ANNALS OF
BOTANY
Founded 1887

Historical biogeography and local adaptation explain population genetic structure in a widespread terrestrial orchid

Alexandra Evans[1],[*],[iD], Hanne de Kort[1],[iD], Rein Brys[2], Karl J. Duffy[3], Jana Jersáková[4], Tiiu Kull[5], Marc-André Selosse[6],[7], Spyros Tsiftsis[8],[iD], Julita Minasiewicz[7] and Hans Jacquemyn[1]

[1]Department of Biology, Plant Conservation and Population Biology, Katholieke Universiteit Leuven, Leuven, Belgium, [2]Research Institute for Forest and Nature, Geraardsbergen, Belgium, [3]Department of Biology, University of Naples Federico II, Complesso Monte Sant'Angelo, Naples 80126, Italy, [4]Department of Biology of Ecosystems, Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic, [5]Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, 5 Kreutzwaldi, 51014 Tartu, Estonia, [6]Institut Systématique Evolution Biodiversité, Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, Paris, France, [7]Department of Plant Taxonomy and Nature Conservation, University of Gdansk, Wita Stwosza 59, 80-308 Gdansk, Poland and [8]Department of Forest and Natural Environment Sciences, International Hellenic University, GR-66132, Drama, Greece
For correspondence. E-mail alexandra.evans@kuleuven.be

• **Background and Aims** Historical changes in environmental conditions and colonization–extinction dynamics have a direct impact on the genetic structure of plant populations. However, understanding how past environmental conditions influenced the evolution of species with high gene flow is challenging when signals for genetic isolation and adaptation are swamped by gene flow. We investigated the spatial distribution and genetic structure of the widespread terrestrial orchid *Epipactis helleborine* to identify glacial refugia, characterize postglacial population dynamics and assess its adaptive potential.
• **Methods** Ecological niche modelling was used to locate possible glacial refugia and postglacial recolonization opportunities of *E. helleborine*. A large single-nucleotide polymorphism (SNP) dataset obtained through genotyping by sequencing was used to define population genetic diversity and structure and to identify sources of postglacial gene flow. Outlier analyses were used to elucidate how adaptation to the local environment contributed to population divergence.
• **Key Results** The distribution of climatically suitable areas was restricted during the Last Glacial Maximum to the Mediterranean, south-western Europe and small areas in the Alps and Carpathians. Within-population genetic diversity was high in *E. helleborine* (mean expected heterozygosity, 0.373 ± 0.006; observed heterozygosity, 0.571 ± 0.012; allelic richness, 1.387 ± 0.007). Italy and central Europe are likely to have acted as important genetic sources during postglacial recolonization. Adaptive SNPs were associated with temperature, elevation and precipitation.
• **Conclusions** Forests in the Mediterranean and Carpathians are likely to have acted as glacial refugia for *Epipactis helleborine*. Postglacial migration northwards and to higher elevations resulted in the dispersal and diversification of *E. helleborine* in central Europe and Italy, and to geographical isolation and divergent adaptation in Greek and Italian populations. Distinguishing adaptive from neutral genetic diversity allowed us to conclude that *E. helleborine* has a high adaptive potential to climate change and demonstrates that signals of adaptation and historical isolation can be identified even in species with high gene flow.

Key words: environment association analysis, environmental niche modelling, *Epipactis helleborine*, genotype–environment associations, genotyping by sequencing, landscape genomics, local adaptation, natural selection.

INTRODUCTION

Historical and ecological factors can have substantial effects on the distribution and evolution of plant and animal species and are often reflected in the large-scale spatial genetic structure of species (Comes and Kadereit, 1998; Hewitt, 2000; de Lafontaine *et al.*, 2018). Populations that survived in glacial refugia can still show signatures of allele frequencies and diversities reflecting long-term persistence and periods of genetic isolation (Tribsch, 2004; Keppel *et al.*, 2012; Reisch and Rosbakh, 2021). However, characterizing postglacial dispersal patterns for wind-dispersed, herbaceous species, such as orchids, is complicated, because their high dispersal potential but specialized habitat requirements (McCormick and Jacquemyn, 2014) result in contrasting gene flow mechanisms (Tremblay *et al.*, 2005; Givnish *et al.*, 2016). This has resulted in gaps in our understanding of the effects of climatic changes on the genetic structure, spatial distribution and evolution of ecologically specialized plants (Peakall, 2007; Givnish *et al.*, 2016).

Glacial periods resulted in the fragmentation of continuous distributions of plant species when large areas became

inhospitable, with isolated populations surviving only in the more stable habitats of glacial refugia (Ashcroft, 2010; Keppel *et al.*, 2012). In Europe, glacial macrorefugia for plant species have been identified mainly around the Mediterranean basin, an area that was relatively sheltered from the permafrost covering more northern areas during this time (Bhagwat and Willis, 2008), but also in smaller isolated areas further north, with the microhabitats of these areas acting as buffered shelters (Stewart and Lister, 2001), such as in the Alps (Tribsch, 2004; Schönswetter *et al.*, 2005; Parisod, 2022) and the Carpathians (Kolář *et al.*, 2016). Populations that expanded from glacial refugia into the newly suitable conditions of postglacial Europe can carry genetic signatures of their isolated source populations, but the evidence for patterns of refugial and non-refugial populations is inconsistent within the literature. In a study of 179 European plant species, populations of alpine species tended to be more genetically distinct but had lower genetic diversity than non-alpine species, presumably attributable, in part, to alpine populations being genetically isolated during glacial periods (Reisch and Rosbakh, 2021). In other cases, refugial populations tend to have higher genetic diversity than offshoot populations owing to founder effects (Hewitt, 2004), and more recently established populations can show stronger spatial genetic structure than older populations (Ilves *et al.*, 2015). This disparity can occur when gene flow between populations is re-established during postglacial expansion and when genetic structure that has arisen in isolated refugial populations is homogenized (Slatkin, 1987; Hutchison and Templeton, 1999; Ellstrand, 2014; García-Navas *et al.*, 2015; Fitzpatrick and Reid, 2019).

The erasure of refugial genetic structure by postglacial gene flow is more likely to occur in species with high gene dispersal capabilities (Loveless and Hamrick, 1984; Givnish, 2010). This has been demonstrated in orchids that, through their easily dispersed dust seeds, have a high potential for gene flow and thus population admixture (Phillips *et al.*, 2012). However, several studies have found substantial adaptive differentiation in plant species with high dispersal capabilities (De Kort *et al.*, 2014; Muir *et al.*, 2014; Hämälä and Savolainen, 2019). These studies demonstrated that landscape features promoting gene flow do not necessarily hamper adaptive evolution and suggest that even plants with high dispersal capabilities can exhibit genetic signals for adaptation to local environments. In orchids, this might be a result of most seeds landing near the mother plant, which would reduce gene flow despite the theoretically high potential for gene dispersal (Jersáková and Malinová, 2007), in addition to specific ecological requirements that dictate where orchids can successfully colonize and persist. Terrestrial orchid distributions in Europe are linked to substrate, elevation and climate (Djordjević and Tsiftsis, 2022) and often to the presence of pollinators and mycorrhizal fungi on which they rely to germinate (McCormick *et al.*, 2018). Differences in environmental pressures and consequent divergent selection of populations can also affect the genetic structure of plant species (Stinchcombe and Hoekstra, 2008; Nosil *et al.*, 2009; Orsini *et al.*, 2013). The habitat requirements of orchids create a potential for strong signals for adaptation, but these signals can be swamped by high gene flow resulting from a high potential for seed dispersal, which complicates inferences about the evolutionary history of orchids.

To shed light on the historical and ecological processes that have led to contemporary patterns of distribution and genetic structure of orchids, we investigated the historical distribution and genetic structure of the widespread terrestrial orchid *Epipactis helleborine*. Ecological niche modelling was used to identify possible locations of glacial refugia and postglacial recolonization opportunities. Genetic structure and diversity were calculated from a large single-nucleotide polymorphism (SNP) dataset to distinguish reproductively isolated populations from admixing populations and to identify outlier loci associated with environmental characteristics to assess whether local adaptation contributed to population divergence. Finally, we explored the relationship between genetic diversity and the change in habitat suitability from the Last Glacial Maximum (LGM) to the near-present and predicted locations of diversification sources to assess how postglacial environmental change affected genetic diversity. We tested the hypothesis that populations located in glacial refugia exhibit some signals of historical isolation, such as lower genetic diversity and presence of adaptive loci, but that postglacial colonization and increased gene flow resulted in reduced population differentiation.

MATERIALS AND METHODS

Study species

Epipactis helleborine is a widespread terrestrial orchid that occurs across Europe and Asia and has been naturalized in America. Although *E. helleborine* is a habitat generalist in comparison to other orchids, its occurrence is still limited to some extent by bedrock, deciduous tree cover, soil and climate (Hemrová *et al.*, 2019; Evans and Jacquemyn, 2020). Reproduction in *E. helleborine* can be either sexual (primarily allogamous, with pollination occurring via insects) or asexual (vegetative via rhizomes) (Delforge, 2006). The seeds produced are very small, wind-dispersed 'dust seeds' that depend on the presence of mycorrhizal fungi in the soil for successful germination and establishment (Bidartondo and Read, 2008; Jacquemyn *et al.*, 2018). The high seed-dispersal capability of *E. helleborine*, combined with its ability to colonize a relatively broad array of habitats, is thought to have given rise to a large number of subspecies or variants through temporal, ecological or geographical isolation (Delforge, 2006), but the validity of their taxonomic status warrants more detailed research (Sramkó *et al.*, 2019; Bateman, 2020).

Ecological niche model

To define the spatial distribution of suitable habitat of *E. helleborine*, we used Maxent v.3.4 (Phillips *et al.*, 2018) to create ecological niche models (ENMs) for *E. helleborine* distributed in Europe. Maxent is limited, like other presence-only modelling methods, and the use of pseudo-presences can lead to incorrect interpretation of the output (Elith *et al.*, 2011). However, when the appropriate parameter settings are used, Maxent performs well (Merow *et al.*, 2013). Occurrence data for *E. helleborine* in Europe from 2000 to 2019 was obtained from the online database GBIF (www.GBIF.org; https://doi.org/10.15468/dl.aicdym, 1 July 2019). Records with a resolution coarser than

100 m were discarded, as were those with uncertain species identification. Using QGIS v.3.4.9 (QGIS Development Team, 2019), the data were aggregated to reduce the effects of spatial clustering by creating a grid of 10 km^2 pixels and extracting the midpoint of each pixel containing at least one species record. This resulted in 6221 midpoints, which were used as the occurrence data in the model.

For the environmental variables, we used elevation from the EarthEnv database (Amatulli *et al.*, 2018), bedrock (dominant parent material) from the European Soil Database (ESDAC; Van Liedekerke *et al.*, 2006), and mean annual temperature and annual precipitation from WorldClim v.2.1 (Fick and Hijmans, 2017). Mean annual temperature and annual precipitation projected by the MIROC-ESM climate models were also obtained for the Last Interglacial period (LI), LGM and Mid-Holocene (MH), from WorldClim v.1.4 (Hijmans *et al.*, 2005; Otto-Bliesner *et al.*, 2006). We also investigated the CC projection, but given that mean precipitation or temperature for our study area did not differ significantly from the MIROC model, we continued with only the MIROC data. All environmental variables were downloaded as rasters with a spatial resolution of 1 km, except for elevation (100 m), which was resampled to 1 km. Rasters were tested for correlation using Pearson correlation coefficients and were excluded if highly correlated (correlation > 0.5). All raster processing (resampling, reprojection and cropping) was performed in R v.4 (R Core Team, 2021) using the 'raster' v.3.3 (Hijmans *et al.*, 2015) and 'RStoolbox' v.0.2.6 (Leutner *et al.*, 2017) packages.

We used the occurrence points and environmental data to run Maxent with 100 replicates, with 75 % of the occurrences used for model training and 25 % for model testing, a random seed, and the other settings at default (including a cloglog output). Area under the curve and test skill statistic values were used to assess model performance. The model was projected onto the three past climate scenarios to generate maps of habitat suitability for *E. helleborine* during the LI, LGM and MH, which were used to infer the locations of habitat that could have acted as refugia for the species during unfavourable climatic conditions.

To quantify the change in habitat of *E. helleborine* resulting from climatic changes, the mean Gibbs value (habitat suitability) of each Maxent map was calculated using QGIS v.3.4.9 (QGIS Development Team, 2019) and compared between climate scenarios with an ANOVA. Gibbs values were also extracted from a 10 km buffer around the centre point of each population that was sampled for genetic data, from which the mean habitat suitability score per population was calculated and used in regressions with genetic diversity (see below, Genetic diversity and population genetic structure). To visualize change in suitable habitat between climate scenarios, pixels above the equal training sensitivity and specificity Cloglog threshold value of the model (0.53) were extracted and plotted to represent 'highly suitable' habitat.

Sampling and genotyping by sequencing

We sampled 20 individuals from each of 23 *E. helleborine* populations in 11 European countries (Fig. 1) in June–August 2019. For the three populations in Italy, only five individuals

per population were sampled. One leaf from each plant was collected and stored in silica gel before DNA extraction (for details, see Supplementary data Text S1). A total of 288 leaves were genotyped using the LGC Genomics protocol (LGC Genomics, Teddington, UK). After removing individual samples and SNP loci that did not meet missing data requirements (for details, see Supplementary data Text S1), the remaining dataset consisted of 893 SNP loci and 206 individuals from 23 populations (9–11 individuals per population, except for Italian populations, each of which retained the original five individuals). All genomic data can be accessed at the NCBI Sequence Read Archive with the accession number SAMN19104601 or through http://www.ncbi.nlm.nih.gov/bioproject/728837.

Genetic diversity and population genetic structure

To investigate gene flow among populations, we calculated genetic diversity [allelic richness (AR) and observed and expected heterozygosity (H_o and H_e)] on a rarefied sample of SNPs using the R package 'hierfstat' v.0.5-9 (Goudet, 2005) to account for differences in sample sizes among populations. Ordinary least-squares regressions were performed to test the relationship between mean genetic diversity and habitat suitability (Gibbs value of the near-present climate) per population. To test whether postglacial environmental change might have affected genetic diversity, we also performed regressions to explore the relationships between genetic diversity and the change in habitat suitability for each population from the LGM to the near-present (the mean near-present Gibbs value minus the mean LGM Gibbs value).

GenAlEx v.6.503 (Peakall and Smouse, 2012) was used to perform an analysis of molecular variance (AMOVA) to analyse partitioning of genetic variation within and between populations and to calculate pairwise F_{ST} values (genetic distance) between populations. A Mantel test was performed to test whether F_{ST} values were correlated with $\log_{10}$-transformed geographical distances between populations, using the 'vegan' package v.2.5 (Oksanen *et al.*, 2019) in R v.4 (R Core Team, 2021) with 9999 iterations.

To assess gene flow and reproductive isolation of populations, genetic structure was characterized with STRUCTURE v.2.3.4 (Pritchard *et al.*, 2000) using 100 000 iterations, with 50 000 as burn-in, and 25 replicates for values of *K* genetic clusters = 1–24. We used an admixture model with population location included as a prior, which allows for the possibility that individuals have recent ancestors from more than one population, under the assumption of no isolation by distance. The multiple runs were consolidated using CLUMPP v.1.1.2 (Jakobsson and Rosenberg, 2007). We then implemented the Evanno method (Evanno *et al.*, 2005) to determine the most likely number of genetic clusters that exist in the dataset using the R package 'pophelper' v.2.3.0 (Francis, 2017). The membership coefficient of each sampled population to each genetic cluster was visualized using 'pophelper' and code from François (2016).

Genetic outlier detection and analyses

To distinguish signals for adaptive potential (outliers) from diversification through genetic drift (neutral SNPs), we used two
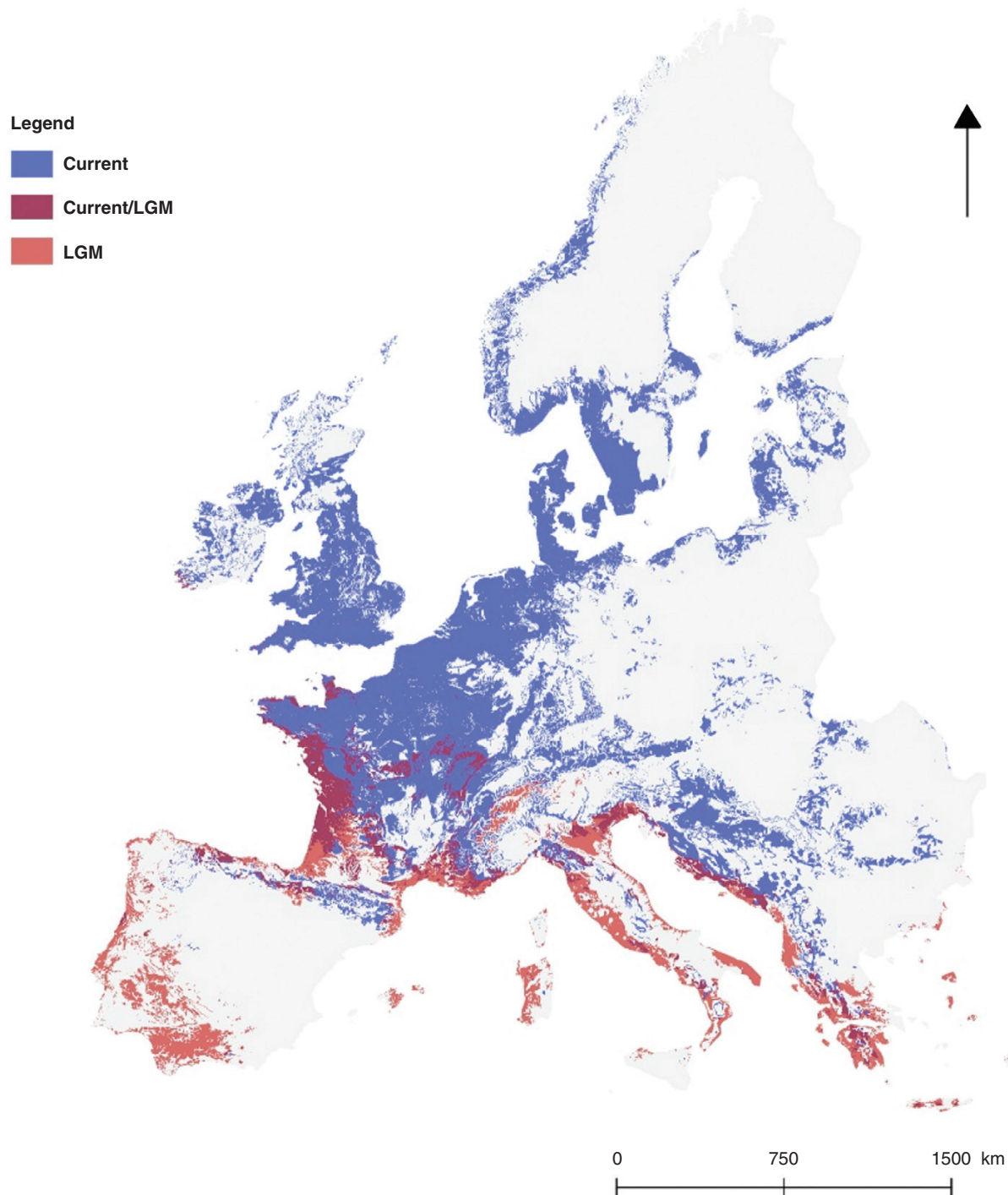
FIG. 1. European distributions of suitable habitat of *Epipactis helleborine* predicted for the current climate and the climate of the Last Glacial Maximum (LGM), showing areas of overlap between current and past climate. The continuous range of habitat probabilities of the Maxent output were converted to binary presence/absence maps by extracting and plotting the pixels above the equal training sensitivity and specificity Cloglog threshold value (0.53).

outlier detection methods and retained SNPs that were detected by both methods. The first method involved Bayescan v.2.1 (Foll and Gaggiotti, 2008), which uses a Bayesian approach to detect outliers based on genetic signatures of diversifying selection. We ran the SNP allele frequency data for 20 pilot runs of 5000 iterations, followed by 100 000 iterations, with

50 000 being used as burn-in and a thinning interval of ten. The F_{ST} results obtained were plotted in R, and outliers were identified using a false discovery rate of 0.01. For the second outlier detection method, we used the R package 'pcadapt' v.4.3.3 (Luu *et al.*, 2017). This method used a principal component analysis to summarize the variation in SNP data, with

the number of clusters detected by STRUCTURE used as the number of components, after which outlier SNPs were identified based on their estimated deviation from the mean population structure. A q-value threshold of 0.01 was again used for the proportion of false positives, below which outlier SNPs were identified. Neutral SNPs were defined as those above the q-value. The mean genetic diversity (H_e, H_o and AR) was calculated for neutral and outlier datasets separately and compared using Student's unpaired t-tests or Kruskal–Wallis tests (with the choice of test being informed by a Shapiro–Wilk test of normality). To test for signals for adaptation to postglacial environmental change, we performed ordinary least-squares regressions between change in habitat suitability from the LGM to the near-present and genetic diversity of outlier and neutral datasets separately.

To detect SNPs associated specifically with the environmental variables used in the ecological niche analyses, we performed a redundancy analysis (RDA) using the 'vegan' package in R (Oksanen *et al.*, 2019). For each continuous environmental variable raster (temperature, precipitation and elevation), values were extracted for each pixel in a 10-km-radius buffer around each population sampled. The mean value of each environmental variable for each population was calculated and multicollinearity was checked using variance inflation factor values. Variables with values below ten were retained. Allele frequencies per population were calculated by dividing counts of the major allele per population by the total major allele count. The redundancy analysis was fitted using the environmental variables as predictors and the allele frequencies of each population as dependent variables, following the method of Forester (2019). SNPs under selection were identified from the loadings of the significant RDA (constrained) axes as those with values ± 1 s.d. ($P = 0.05$) from the mean loading.

Geogenetic analyses

To identify sources of diversification and to visualize the postglacial movement of genetic material between populations, we used the package 'spacemix' v.0.13 (Bradburd *et al.*, 2016). This software uses allele frequency covariance and sample coordinates to model gene flow among populations. The resulting 'geogenetic' maps aid in visualizing genetic distances between populations in a spatial context and inferring long-distance dispersal events and sources of admixture. We used count data in a 'source-and-target' model (with admixture), which was run for ten million Markov chain Monte Carlo generations, with a sample frequency (thinning rate) of 100 after a burn-in of ten repetitions of 100 000 generations. Chain mixing and model adequacy were assessed visually using trace plots and plots of sample covariance vs. parametric covariance, as described in the package vignette.

RESULTS

Ecological niche models

The mean area under the curve value ± s.e. for the Maxent runs was 0.80 ± 0.021, and the true skill statistic was 0.51, indicating

decent model fit for a widespread species with a large background area. Gibbs values (relative habitat suitability scores) for the near-present climate ranged from 0.208 ± 0.011 (population IT-A) to 0.729 ± 0.007 (population BE-A), with a mean of 0.457 ± 0.010 (Supplementary data Table S1). Precipitation contributed 31.5 % to the mean ecological niche of *E. helleborine*, followed by temperature (27.3 %), bedrock (25.0 %) and elevation (16.2 %).

The suitable habitat of *E. helleborine* in the LI and the current distribution was widely distributed across Europe, with the most suitable habitat situated in the north-west (Fig. 1; Supplementary data Appendix S1). During the LGM, the suitability of habitat was reduced significantly [change (Δ) in mean Gibbs value = −0.292, $Z = −4.160$, $P < 0.001$], with most of northern and eastern Europe becoming unsuitable (Fig. 1). Glacial refugia, or areas of suitable habitat during the LGM, were located on the Mediterranean coast (particularly Greek and Italian peninsulas and southern France), the western coasts of France and the Iberian peninsula, and small isolated areas in the Alps and central European Carpathians. Between the LGM and near-current climate, habitat returned to a similar mean suitability and distribution to that in the LI period (Fig. 1; Δ mean Gibbs value = 0.334, $Z = 4.990$, $P < 0.001$).

Genetic diversity and population genetic structure

Allelic richness ranged from 1.353 ± 0.008 s.e. (population GK-C) to 1.403 ± 0.008 (population IT-B) (Supplementary data Table S1). Observed heterozygosity (H_o) values were between 0.551 ± 0.012 (population IT-C) and 0.592 ± 0.012 (population FR-B) and expected heterozygosity (H_e) between 0.340 ± 0.007 (population GK-C) and 0.386 ± 0.006 (population ES-C). Greek, Italian and Danish populations tended to have measures of within-population genetic diversity below the mean for observed and expected heterozygosity, whereas Belgian, Polish and Estonian populations had higher diversity than the mean (Supplementary data Table S1). However, Italian populations had high allele richness. Mean habitat suitability was not correlated with an increase in genetic diversity. However, populations located in areas that experienced a greater change in habitat suitability from the LGM to near-present demonstrated higher expected heterozygosity ($R^2 = 0.45$, $F_{1,21} = 17.17$, $P = 0.0005$) and observed heterozygosity ($R^2 = 0.52$, $F_{1,21} = 22.64$, $P = 0.0001$).

The AMOVA indicated that most of the genetic variation observed (95 %) occurs within populations, and the remaining 5 % occurs between populations. The mean F_{ST} (the mean of the pairwise distance values between populations) was 0.089 ± 0.003. A Mantel test indicated that genetic distance (F_{ST}) was not a function of geographical distance between populations ($r_M = 0.014$, $P = 0.45$; Supplementary data Appendix S2).

Thirteen genetic clusters were detected in the STRUCTURE analyses (Supplementary data Fig. S1), with substantial admixture among most sampled populations (Fig. 2). The Greek populations showed clear membership of a genetic cluster. The Slovak, Slovenian and Danish populations, and one French and one Hungarian population, showed strong genetic structure.
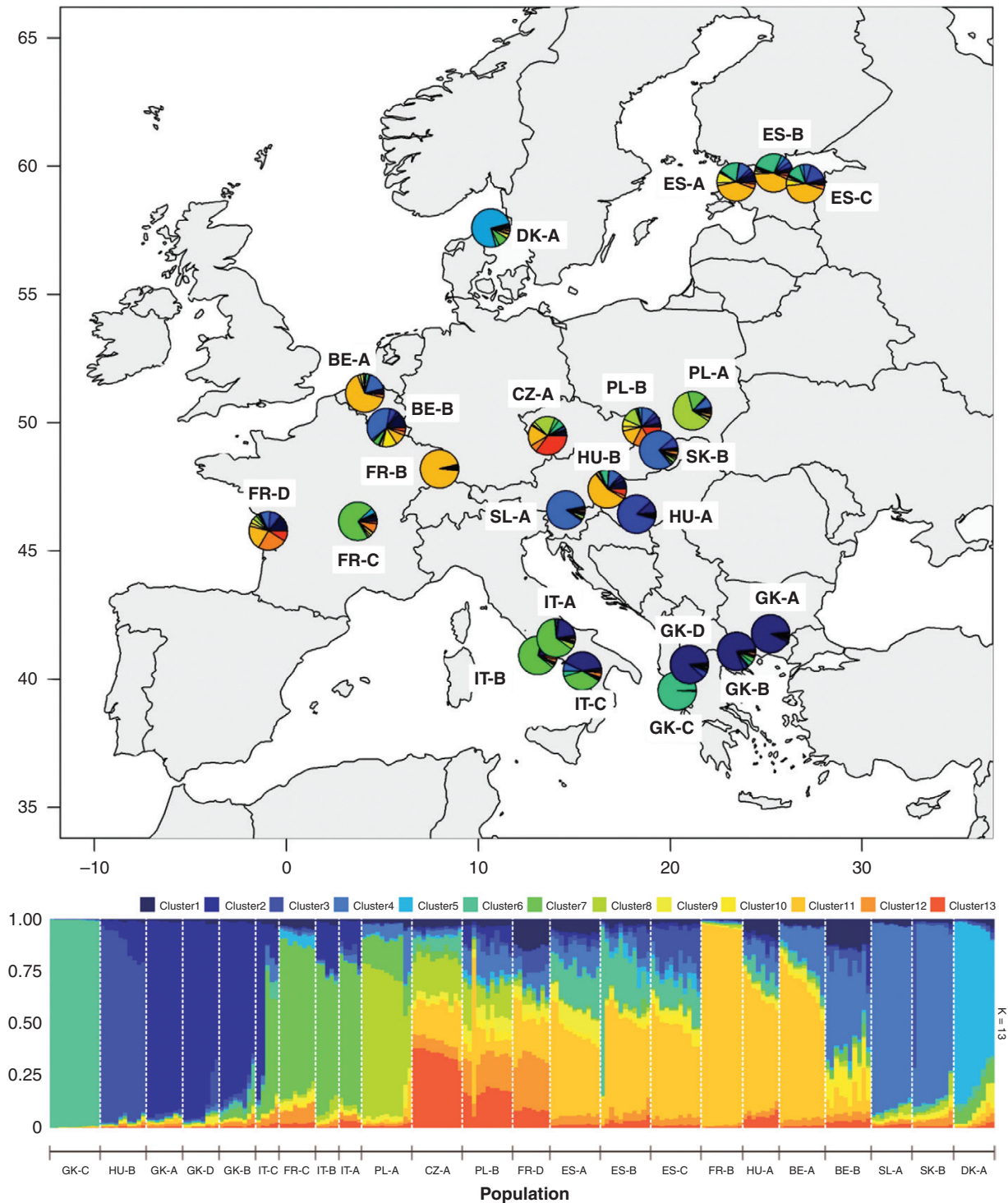
FIG. 2. Genetic clusters detected among the samples from 23 sampling locations (populations) in Europe of *Epipactis helleborine*. The membership coefficient of each sampled population is represented as a proportion of each colour-coded genetic cluster.

Outlier detection and analyses

There were 117 outlier SNP loci detected by Bayescan and 101 by PCAdapt, with 76 outlier loci and 181 neutral loci shared by the two methods, of the 893 total loci (Supplementary data Fig. S2). The mean observed heterozygosity was significantly higher for outliers (0.59 ± 0.007 s.e.) than for neutral SNPs (0.52 ± 0.007; t_{21} = 10.62, P < 0.0001), but there was no significant difference between SNP types detected for expected heterozygosity or allele richness. All genetic diversity indices had significant positive relationships with an increase in habitat suitability for the outlier dataset (Table 1), indicating adaptive

TABLE 1. *Results of regressions testing the effects of change in habitat suitability from the Last Glacial Maximum to the near-present on genetic diversity, for outlier and neutral markers separately. The genetic diversity measures tested were the observed and expected heterozygosity and allele richness (H_o, H_e and AR, respectively) of a rarefied sample*

Diversity measure	Outlier SNPs			Neutral SNPs		
	Adjusted R^2	*F*-statistic	*P*-value	Adjusted R^2	*F*-statistic	*P*-value
H_o	0.141	4.459	**0.048**	−0.045	0.092	0.765
H_e	0.149	4.666	**0.043**	−0.018	0.636	0.435
AR	0.152	4.77	**0.041**	−0.007	0.857	0.366

Abbreviation: SNP, single-nucleotide polymorphism.
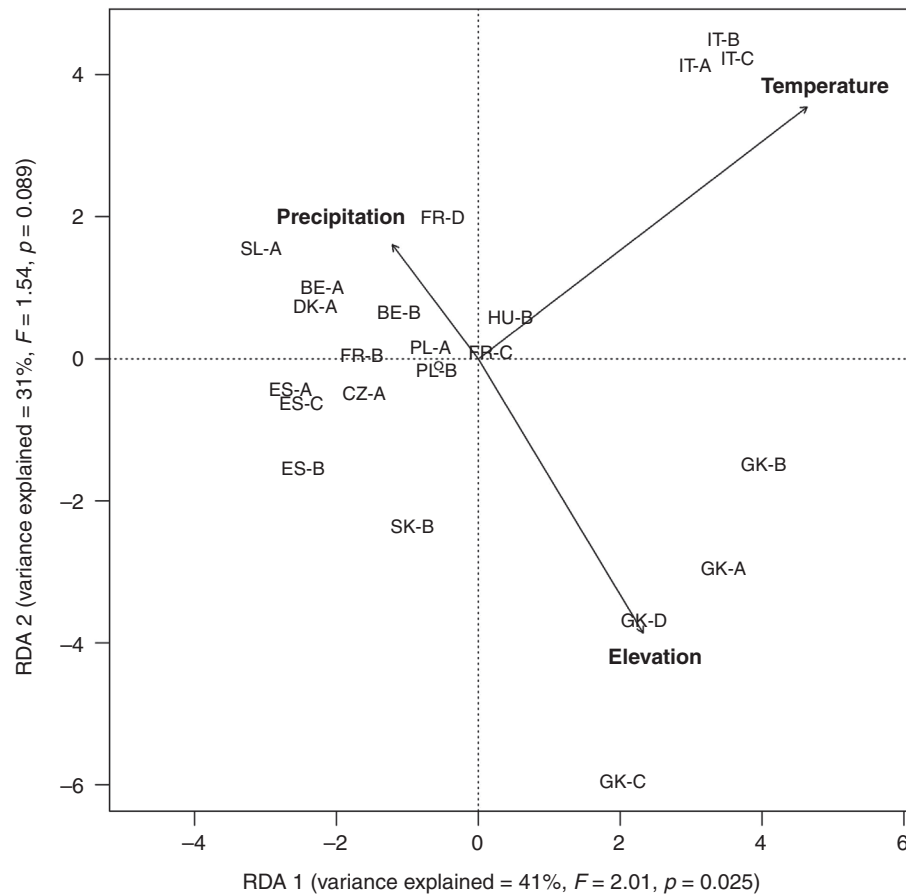
Significant *P*-values are in bold.

FIG. 3. Redundancy analysis (RDA) of *Epipactis helleborine* single-nucleotide polymorphisms, showing the effects of annual precipitation, mean annual temperature and elevation on genetic structure.

potential in these measures with regard to change in suitable habitat. For the neutral SNP dataset, no diversity indices were significantly associated with the change in habitat suitability.

The first axis of the RDA was significant ($F = 2.01$, $P = 0.025$), and from the loadings of this axis we detected 91 outlier SNPs associated with temperature, 52 with elevation and 28 with precipitation. The RDA indicated that Greek and Italian populations were distinct from the rest of the populations (Fig. 3). Greek populations were strongly associated with higher elevations and Italian populations with higher temperatures, while the remaining populations were moderately associated with lower rainfall.

Geogenetic analysis

The geogenetic analyses indicated sources of genetic enrichment (i.e. the predicted historical spatial location of genetic source populations that contributed to the current gene pool of the sampled populations) mainly in Italy and central Europe, with a few sources scattered in northern Europe and
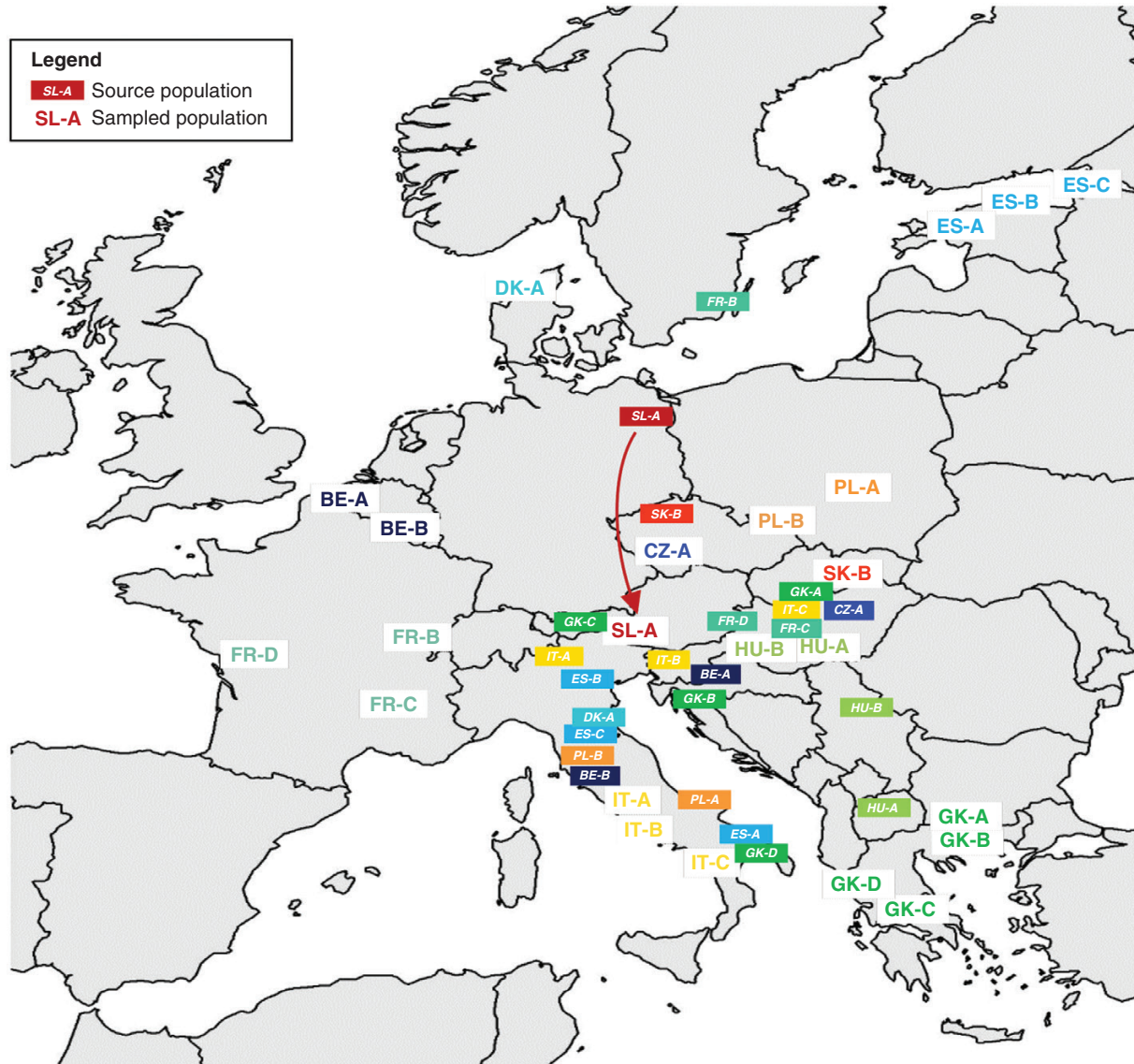
FIG. 4. Locations of main sources of genetic enrichment (coloured blocks with white text) predicted by spacemix for each sampled population (coloured text) of *Epipactis helleborine* in Europe. An example is shown by an arrow of the potential movement of genetic material from a hypothetical historical source population to the present sampled population, SL-A.

one in Albania (Fig. 4). The genetic sources in Italy dispersed throughout Europe, mainly north-east towards central and northern Europe, including Denmark and Estonia, while the sources in central Europe contributed to western populations in France and Italy. The central European sources also contributed to populations within central Europe, such as the Slovak and Czech populations.

DISCUSSION

Refugia and genetic isolation

Considering that dispersal of *E. helleborine* is most probably not distance limited, genetic signals for restricted gene flow in these populations are likely to be attributable to geographical barriers and historical isolation. The wide northern European distribution of suitable habitat of *E. helleborine* during the LI shrank and shifted to areas around the Mediterranean and Aegean, the western coasts of France and the Iberian Peninsula, and small areas of the Alps and Carpathians during the LGM (Fig. 1). Previous research has indicated that broad-leafed forests were able to persist through the LGM in glacial refugia (Bhagwat and Willis, 2008), notably in mountainous areas on the Greek and Italian peninsulas and the central European Carpathians (Tzedakis, 2004; Giesecke, 2013). These habitats might have provided buffered microclimates to *E. helleborine*, which, being herbaceous, small-seeded and dependent on mycorrhizal fungi, would not otherwise have survived glaciation (Keppel *et al.*, 2012).

The steep environmental gradients present in mountainous areas limit the exchange of genetic material between

populations through physical barriers to pollination and seed dispersal (Riebesell, 1982; Steinbauer *et al.*, 2016; Wu *et al.*, 2020) and a mismatch in flowering times owing to temperature differences at different elevations (Gustafsson and Lönn, 2003; Gaudinier and Blackman, 2020). Populations of *E. helleborine* in Greece demonstrated low heterozygosity and strong genetic structure (Fig. 2), which might be indicative of geographical isolation (Habel and Schmitt, 2012; Carvalho *et al.*, 2019) and elevation-associated adaptive alleles (Fig. 3). Populations that survived in mountain glacial refugia in Greece probably tracked the suitable climate up to the high elevations where they occur at today, following the migration of boreal forests (Hewitt, 1999), as surrounding conditions became too warm during the climate amelioration (Fig. 1). This geographical isolation created a genetic bottleneck and resulted in relatively lower genetic diversity and strong genetic structure (Nei *et al.*, 1975). Although heterozygosity was low relative to the other populations, it was not as low as might be expected in geographically isolated populations of other species, possibly owing to the role of vegetative reproduction in reducing the loss of genetic diversity through genetic drift (Gargiulo *et al.*, 2019).

The importance of Mediterranean refugia has been widely recognized in the literature (Médail and Diadema, 2009), and the presence of central European microrefugia is becoming more apparent (see Schmitt & Varga, 2012). Although the spatial resolution we used did not allow us to detect microrefugia, some evidence of suitable habitat during the LGM can be seen in the vicinity of Slovakia and Hungary (Fig. 1), and glacial refugia have been identified here for species with which *E. helleborine* commonly co-occurs, such as beech (Delforge, 2006; Giesecke, 2013). Additionally, these areas contain populations that are genetically distinct (Fig. 2). Populations in the French Alps and Denmark also demonstrated signs of geographical isolation that might be a result of geographical barriers, but their genetic distinctness was not linked to adaptation to the environmental variables we tested. Considering that these areas were glaciated in the LGM, it is likely that these populations originated from more southern refugia and became isolated after following the migration of cold-adapted forest trees north and towards higher elevations (Hewitt, 1999; Parisod, 2022).

Postglacial radiation and gene flow

As the distribution of suitable habitat expanded northwards following the LGM, Italy and central Europe acted as centres of genetic enrichment for other populations in Europe during postglacial recolonization of the continent (Fig. 4). The dispersal of genetic material from source populations of the species was possible owing to the high dispersal capabilities of orchid 'dust seeds', and despite most seeds landing near the mother plant, the small proportion of seeds that disperses from the parental population is likely to be enough to allow for gene flow across populations and the colonization of new habitats (Arditti and Ghani, 2000; Jersáková and Malinová, 2007; Kotilínek *et al.*, 2020). This is reflected in the relatively high levels of genetic diversity and low structuring in non-refugial populations, which is in agreement with previous studies on the genetic structure of other European *E. helleborine* populations (Tranchida-Lombardo *et al.*, 2011; Jacquemyn *et al.*, 2018,

2020) and species capable of long-distance dispersal (Phillips *et al.*, 2012; Helsen *et al.*, 2015). Populations that were situated in areas that experienced a great increase in suitability of habitat between the LGM and near-present (i.e. glaciated to habitable) had the highest measures of heterozygosity (Table 1). This might be attributable to increased gene flow that came with improved survival and establishment and therefore greater opportunities for reproduction and population growth, which increased the gene pool and chance of admixture.

Implications for adaptive potential

The observed patterns of genomic divergence support previous claims that the genus *Epipactis* is currently undergoing an evolutionary radiation driven by a wide spectrum of genotypic, phenotypic and environmental factors and that *E. helleborine* might serve as a source of local variants from which new ecotypes have arisen (Hollingsworth *et al.*, 2006; Tranchida-Lombardo *et al.*, 2011; Jacquemyn *et al.*, 2020). However, the observed levels of genomic divergence are most probably not strong enough to have led to complete reproductive isolation between populations, and local variants can at best be described as varieties or ecotypes (Jacquemyn *et al.*, 2018; Sramkó *et al.*, 2019).

When adaptive markers are analysed separately from neutral markers, one can disentangle the effects of natural selection and neutral evolution on genetic diversity and identify the potential of a species to adapt to environmental change (Kirk and Freeland, 2011). The higher genetic diversity in adaptive markers that was observed in populations located in areas that underwent a greater change in habitat suitability supports the potential of this species for adapting to and surviving changes in climate. This is especially apparent in the high allelic richness of Italian populations, indicating adaptive potential (Vilas *et al.*, 2015), in addition to the evidence of adaptive alleles associated with higher temperatures. In contrast to the climate tracking of Greek populations, the genetic structure of Italian populations might have arisen more recently, through adaptation to increasing postglacial temperatures (Fig. 3), resulting in ecological isolation reflected by similarly low heterozygosity and moderate genetic structure. However, the Greek populations will become more isolated if they continue to track suitable climate up to higher elevations, and their lower neutral genetic diversity reduces the chance of adaptive alleles arising (Luikart *et al.*, 2003; Kirk and Freeland, 2011). Unless outlier loci are directly linked to a phenotypic trait that demonstrates adaptation, linking outlier loci to environmental gradients should be assessed with caution, because genetic drift can cause populations to diverge at neutral loci through reproductive isolation (Stinchcombe and Hoekstra, 2008; Kirk and Freeland, 2011). Considering that there was no significant relationship between genetic diversity in neutral SNPs and habitat suitability, we believe that adaptation to local environments contributed more than genetic drift to population structure.

Conclusions

The observed spatial genetic structure of *E. helleborine* results from a complex interplay between changes in habitat

availability since the LGM, gene flow, and adaptation to local environmental conditions. Our results support the occurrence of central European glacial refugia, in addition to Mediterranean refugia, and their importance in contributing to postglacial recolonization and genetic enrichment of forest species. The assumption that refugial populations have higher genetic diversity than offshoot populations might not be consistent for species with highly dispersible seeds and complex environmental associations, such as *E. helleborine*, where colonizing populations are subject to founder effects but can also act as sources of genetic enrichment. Adaptive alleles associated with temperature and elevation contributed to the divergence of Italian and Greek populations, with different implications for the adaptive potential of these populations regarding climate change. This study demonstrates how the evolutionary histories of herbaceous, small-seeded species that are not readily fossilized can be characterized using a combination of ecological niche modelling and genomics and might serve as a basis for future studies aiming to disentangle the phylogeny of the *E. helleborine* species complex.

SUPPLEMENTARY DATA

Supplementary data are available online at https://academic.oup.com/aob and consist of the following.

Appendix S1: ecological niche distributions of *Epipactis helleborine* in Europe, projected for three past climatic periods: (a) Last Interglacial period; (b) Last Glacial Maximum; (c) Mid-Holocene; and (d) the near-present climate (1970–2000). Appendix S2: results of a Mantel test, showing the relationship between pairwise $\log_{10}$-transformed geographical distance and F_{ST} values among *Epipactis helleborine* populations in Europe. Text S1: DNA extraction, sequencing and filtering. Table S1: mean Gibbs values of near-present climate conditions and mean genetic diversity with standard errors of 23 *Epipactis helleborine* populations sampled in Europe. Fig. S1: delta *K* graph indicating the most likely number of genetic clusters in the sample of *Epipactis helleborine*, detected by the Evanno method. Fig. S2: results of Bayescan and PCAdapt for detecting genetic outliers in a single-nucleotide polymorphism dataset of *Epipactis helleborine* samples collected in Europe.

ACKNOWLEDGEMENTS

FUNDING

LITERATURE CITED

Amatulli G, **Domisch S**, **Tuanmu M-N**, *et al*. **2018**. A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific Data* **5**: 180040.

Arditti J, **Ghani AKA. 2000**. Tansley Review No. 110. Numerical and physical properties of orchid seeds and their biological implications. *New Phytologist* **145**: 367–421.

Ashcroft MB. 2010. Identifying refugia from climate change. *Journal of Biogeography* **37**: 1407–1413.

Bateman RM. 2020. Implications of next-generation sequencing for the systematics and evolution of the terrestrial orchid genus *Epipactis*, with particular reference to the British Isles. *Kew Bulletin* **75**: 4.

Bhagwat SA, **Willis KJ. 2008**. Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *Journal of Biogeography* **35**: 464–482.

Bidartondo MI, **Read DJ. 2008**. Fungal specificity bottlenecks during orchid germination and development. *Molecular Ecology* **17**: 3707–3716.

Bradburd GS, **Ralph PL**, **Coop GM. 2016**. A spatial framework for understanding population structure and admixture. *PLoS Genetics* **12**: e1005703.

Carvalho YGS, **Vitorino LC**, **de Souza UJB**, **Bessa LA. 2019**. Recent trends in research on the genetic diversity of plants: implications for conservation. *Diversity* **11**: 62.

Comes HP, **Kadereit JW. 1998**. The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science* **3**: 432–438.

De Kort H, **Vandepitte K**, **Bruun HH**, **Closset-Kopp D**, **Honnay O**, **Mergeay J. 2014**. Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*. *Molecular Ecology* **23**: 4709–4721.

de Lafontaine G, **Napier JD**, **Petit RJ**, **Hu FS. 2018**. Invoking adaptation to decipher the genetic legacy of past climate change. *Ecology* **99**: 1530–1546.

Delforge P. 2006. *Orchids of Europe, North Africa and the Middle East*. Portland, OR: Timber Press.

Djordjević V, **Tsiftsis S. 2022**. The role of ecological factors in distribution and abundance of terrestrial orchids. In: **Mérillon J-M**, **Kodja H**, eds. Reference Series in Phytochemistry. *Orchids Phytochemistry, Biology and Horticulture*. Cham: Springer International, 3–72.

Elith J, **Phillips SJ**, **Hastie T**, **Dudík M**, **Chee YE**, **Yates CJ. 2011**. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**: 43–57.

Ellstrand NC. 2014. Is gene flow the most important evolutionary force in plants? *American Journal of Botany* **101**: 737–753.

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.

Evans A, **Jacquemyn H. 2020**. Impact of mating system on range size and niche breadth in *Epipactis* (Orchidaceae). *Annals of Botany* **126**: 1203–1214.

Fick SE, **Hijmans RJ. 2017**. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**: 4302–4315.

Fitzpatrick SW, **Reid BN. 2019**. Does gene flow aggravate or alleviate maladaptation to environmental stress in small populations? *Evolutionary Applications* **12**: 1402–1416.

Foll M, **Gaggiotti O. 2008**. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**: 977–993.

Forester B. 2019. Detecting multilocus adaptation using redundancy analysis (RDA). *Population genetics in R*. https://popgen.nescent.org/2018-03-27_RDA_GEA.html (accessed 19 December 2022).

Francis RM. 2017. pophelper: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* **17**: 27–32.

François O. 2016. Running structure-like population genetic analyses with R. *R tutorials in population genetics, U. Grenoble-Alpes*: 1–9. http://membres-timc.imag.fr/Olivier.Francois/tutoRstructure.pdf

García-Navas V, **Bonnet T**, **Waldvogel D**, **Wandeler P**, **Camenisch G**, **Postma E. 2015**. Gene flow counteracts the effect of drift in a Swiss

population of snow voles fluctuating in size. *Biological Conservation* **191**: 168–177.

Gargiulo R, **Pironon S**, **Zheleznaya E**, *et al*. **2019**. Phylogeography and post-glacial dynamics in the clonal-sexual orchid *Cypripedium calceolus* L. *Journal of Biogeography* **46**: 526–538.

Gaudinier A, **Blackman BK. 2020**. Evolutionary processes from the perspective of flowering time diversity. *New Phytologist* **225**: 1883–1898.

Giesecke T. 2013. Changing plant distributions and abundances. In: **Mock C, Elias S**, eds. *Encyclopedia of Quaternary Science*, 2nd edn. Amsterdam: Elsevier, 854–860.

Givnish TJ. 2010. Ecology of plant speciation. *Taxon* **59**: 1326–1366.

Givnish TJ, **Spalink D**, **Ames M**, *et al*. **2016**. Orchid historical biogeography, diversification, Antarctica and the paradox of orchid dispersal. *Journal of Biogeography* **43**: 1905–1916.

Goudet J. 2005. ʜɪᴇʀғsᴛᴀᴛ, a package for ʀ to compute and test hierarchical *F*-statistics. *Molecular Ecology Notes* **5**: 184–186.

Gustafsson S, **Lönn M. 2003**. Genetic differentiation and habitat preference of flowering-time variants within *Gymnadenia conopsea*. *Heredity* **91**: 284–292.

Habel JC, **Schmitt T. 2012**. The burden of genetic diversity. *Biological Conservation* **147**: 270–274.

Hämälä T, **Savolainen O. 2019**. Genomic patterns of local adaptation under gene flow in *Arabidopsis lyrata*. *Molecular Biology and Evolution* **36**: 2557–2571.

Helsen K, **Jacquemyn H**, **Honnay O. 2015**. Hidden founder effects: small-scale spatial genetic structure in recently established populations of the grassland specialist plant *Anthyllis vulneraria*. *Molecular Ecology* **24**: 2715–2728.

Hemrová L, **Kotilínek M**, **Konečná M**, *et al*. **2019**. Identification of drivers of landscape distribution of forest orchids using germination experiment and species distribution models. *Oecologia* **190**: 411–423.

Hewitt GM. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**: 87–112.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal B: Biological Sciences* **359**: 183–95; discussion 195.

Hijmans RJ, **Cameron SE**, **Parra JL**, **Jones PG**, **Jarvis A. 2005**. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

Hijmans RJ, **Van Etten J**, **Cheng J**, *et al*. **2015**. Package 'raster'. *R package*. https://cran.r-project.org/web/packages/raster/index.html

Hollingsworth P, **Squirrell J**, **Hollingsworth M**, **Richards A**, **Bateman R. 2006**. Taxonomic complexity, conservation and recurrent origins of self-pollination in *Epipactis* (Orchidaceae). In: **Bailey JP** and **Ellis RG**, eds. *Current taxonomic research on the British and European flora*. London: Botanical Society of the British Isles, 27–44.

Hutchison DW, **Templeton AR. 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.

Ilves A, **Metsare M**, **Tali K**, **Kull T. 2015**. The impact of recent colonization on the genetic diversity and fine-scale genetic structure in *Orchis militaris* (L.). *Plant Systematics and Evolution* **301**: 1875–1886.

Jacquemyn H, **De Kort H**, **Vanden Broeck A**, **Brys R. 2018**. Immigrant and extrinsic hybrid seed inviability contribute to reproductive isolation between forest and dune ecotypes of *Epipactis helleborine* (Orchidaceae). *Oikos* **127**: 73–84.

Jacquemyn H, **De Kort H**, **Vanden Broeck A**, **Brys R. 2020**. Low genetic divergence and variation in coastal dune populations of the widespread terrestrial orchid *Epipactis helleborine*. *Botanical Journal of the Linnean Society* **193**: 419–430.

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

Jersáková J, **Malinová T. 2007**. Spatial aspects of seed dispersal and seedling recruitment in orchids. *New Phytologist* **176**: 237–241.

Keppel G, **Niel KPV**, **Wardell-Johnson GW**, *et al*. **2012**. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography* **21**: 393–404.

Kirk H, **Freeland JR. 2011**. Applications and implications of neutral versus non-neutral markers in molecular ecology. *International Journal of Molecular Sciences* **12**: 3966–3988.

Kolář F, **Fuxová G**, **Záveská E**, *et al*. **2016**. Northern glacial refugia and altitudinal niche divergence shape genome-wide differentiation in the emerging plant model *Arabidopsis arenosa*. *Molecular Ecology* **25**: 3929–3949.

Kotilínek M, **Těšitelová T**, **Košnar J**, *et al*. **2020**. Seed dispersal and realized gene flow of two forest orchids in a fragmented landscape. *Plant Biology* **22**: 522–532.

Leutner B, **Horning N**, **Schwalb-Willmann J**, **Hijmans R. 2017**. RStoolbox: tools for remote sensing data analysis. *R package version 0.2.6*. https://cran.r-project.org/web/packages/RStoolbox/index.html

Loveless MD, **Hamrick JL. 1984**. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* **15**: 6533–6595.

Luikart G, **England PR**, **Tallmon D**, **Jordan S**, **Taberlet P. 2003**. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* **4**: 981–994.

Luu K, **Bazin E**, **Blum MGB. 2017**. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* **17**: 67–77.

McCormick MK, **Jacquemyn H. 2014**. What constrains the distribution of orchid populations? *New Phytologist* **202**: 392–400.

McCormick MK, **Whigham DF**, **Canchani-Viruet A. 2018**. Mycorrhizal fungi affect orchid distribution and population dynamics. *New Phytologist* **219**: 1207–1215.

Médail F, **Diadema K. 2009**. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography* **36**: 1333–1345.

Merow C, **Smith MJ**, **Silander JA Jr. 2013**. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* **36**: 1058–1069.

Muir AP, **Biek R**, **Thomas R**, **Mable BK. 2014**. Local adaptation with high gene flow: temperature parameters drive adaptation to altitude in the common frog (*Rana temporaria*). *Molecular Ecology* **23**: 561–574.

Nei M, **Maruyama T**, **Chakraborty R. 1975**. The bottleneck effect and genetic variability in populations. *Evolution* **29**: 1–10.

Nosil P, **Funk DJ**, **Ortiz-Barrientos D. 2009**. Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* **18**: 375–402.

Oksanen J, **Blanchet F**, **Friendly M**, *et al*. **2019**. v*egan: community ecology package. R package version 2.5-6*. 2019. https://cran.r-project.org/web/packages/vegan/index.html

Orsini L, **Andrew R**, **Eizaguirre C. 2013**. Evolutionary ecological genomics. *Molecular Ecology* **22**: 527–531.

Otto-Bliesner BL, **Marshall SJ**, **Overpeck JT**, **Miller GH**, **Hu A. 2006**. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science* **311**: 1751–1753.

Parisod C. 2022. Plant speciation in the face of recurrent climate changes in the Alps. *Alpine Botany* **132**: 21–28.

Peakall R. 2007. Speciation in the Orchidaceae: confronting the challenges. *Molecular Ecology* **16**: 2834–2837.

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

Phillips RD, **Dixon KW**, **Peakall R. 2012**. Low population genetic differentiation in the Orchidaceae: implications for the diversification of the family. *Molecular Ecology* **21**: 5208–5220.

Phillips SJ, **Dudík M**, **Schapire RE. 2018**. *Maxent software for modeling species niches and distributions*. http://biodiversityinformatics.amnh.org/open_source/maxent/

Pritchard JK, **Stephens M**, **Donnelly P. 2000**. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

QGIS Development Team. **2019**. *QGIS geographic information system*. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

R Core Team. **2021**. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/

Reisch C, **Rosbakh S. 2021**. Patterns of genetic variation in European plant species depend on altitude. *Diversity and Distributions* **27**: 157–163.

Riebesell JF. 1982. Arctic-Alpine plants on mountaintops: agreement with island biogeography theory. *The American Naturalist* **119**: 657–674.

Schmitt T, **Varga Z. 2012**. Extra-Mediterranean refugia: the rule and not the exception? *Frontiers in Zoology* **9**: 22.

Schönswetter P, **Stehlik I**, **Holderegger R**, **Tribsch A. 2005**. Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology* **14**: 3547–3555.

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science* **236**: 787–792.

Sramkó G, **Paun O**, **Brandrud MK**, **Laczkó L**, **Molnár A**, **Bateman RM. 2019**. Iterative allogamy–autogamy transitions drive actual and incipient speciation during the ongoing evolutionary radiation within the orchid genus *Epipactis* (Orchidaceae). *Annals of Botany* **124**: 481–497.

Steinbauer MJ, **Field R**, **Grytnes J**, *et al*. **2016**. Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecology and Biogeography* **25**: 1097–1107.

Stewart JR, **Lister AM. 2001**. Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution* **16**: 608–613.

Stinchcombe JR, **Hoekstra HE. 2008**. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* **100**: 158–170.

Tranchida-Lombardo V, **Cafasso D**, **Cristaudo A**, **Cozzolino S. 2011**. Phylogeographic patterns, genetic affinities and morphological differentiation between *Epipactis helleborine* and related lineages in a Mediterranean glacial refugium. *Annals of Botany* **107**: 427–436.

Tremblay RL, **Ackerman JD**, **Zimmerman JK**, **Calvo RN. 2005**. Variation in sexual reproduction in orchids and its evolutionary consequences: a spasmodic journey to diversification. *Biological Journal of the Linnean Society* **84**: 1–54.

Tribsch A. 2004. Areas of endemism of vascular plants in the Eastern Alps in relation to Pleistocene glaciation. *Journal of Biogeography* **31**: 747–760.

Tzedakis PC. 2004. The Balkans as prime glacial refugial territory of European temperate trees. In: **Griffiths HI**, **Kryštufek B**, **Reed JM**, eds. *Balkan biodiversity: pattern and process in the European hotspot*. Dordrecht: Springer Netherlands, 49–68.

Van Liedekerke M, **Jones A**, **Panagos P. 2006**. ESDBv2 raster library—a set of rasters derived from the European Soil Database Distribution v2.0. *European Commission and the European Soil Bureau Network, CDROM, EUR* **19945**. https://esdac.jrc.ec.europa.eu/content/european-soil-database-v2-raster-library-1kmx1km.

Vilas A, **Pérez-Figueroa A**, **Quesada H**, **Caballero A. 2015**. Allelic diversity for neutral markers retains a higher adaptive potential for quantitative traits than expected heterozygosity. *Molecular Ecology* **24**: 4419–4432.

Wu Z, **Li X**, **Xie D**, *et al*. **2020**. Spatial genetic structuring in a widespread wetland plant on a plateau: effects of elevation-driven geographic isolation and environmental heterogeneity. *Freshwater Biology* **65**: 1596–1607.