *Hypochaeris* within a fragmented landscape, IBD extended only up to 3·5 km, and beyond that populations were effectively isolated (Mix *et al*. 2006). This discrepancy, however, is no contradiction as such small-scale patterns may be due to more recent and local anthropogenic changes such as habitat fragmentation.

The observed predominance of IBD across the investigated species is in line with Sexton, Hangartner & Hoffmann (2014) who found that in plants IBD is more common than IBE. We found IBE driven by climate seasonality in only three species. However, because of the cross-correlation of climate and geographic distance, it is difficult to disentangle IBE from IBD (but see Wang & Bradburd 2014). Similarly, clinal variation of other environmental factors such as topography and geology, with pleistocene lowlands in the north of Germany and geologically older uplands in the south, covaries with geographic distance. *Arrhenatherum* showed both IBD and IBE, which is surprising because it had the lowest overall level of genetic differentiation. This strongly suggests that we must be cautious with interpreting overall levels of genetic differentiation and that there can be local adaptation despite seeming genetic homogeneity (McKay *et al*. 2001).

Three of our study species – *Centaurea*, *Daucus* and *Lychnis* – did not show IBD at all, indicating that they were not in gene flow–drift equilibrium (Hutchison & Templeton 1999). For *Daucus*, pairwise F_{ST} values were small and generally rather similar, which indicates small relative drift effects, efficient gene flow and/or large population sizes that are not prone to drift (but see below). In contrast, for *Centaurea* and especially *Lychnis*, the large scatter of F_{ST} values indicates that the effects of random genetic drift are not outweighed by gene flow, resulting in unpredictable and strong genetic isolation among provenances.

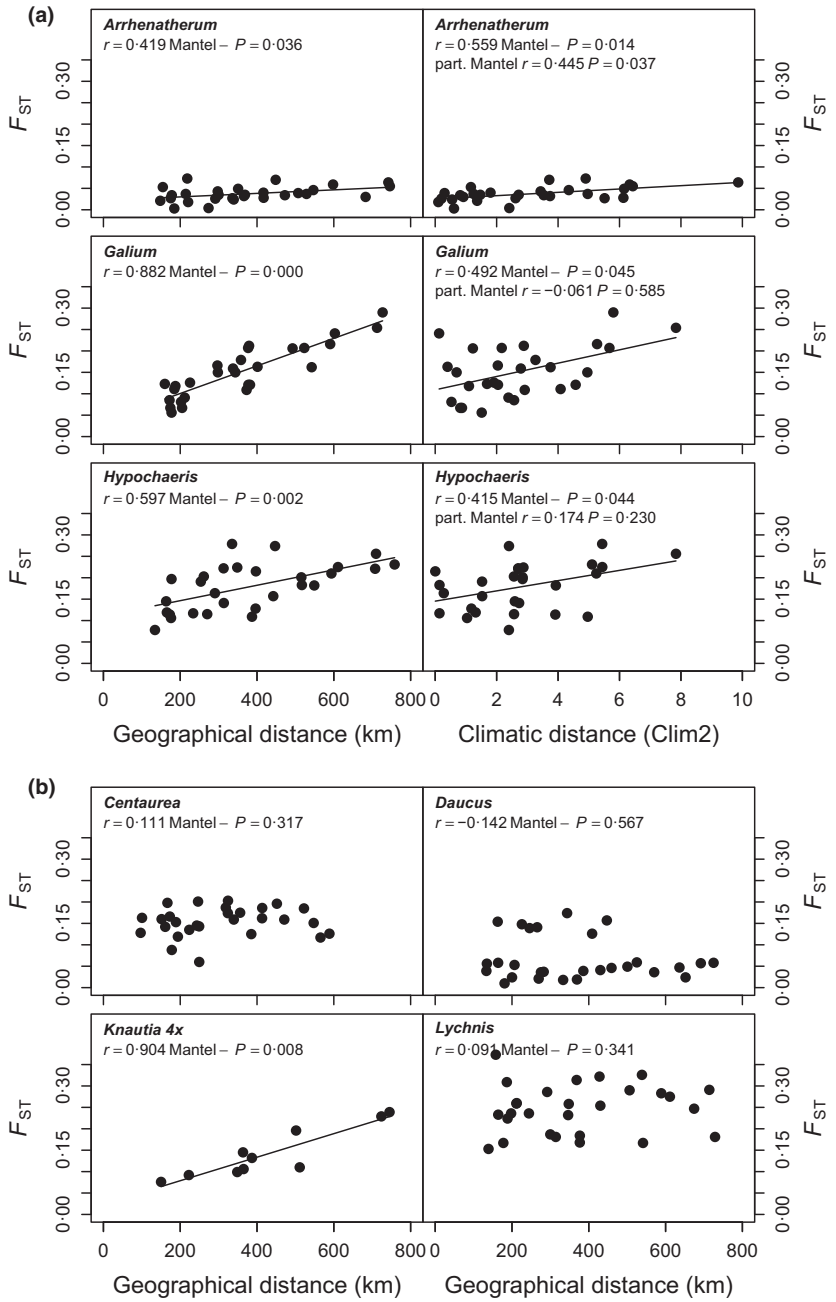

Fig. 4. Genetic differentiation (F_{ST}) as a function of geographic (isolation by distance [IBD]) or environmental (isolation by environment [IBE]) distance among provenances from different seed transfer zones of common grassland plants in Germany. In (a) both IBD and IBE are shown for species that show a significant IBE; in (b) only IBD patterns are shown for the other species.

SPECIES-SPECIFIC PATTERNS WITH GENERAL IMPLICATIONS

The grass *Arrhenatherum* showed low overall genetic differentiation, which is not unexpected for an outcrossing and wind-pollinated species (see also Michalski *et al.* 2010). However, another possible cause of this genetic homogeneity is that *Arrhenatherum* became abundant in Europe only quite recently. Although the species was hypothesized to be not native at all to Central Europe (Poschlod & WallisDeVries 2002), *Arrhenatherum* appears to be native, but was rather rare prior to the increase in fertilized hay meadows in the early 18th century (Hejcman *et al.* 2013). Since then, however, *Arrhenatherum* and other forage grass species were likely managed by on-farm propagation of local landraces, or by sowing of commercial seed stock (Kauter 2001). The most important cultivation area for *Arrhenatherum* seeds was in south-east France (e.g. Young 1792 cited in Kauter 2001), and such genotypes of 'French Ryegrass' were marketed across Europe and could have contributed to the rather homogeneous current gene pools (Kauter 2001). Note, however, that we only used seeds collected from conservation sites without seed addition for at least 40 years and likely much longer. Therefore, even if seeds had been sown historically, there was some time for regional genetic differentiation and adaptation.

Table 3. Mantel statistics r of Mantel or partial Mantel tests examining the association between genetic differentiation among provenances (F_{ST}) and geographic distance (Geo), or one of two climatic distances (Clim1 and Clim2)

Test[†]	*Arrhenatherum*	*Centaurea*	*Daucus*	*Galium*	*Hypochaeris*	*Knautia* (2x + 4x)	*Knautia* (4x)	*Lychnis*
F_{ST} ~ Geo	**0·419***	0·111	−0·142	**0·841*****	**0·597****	0·086	**0·904****	0·091
F_{ST} ~ Clim1	0·141	−0·083	−0·282	−0·099	−0·199	0·002	−0·133	0·058
Geo ~ Clim1	0·086	0·112	−0·016	−0·187	−0·191	−0·041	−0·069	−0·238
F_{ST} ~ Geo (Clim1)	**0·413***	0·121	−0·153	**0·842*****	**0·581****	0·086	**0·905****	0·108
F_{ST} ~ Clim1 (Geo)	0·116	−0·096	−0·287	0·109	−0·108	0·006	−0·166	0·083
F_{ST} ~ Clim2	**0·559***	0·063	0·007	0·407	**0·415***	−0·016	0·133	0·102
Geo ~ Clim2	0·499	0·083	**0·653****	**0·583***	**0·492***	0·498·	0·358	**0·602****
F_{ST} ~ Geo (Clim2)	0·195	0·106	−0·193	**0·814*****	**0·495***	0·109	**0·926****	0·038
F_{ST} ~ Clim2 (Geo)	**0·445***	0·054	0·133	−0·192	0·174	−0·068	−0·478	0·059

***$P < 0·001$, **$P < 0·01$, *$P < 0·05$, ·$P < 0·1$. Bold values are considered significant.
[†]F_{ST} ~ Clim1 (Geo) denotes a partial Mantel test in which the partial correlation between F_{ST} and Clim1 is tested after accounting for geographic distance.

In *Daucus*, one provenance turned out to be differentiated from all others. This could either indicate strong effects of genetic drift, for example, a population bottleneck in this outlier provenance. However, as *Daucus carota* is also a cultivated species, for which gene flow into natural populations has been observed (Iorizzo *et al.* 2013), introgression from cultivated carrot might be possible. However, preliminary AFLP analyses showed no indications of introgression of carrots (data not shown), suggesting a demographic cause. This case, however, shows that it is important to carefully select and scrutinize source populations used for seed collection.

The observed two cytotypes of *Knautia* were known before, but their geographic distribution within Germany is still little understood. Our results indicate that diploids are more widespread than previously hypothesized (Kolář *et al.* 2009). However, more detailed studies are needed to assess the distribution of cytotypes at smaller scales and the presence of mixed ploidy populations. More generally, these results show that species with several cytotypes can show stronger and more complex patterns of genetic differentiation, which further stresses the importance of appropriate seed transfer zones for their management.

MOLECULAR MARKERS VS. ADAPTIVE TRAITS

Molecular markers such as AFLP are anonymous and mostly neutral and thus do not represent the functional genome. They have been criticized as ineffective for studying local adaptation, which is best analysed at the phenotypic level (McKay *et al.* 2005). However, first, we found a number of genetic markers that departed from a neutral model, indicating a potential role in regional adaptation. Secondly, phenotypic differentiation and local adaptation may also be affected by drift or constrained by gene flow (Lenormand 2002), resulting in neutral phenotypic divergence. On the other hand, local adaptation may lead to neutral divergence and patterns of IBE (Nosil, Funk & Ortiz-Barrientos 2009), even at small spatial scales (Shi *et al.* 2011). While a comparative analysis of molecular marker divergence and adaptive phenotypic divergence in our study species is out of the scope of this paper, there is a reasonable match. Across species, the overall phenotypic divergence of biomass production among provenances expressed as P_{ST} (see Bucharova *et al.* 2016) is significantly correlated with overall genetic differentiation (F_{ST}, $r^2 = 0·72$, $P = 0·016$). This indicates that trait divergence and genetic differentiation are at least partly driven by the same processes. Ultimately, only a combined analysis of neutral and adaptive divergence will allow for a more nuanced understanding of genetic and evolutionary processes.

PRACTICAL IMPLICATIONS FOR SEED TRANSFER ZONES

The current system of regional seed use in Germany (Prasse, Kunzmann & Schröder 2010; ErMiV 2011) combines, to some degree, both sides of the 'Mix or Match' debate (Lesica & Allendorf 1999) that is related to various other seed-sourcing strategies, such as 'composite' or 'predictive' seed sourcing that are currently discussed (Williams, Nevill & Krauss 2014). First, the use of seeds is restricted to within one of 22 seed transfer zones, because species are supposed to be regionally adapted. Secondly, within each seed transfer zone, the mixing of seeds from several source populations is pursued (or at least recommended), to encompass different locally adapted populations and to increase genetic variation and restoration success. So it is in fact a 'Mix within Match' strategy. Our results of genetically differentiated provenances, often showing isolation by distance or isolation by climate, clearly support this strategy. However, in fact only 21% of the provenances used in our study were mixed from more than one source site, indicating the necessity for the seed producers to broaden the spatial and genetic basis of propagated populations. Notwithstanding such regional differentiation, the studied species likely harbour additional, more local adaptive differentiation, for example related to different levels of soil pH or soil moisture (Bischoff *et al.* 2006; Raabova, Münzbergova & Fischer 2007). Thus, because the eight investigated seed

transfer zones per species only represent a part of the system, the other, geographically intermediate, seed transfer zones, as well as multiple source populations within each zone, likely add more complexity to the patterns of genetic variation and will help to conserve a substantial part of intraspecific genetic variation.

Across the studied species, the patterns of genetic differentiation were species specific, as was also found in other cross-species analyses of plant population structure (Jorgensen *et al*. 2014). This indicates that, in theory, each species could be managed with a species-specific number and extent of seed transfer zones. However, with over 150 species that are currently commercially produced for ecological restoration in Germany (Rieger, Feucht & Wieden 2014), such species-specific management seems unfeasible. Legally, the seed transfer zones are identically defined for all grassland species (ErMiV 2011; but see Wieden 2015 for practical considerations). We found that most of the studied seed transfer zones representing a region in fact also represented a unique gene pool in many species. This indicates that although individual species could be managed with a smaller number of zones, across multiple species a larger number of zones are appropriate for maintaining genetic variation in the majority of species. Given the species-specific patterns of genetic differentiation, it is likely that for the 150 grassland species currently managed, not only eight regions as investigated here, but also the current number of 22 seed transfer zones is justified. However, although we covered the whole of Germany through the eight regions, our study included only about one-third of the seed transfer zones and less than 10% of the currently marketed species. Thus, we cannot discuss the geographic extent of individual seed transfer zones and differentiation patterns within regions or even within seed transfer zones. This would require a much more in-depth analysis. Ideally, several source populations within each seed transfer zone applying a grid-based sampling would have to be analysed across several species for a thorough assessment of the spatial scales of genetic differentiation (Malaval *et al*. 2010; Michalski & Durka 2012). Thus, it would also be very desirable to conduct similar analyses as presented here for all of the 150 grassland species used, to put the entire seed provenancing system on a solid empirical basis.

In conclusion, our study shows that grassland plants are indeed strongly genetically differentiated across Germany, supporting the strategy of seed transfer zones for ecological restoration. Although the predefined seed transfer zones are unlikely to match the exact genetic structure of many species, they serve their purpose by capturing a substantial amount of intraspecific genetic variation across species.

Acknowledgements

We are grateful to Ernst Rieger, Hans Krimmer, Matthias Stolle and Markus Wieden for providing seeds and accession details, to Valeska Vees for laboratory work and to Dierk Kunzmann for GIS data.

Data accessibility

AFLP data: Dryad Digital Repository doi:10.5061/dryad.821b5 (Durka *et al*. 2016).

References

Bakker, J.P., van Diggelen, R., Bekker, R.M. & Marrs, R.H. (2012) Restoration of dry grasslands and heathlands. *Restoration Ecology: The New Frontier* (eds J. van Andel & J. Aronson), pp. 173–187. John Wiley & Sons, Chichester.

Balfourier, F., Imbert, C. & Charmet, G. (2000) Evidence for phylogeographic structure in *Lolium* species related to the spread of agriculture in Europe. A cpDNA study. *Theoretical and Applied Genetics*, **101**, 131–138.

Bischoff, A., Steinger, T. & Müller-Schärer, H. (2010) The importance of plant provenance and genotypic diversity of seed material used for ecological restoration. *Restoration Ecology*, **18**, 338–348.

Bischoff, A., Cremieux, L., Smilauerova, M., Lawson, C.S., Mortimer, S.R., Dolezal, J. *et al*. (2006) Detecting local adaptation in widespread grassland species – the importance of scale and local plant community. *Journal of Ecology*, **94**, 1130–1142.

Bower, A.D., Clair, J.B.S. & Erickson, V. (2014) Generalized provisional seed zones for native plants. *Ecological Applications*, **24**, 913–919.

Bucharova, A., Michalski, S.G., Hermann, J.M., Heveling, K., Durka, W., Hölzel, N., Kollmann, J. & Bossdorf, O. (2016) Genetic differentiation and regional adaptation among seed origins used for grassland restoration: lessons from a multi-species transplant experiment. *Journal of Applied Ecology*, doi: 10.1111/1365-2664.12645.

De Kort, H., Mergeay, J., Vander Mijnsbrugge, K., Decocq, G., Maccherini, S., Kehlet Bruun, H.H., Honnay, O. & Vandepitte, K. (2014) An evaluation of seed zone delineation using phenotypic and population genomic data on black alder *Alnus glutinosa*. *Journal of Applied Ecology*, **51**, 1218–1227.

Durka, W., Michalski, S.G., Berendzen, K.W., Bossdorf, O., Bucharova, A., Hermann, J.-M., Hölzel, N. & Kollmann, J. (2016) Data from: Genetic differentiation within multiple common grassland plants supports seed transfer zones for ecological restoration. *Dryad Digital Repository*, http://dx.doi.org/10.5061/dryad.821b5.

ErMiV (2011) Verordnung über das Inverkehrbringen von Saatgut von Erhaltungsmischungen (Erhaltungsmischungsverordnung). *Bundesgesetzblatt*, **Teil I, Nr. 65**, 2641–2646.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction sites. *Genetics*, **131**, 479–491.

Falush, D., Stephens, M. & Pritchard, J.K. (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Fischer, M.C., Foll, M., Excoffier, L. & Heckel, G. (2011) Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (*Microtus arvalis*). *Molecular Ecology*, **20**, 1450–1462.

Fjellheim, S., Rognli, O.A., Fosnes, K. & Brochmann, C. (2006) Phylogeographical history of the widespread meadow fescue (*Festuca pratensis* Huds.) inferred from chloroplast DNA sequences. *Journal of Biogeography*, **33**, 1470–1478.

FoVHgV (2003) Verordnung über Herkunftsgebiete für forstliches Vermehrungsgut (Forstvermehrungsgut-Herkunftsgebietsverordnung – FoVHgV). *Bundesgesetzblatt*, **I**, 238.

Frei, E.S., Scheepens, J.F., Armbruster, G.F.J. & Stöcklin, J. (2012) Phenotypic differentiation in a common garden reflects the phylogeography of a widespread Alpine plant. *Journal of Ecology*, **100**, 297–308.

Harter, D.E.V., Jentsch, A. & Durka, W. (2015) Holocene re-colonisation, central-marginal-distribution and habitat specialisation shape population genetic patterns within an Atlantic European grass species. *Plant Biology*, **17**, 684–693.

Hejcman, M., Hejcmanová, P., Pavlů, V. & Beneš, J. (2013) Origin and history of grasslands in Central Europe – a review. *Grass and Forage Science*, **68**, 345–363.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hufford, K.M. & Mazer, S.J. (2003) Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution*, **18**, 147–155.

Hutchison, D.W. & Templeton, A.R. (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898–1914.

Iorizzo, M., Senalik, D.A., Ellison, S.L., Grzebelus, D., Cavagnaro, P.F., Allender, C. *et al*. (2013) Genetic structure and domestication of carrot (*Daucus carota* subsp. *sativus*) (Apiaceae). *American Journal of Botany*, **100**, 930–938.

Jakobsson, M. & Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jorgensen, M.H., Elameen, A., Klemsdal, S. & Fjellheim, S. (2014) Use of molecular markers for defining site specific seed material for restoration in Norway. *Guidelines for Native Seed Production and Grassland Restoration* (eds K. Kiehl, A. Kirmer, N. Shaw & S. Tischew), pp. 57–74. Cambridge Scholars Publishing, Newcastle upon Type.

Kauter, D. (2001) "Sauergras" and "Wegbreit" - The development of meadows in Central Europe between 1500 and 1900 [in German]. *Berichte des Institutes für Landschafts- und Pflanzenökologie der Universität Hohenheim. Beiheft*, **14**, 226 p.

Kiehl, K., Kirmer, A., Shaw, N. & Tischew, S. (2014) *Guidelines for Native Seed Production and Grassland Restoration*. Cambridge Scholars Publishing, Newcastle upon Tyne.

Kloss, L., Fischer, M. & Durka, W. (2011) Land-use effects on genetic structure of a common grassland herb: a matter of scale. *Basic and Applied Ecology*, **12**, 440–448.

Köhler, C., Mittelsten Scheid, O. & Erilova, A. (2010) The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics*, **26**, 142–148.

Kolář, F., Štech, M., Trávníček, P., Rauchova, J., Urfus, T., Vit, P., Kubesova, M. & Suda, J. (2009) Towards resolving the *Knautia arvensis* agg. (Dipsacaceae) puzzle: primary and secondary contact zones and ploidy segregation at landscape and microgeographic scales. *Annals of Botany*, **103**, 963–974.

Krauss, S.L., Sinclair, E.A., Bussell, J.D. & Hobbs, R.J. (2013) An ecological genetic delineation of local seed-source provenance for ecological restoration. *Ecology and Evolution*, **3**, 2138–2149.

Leimu, R. & Fischer, M. (2008) A meta-analysis of local adaptation in plants. *PLoS One*, **3**, e4010.

Lenormand, T. (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.

Lesica, P. & Allendorf, F.W. (1999) Ecological genetics and the restoration of plant communities: mix or match? *Restoration Ecology*, **7**, 42–50.

Malaval, S., Lauga, B., Regnault-Roger, C. & Largier, G. (2010) Combined definition of seed transfer guidelines for ecological restoration in the French Pyrenees. *Applied Vegetation Science*, **13**, 113–124.

McKay, J.K., Bishop, J.G., Lin, J.Z., Richards, J.H., Sala, A. & Mitchell-Olds, T. (2001) Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **268**, 1715–1721.

McKay, J.K., Christian, C.E., Harrison, S. & Rice, K.J. (2005) "How local is local?" – a review of practical and conceptual issues in the genetics of restoration. *Restoration Ecology*, **13**, 432–440.

Meynen, E. & Schmithüsen, J. (1953-1962) Handbuch der naturräumlichen Gliederung Deutschlands. Selbstverlag der Bundesanstalt für Landeskunde, Bad Godesberg.

Michalski, S.G. & Durka, W. (2012) Assessment of provenance delineation by genetic differentiation patterns and estimates of gene flow in the common grassland plant *Geranium pratense*. *Conservation Genetics*, **13**, 581–592.

Michalski, S.G., Durka, W., Jentsch, A.V., Kreyling, J., Pompe, S., Schweiger, O., Willner, E. & Beierkuhnlein, C. (2010) Evidence for genetic differentiation and divergent selection in an autotetraploid forage grass (*Arrhenatherum elatius*). *Theoretical and Applied Genetics*, **120**, 1151–1162.

Millar, M.A., Byrne, M. & Coates, D.J. (2008) Seed collection for revegetation: guidelines for Western Australian flora. *Journal of the Royal Society of Western Australia*, **91**, 293–299.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis*. Island Press, Washington, DC.

Miller, S.A., Bartow, A., Gisler, M., Ward, K., Young, A.S. & Kaye, T.N. (2011) Can an ecoregion serve as a seed transfer zone? Evidence from a common garden study with five native species. *Restoration Ecology*, **19**, 268–276.

Mix, C., Arens, P.F.P., Rengelink, R., Smulders, M.J.M., van Groenendael, J.M. & Ouborg, N.J. (2006) Regional gene flow and population structure of the wind-dispersed plant species *Hypochaeris radicata* (Asteraceae) in an agricultural landscape. *Molecular Ecology*, **15**, 1749–1758.

Nosil, P., Funk, D.J. & Ortiz-Barrientos, D. (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375–402.

Nybom, H. (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, **13**, 1143–1155.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J. & De Meester, L. (2013) Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, **22**, 5983–5999.

Peakall, R. & Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, **28**, 2537–2539.

Poschlod, P. & WallisDeVries, M.F. (2002) The historical and socioeconomic perspective of calcareous grasslands – lessons from the distant and recent past. *Biological Conservation*, **104**, 361–376.

Prasse, R., Kunzmann, D. & Schröder, R. (2010) *Entwicklung und praktische Umsetzung naturschutzfachlicher Mindestanforderungen an einen Herkunftsnachweis für gebietseigenes Wildpflanzensaatgut krautiger Pflanzen*. Abschlußbericht zum Forschungsprojekt (DBU FKZ: 23931), Hannover.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Raabova, J., Münzbergova, Z. & Fischer, M. (2007) Ecological rather than geographic or genetic distance affects local adaptation of the rare perennial herb, *Aster amellus*. *Biological Conservation*, **139**, 348–357.

Reisch, C. & Bernhardt-Römermann, M. (2014) The impact of study design and life history traits on genetic variation of plants determined with AFLPs. *Plant Ecology*, **215**, 1493–1511.

Rieger, E., Feucht, B. & Wieden, M. (2014) Agricultural propagation of native seeds and development of a certification procedure in Germany. *Guidelines for Native Seed Production and Grassland Restoration* (eds K. Kiehl, A. Kirmer, N. Shaw & S. Tischew), pp. 101–116. Cambridge Scholars Publishing, Newcastle upon Type.

Sackville Hamilton, N.R. (2001) Is local provenance important in habitat creation? A reply. *Journal of Applied Ecology*, **38**, 1374–1376.

Sexton, J.P., Hangartner, S.B. & Hoffmann, A.A. (2014) Genetic isolation by environment or distance: which pattern of gene flow is most common? *Evolution*, **68**, 1–15.

Sgrò, C.M., Lowe, A.J. & Hoffmann, A.A. (2011) Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, **4**, 326–337.

Shi, M., Michalski, S.G., Chen, X.Y. & Durka, W. (2011) Isolation by elevation: genetic structure at neutral and putatively non-neutral loci in a dominant tree of subtropical forest, *Castanopsis eyrei*. *PLoS One*, **6**, e21302.

SKEW (2009) Empfehlungen für den Anbau und die Verwendung von Pflanz- und Saatgut einheimischer Wildpflanzen. www.cps-skew.ch.

St Clair, J.B., Kilkenny, F.F., Johnson, R.C., Shaw, N.L. & Weaver, G. (2013) Genetic variation in adaptive traits and seed transfer zones for *Pseudoroegneria spicata* (bluebunch wheatgrass) in the northwestern United States. *Evolutionary Applications*, **6**, 933–948.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.G. & Cosson, J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Treier, U.A. & Müller-Schärer, H. (2011) Differential effects of historical migration, glaciations and human impact on the genetic structure and diversity of the mountain pasture weed *Veratrum album* L. *Journal of Biogeography*, **38**, 1776–1791.

Vander Mijnsbrugge, K., Bischoff, A. & Smith, B. (2010) A question of origin: where and how to collect seed for ecological restoration. *Basic and Applied Ecology*, **11**, 300–311.

Wang, I.J. & Bradburd, G.S. (2014) Isolation by environment. *Molecular Ecology*, **23**, 5649–5662.

Wieden, M. (2015) Wildpflanzensaatgut im Spannungsfeld des Naturschutzes. *Naturschutz und Landschaftsplanung*, **47**, 181–190.

Williams, A.V., Nevill, P.G. & Krauss, S.L. (2014) Next generation restoration genetics: applications and opportunities. *Trends in Plant Science*, **19**, 529–537.

Young, A. (1792) *Travels during the Years 1787, 1788, and 1789: Undertaken more particularly with a view of ascertaining the cultivation, wealth, resources, and national prosperity of the Kingdom of France*. Richardson, London.

Received 21 August 2015; accepted 24 February 2016
Handling Editor: Shiqiang Wan

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Table S1. Collection sites in the eight regions and seed transfer zones.

Table S2. Primer combinations used in AFLP analysis, number of loci, mean error rate and number of duplicates analysed.

Table S3. Mean inbreeding coefficients used as Beta prior in outlier locus analysis with BayeScan.

Table S4. AFLP outlier loci identified by BayeScan.

Appendix S1. Methods for Flow cytometry.

Appendix S2. Detailed results of the STRUCTURE analysis.