



Phylogeography and genetic structure of the orchid *Himantoglossum hircinum* (L.) Spreng. across its European central–marginal gradient

Marion Pfeifer^{1*}, Bertrand Schatz², F. Xavier Picó³, Nicodemo G. Passalacqua⁴, Michael F. Fay⁵, Pete D. Carey⁶ and Florian Jeltsch¹

¹Institute of Vegetation Ecology and Nature Conservation, University of Potsdam, Maulbeerallee 2, Potsdam, Germany, ²Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 Montpellier, France, ³Estación Biológica de Doñana, CSIC, Avda. María Luisa s/n, Sevilla, Spain, ⁴Museo di Storia Naturale della Calabria ed Orto Botanico, Università della Calabria, 87030 Arcavacata di Rende, Cosenza, Italy, ⁵Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey, UK and ⁶Centre for Ecology and Hydrology, Monks Wood, Abbots Ripton, Huntingdon, Cambridgeshire, UK

ABSTRACT

Aim This study aims to link demographic traits and post-glacial recolonization processes with genetic traits in *Himantoglossum hircinum* (L.) Spreng (Orchidaceae), and to test the implications of the central–marginal concept (CMC) in Europe.

Location Twenty sites covering the entire European distribution range of this species.

Methods We employed amplified fragment length polymorphism (AFLP) markers and performed a plastid microsatellite survey to assess genetic variation in 20 populations of *H. hircinum* located along central–marginal gradients. We measured demographic traits to assess population fitness along geographical gradients and to test for correlations between demographic traits and genetic diversity. We used genetic diversity indices and analyses of molecular variance (AMOVA) to test hypotheses of reduced genetic diversity and increased genetic differentiation and isolation from central to peripheral sites. We used Bayesian simulations to analyse genetic relationships among populations.

Results Genetic diversity decreased significantly with increasing latitudinal and longitudinal distance from the distribution centre when excluding outlying populations. The AMOVA revealed significant genetic differentiation among populations ($F_{ST} = 0.146$) and an increase in genetic differentiation from the centre of the geographical range to the margins (except for the Atlantic group). Population fitness, expressed as the ratio N_R/N , decreased significantly with increasing latitudinal distance from the distribution centre. Flower production was lower in most eastern peripheral sites. The geographical distribution of microsatellite haplotypes suggests post-glacial range expansion along three major migratory pathways, as also supported by individual membership fractions in six ancestral genetic clusters (C1–C6). No correlations between genetic diversity (e.g. diversity indices, haplotype frequency) and population demographic traits were detected.

Main conclusions Reduced genetic diversity and haplotype frequency in *H. hircinum* at marginal sites reflect historical range expansions. Spatial variation in demographic traits could not explain genetic diversity patterns. For those sites that did not fit into the CMC, the genetic pattern is probably masked by other factors directly affecting either demography or population genetic structure. These include post-glacial recolonization patterns and changes in habitat suitability due to climate change at the northern periphery. Our findings emphasize the importance of distinguishing

*Correspondence: Marion Pfeifer, Institute of Vegetation Ecology and Nature Conservation, University of Potsdam, Maulbeerallee 2, 14469 Potsdam, Germany.
E-mail: marion.pfeifer@googlemail.com

historical effects from those caused by geographical variation in population demography of species when studying evolutionary and ecological processes at the range margins under global change.

Keywords

AFLP, demography, Europe, geographical genetic structure, *Himantoglossum hircinum*, orchids, phylogeography.

INTRODUCTION

Previous studies have tested the demographic assumptions of the 'abundant centre' model (or 'central–marginal concept', CMC) and its implications for geographical genetic structuring of species within Europe (reviewed by Sagarin & Gaines, 2002; Eckert *et al.*, 2008). The model suggests that demographic and evolutionary processes at the margins of the range of a species determine not only its present-day distribution but also its response to changing environmental conditions (Kirkpatrick & Barton, 1997; Case & Taper, 2000). However, empirical evidence for model assumptions remains inconclusive, demonstrating that the CMC is far from fully understood (Eckert *et al.*, 2008).

Declining environmental favourability towards the periphery of the geographical range (Hengeveld & Haeck, 1982) is predicted to result in a reduction in population size and density (Brown, 1984) and a decrease in both plant growth and reproduction (Parsons, 1991) towards range boundaries of species. However, whereas higher survival, fecundity and/or growth have been observed in central than in peripheral populations in some species, the opposite has been found in others (Sagarin & Gaines, 2002). Decline in population densities towards the high latitudinal distribution limits has been linked to reduced fecundity in peripheral populations of *Decodon verticillatus* (Dorken & Eckert, 2001), *Cirsium acaule* and *Cirsium heterophyllum* (Jump & Woodward, 2003) and greater interannual variation in life-state transition rates of passerine birds (Mehlman, 1997). However, reductions in fitness components do not necessarily have an impact on population growth and persistence, and thus fail to explain the position of range boundaries (Nantel & Gagnon, 1999; Kluth & Bruelheide, 2005). Also, geographical ranges examined in many species did not contract inwards when a species became endangered, and remnant populations of many species occurred exclusively in the periphery of the historical range (Channell & Lomolino, 2000).

The genetic implications of the CMC are numerous. Genetic diversity is hypothesized to decrease towards the range edge, because peripheral populations are expected to have lower effective population sizes and may consequently suffer from increased genetic drift and inbreeding (Giles & Goudet, 1997). Declining availability of suitable habitat patches and/or a higher probability of stochastic extinctions due to stronger demographic fluctuations towards range margins (Maurer &

Taper, 2002) are expected to increase isolation between populations. This, in combination with smaller population sizes, may lead to reduced gene flow and increased genetic differentiation (F_{ST}) between populations (Ellstrand & Hoffman, 1990). This is relevant for understanding the distribution patterns of species, because a stochastic reduction in the genetic diversity of peripheral populations may limit their evolutionary potential, inhibiting the capacity of the species to adapt to conditions beyond the range limits (Blows & Hoffmann, 2005).

However, empirical studies show that patterns of population genetic structure across large spatial scales are variable and species-specific (Eckert *et al.*, 2008). Peripheral populations may not show decreased genetic diversity (e.g. *Cirsium heterophyllum*, Jump & Woodward, 2003) or may contain genotypes evolved under variable, extreme and suboptimal conditions. Eckert *et al.* (2008) criticized many of the studies they reviewed for not considering historical influences in a phylogeographical framework and for not assessing any of the possible causes of reduced peripheral genetic diversity or greater differentiation, including estimates of population size, degree of spatial isolation among populations and demographic turnover.

This study attempts to show how the distribution of species is shaped by historical processes, demographic traits and genetic processes interacting over geographical scales, and consequently how these processes determine the capacity of the species to respond to environmental changes. We studied fitness and the geographical genetic structure of populations of the lizard orchid, *Himantoglossum hircinum* (L.) Spreng. (Orchidaceae), along central–marginal gradients covering its entire distribution area in Europe, using amplified fragment length polymorphism (AFLP) markers and demographic monitoring. Plastid microsatellite markers were used to infer migratory pathways of the species and to assess the potential influence of historical processes on geographical genetic structure. *Himantoglossum hircinum* is a climate-sensitive plant that has recently experienced increases in abundance and population numbers at its north-eastern range margin, indicating a range shift (cf. Good, 1936; Carey, 1999). However, it is unclear whether these changes will compensate for the decline in abundance observed along the southern range margin of the species.

The following three questions were addressed in this study. (1) Can geographical trends be detected in population genetic

structure? (2) If geographical genetic structure is present, can the observed pattern be related to demographic traits (e.g. population fitness). (3) How important are historical processes for the current genetic structure of the species?

MATERIALS AND METHODS

Study organism

Himantoglossum hircinum is a long-lived terrestrial orchid that perennates via tubers. Leaves emerge from below-ground tubers in late autumn and grow over the winter months; plants start to develop inflorescences in late April (Carey & Farrell, 2002). Population growth rate, transition probabilities between life stages and flowering probability of *H. hircinum* populations at the north-eastern margin of its range have been shown to be strongly governed by variability in weather (Pfeifer *et al.*, 2006a,b). Flowering takes place between May and June, although a shift towards an earlier onset of flowering in some populations has been observed (M. Pfeifer, unpublished data). Plants are pollinated by generalist insects, precluding the possibility of range limitation due to an absence of pollinators (Bournérias & Prat, 2005). Seeds often remain lodged in the capsules, which are dropped beside the parent plant after maturation.

Study sites, measurements and sampling

Twenty sites distributed throughout the entire range of the species in Europe were sampled in 2007 (Table 1). The study covered three central sites, two sites in the northern range periphery, a number of sites along a gradient towards the eastern range margin and three sites in the southern range periphery of its sub-Mediterranean distribution (Fig. 1). Populations located in the mountains in northern Spain were grouped as Atlantic sites because of the likely influence of the Atlantic on their environment (e.g. weather conditions).

Population size (N) was estimated at each sample site in March. Leaf number was used as a measure of plant size, assuming that leaf number and summed leaf area are correlated (shown for 50 plants in 2004 at TH1: $R^2 = 0.87$, $P < 0.001$). In June each year, populations were revisited and the number of reproductive plants (N_R) was estimated. Because flowering and large plants contributed most to the finite rate of increase of the population (Pfeifer *et al.*, 2006a), we used the ratio of N_R/N and the number of flowers produced per plant as indicators of the condition of the population at each site. At each site, up to 16 flowering individuals were selected randomly for measurements and collection of three flowers each (Table 1). Plant size, inflorescence height and the number of flowers produced per plant were recorded. Fewer

Table 1 Traits of 20 populations of *Himantoglossum hircinum* sampled at five geographical regions within Europe. All measurements were carried out in May–June 2007. Geographical coordinates are given in decimal degrees (World Geodetic System 84). Bold type represents the presence of long-term demographic monitoring plots.

Site	Name		Longitude	Latitude	Alt (m)	N	N_R	N_s
Distribution centre								
CH1	Eclepens	Switzerland	E6.55219	N46.65769	457	5	3	15
FR1	Quintanel	France	E3.49710	N43.94373	774	4	3	12
FR3	Borie	France	E3.53460	N43.88331	570	5	3	15
Northern range margin								
SE1	Burnham	England	W3.00989	N51.25912	21	5	3	7
SE4	Sandwich	England	E1.37919	N51.27925	−2	7	5	8
Southern range margin								
IT3	Sassano	Italy	E15.49464	N40.31806	1085	3	1	9
SP1	Zahara	Spain	W5.38168	N36.79092	1069	4	1	12
SP2	El Bosque	Spain	W5.42007	N36.75565	944	1	1	2
SP3	Grazalema	Spain	W5.39117	N36.81252	690	4	3	10
Atlantic group								
SP4	Sena de Luna	Spain	W5.95016	N42.92917	1141	2	1	9
SP5	Rabonal de Luna	Spain	W5.97038	N42.93824	1137	3	2	10
SP6	Massa	Spain	W3.73995	N42.61883	1017	3	2	9
Eastern range margin								
TH1	Leutrat	Germany	E11.57538	N50.87167	219	7	2	16
BW1	Hirrlingen	Germany	E8.93076	N48.41263	460	6	2	15
BW2	Hirschhalde	Germany	E8.99943	N48.52473	376	4	3	5
BW3	Bobstadt	Germany	E9.68141	N49.46573	306	6	4	16
BW5	Kahlberg	Germany	E9.61833	N49.67374	287	4	3	14
SA1	Billenberg	Germany	E7.31562	N49.16152	295	3	1	8
SA2	Grossbirkel	Germany	E7.30927	N49.15064	304	3	1	4
SA3	Habkirchen	Germany	E7.13894	N49.13732	235	3	2	15

Alt, height above sea level (m); N and N_R , estimates of population size and number of flowering plants (class 1, < 20 plants; class 2, 20–50 plants; class 3, 50–100 plants; class 4, 100–200 plants; class 5, 200–500 plants; class 6, 500–1000 plants; class 7, > 1000 plants); N_s , number of plants sampled per site.

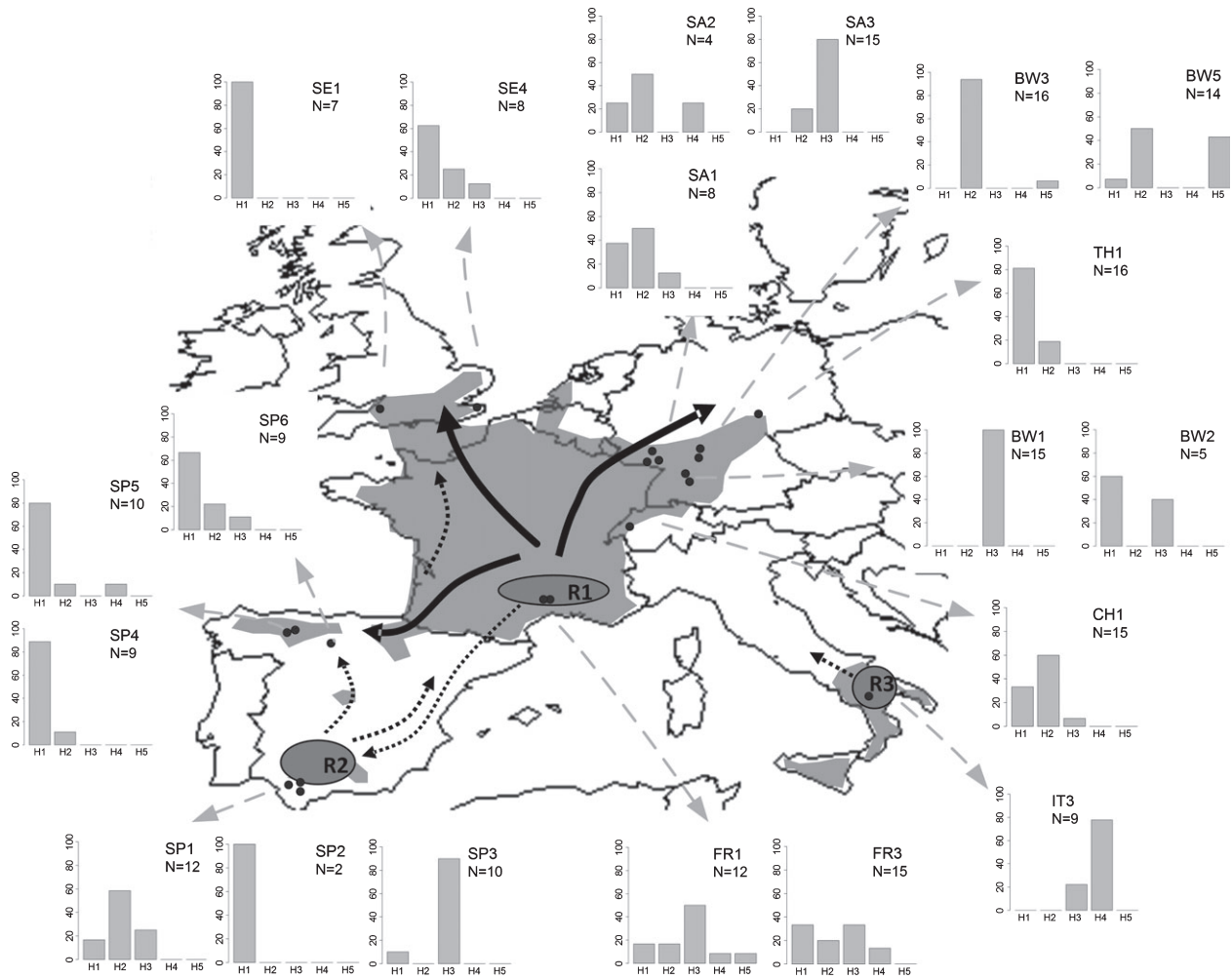


Figure 1 Current distribution (light grey pattern, simplified) and major migration pathways (black arrows) of *Himantoglossum hircinum* in Europe during post-glacial expansion and more recent climate-induced northward range shifts. The dashed arrows indicate alternative pathways. The species probably survived in refugial zones (dark grey, encircled; see Schmitt, 2007) during the last ice age. R1 is the refugial zone of *H. hircinum* in southern France. R2 (Atlantic–Mediterranean element) and R3 (Adriatic–Mediterranean element) represent two major refugial and differentiation centres for many biota in southern Europe. The graphs show the percentage of haplotypes (H1 to H5) at the MSpa1 locus in the sampled populations (black dots). Definitions of the abbreviations for population names are given in Table 1.

plants were sampled if the number of flowering plants was < 10. In total, 211 plants were sampled. Flowers were stored in sealable plastic bags filled with silica gel (white, fine mesh size: 28–200 µm; Sigma-Aldrich Corp., Steinheim, Germany) to ensure rapid drying.

DNA extraction and AFLP analysis

DNA was extracted from dried flowers following a scaled-down 2 × cetyl trimethyl ammonium bromide (CTAB) extraction protocol based on Doyle & Doyle (1987) with minor modifications. DNA samples were cleaned using QIAquick columns (Qiagen, Valencia, CA, USA) and their quality was checked by running the extracts on a 1.0% agarose gel. The DNA concentration was determined using spectrophotometry. Following an initial screening of genetic variation in 16 plants from six different locations, two (*EcoRI*-ACG/

MseI-CAC, *EcoRI*-ACG/*MseI*-CTA) out of 17 AFLP primer combinations tested (see Appendix S1 in Supporting Information) appeared sufficiently polymorphic to discriminate clones within populations and were used in subsequent analyses of 205 plants (Table 2). The AFLP reactions (Vos et al., 1995) were conducted according to a modified AFLP® Plant Mapping Protocol of PE Applied Biosystems Inc. using *EcoRI* and *MseI* with 500 ng of isolated genomic DNA per sample. The success of each step was tested by running the PCR products on a 1.5% agarose gel (20 min, ±100 V). Fragments were run on an ABI Prism® 3100 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA) with 10 µL High Dye (deionized formamide) and 0.2 µL GeneScan™ 500 ROX™ Size Standard per sample. Separated fragments were analysed using GENE SCAN® v2.0 and GENOTYPER® version 2.1 software (Applied Biosystems). Amplified fragments between 55 and 400 bp were scored by visual inspection of

Table 2 Genetic diversity and flowering performance computed for all sample sites of *Himantoglossum hircinum* in Europe (described in Table 1). A total of 205 plants were sampled. Genetic diversity (H_e) and differentiation (F_{ST}) differed among the five geographical regions. Linear modelling using the R 2.7.0 package showed that size had a significant positive effect on number of flowers produced per plant at most sites, while the strength of the relationship differed among sites (R^2_{adj}).

Site	N_s	%p	PB	$H_e \pm SE$	$I \pm SE$	S	Mean(FPP)	R^2_{adj}
Distribution centre: $F_{ST} = 0.111$, $\text{mean}(H_e) = 0.176 \pm 0.02$								
CH1	15	36.7	2	0.160 ± 0.016	0.103 ± 0.011	5	57.6 ± 22.6	0.35
FR1	11	41.9	0	0.203 ± 0.018	0.134 ± 0.013	7	62.0 ± 17.3	0.60
FR3	15	38.1	1	0.165 ± 0.017	0.106 ± 0.011	6	65.7 ± 29.0	0.23†
Northern range margin: $F_{ST} = 0.126$, $\text{mean}(H_e) = 0.127 \pm 0.02$								
SE1	7	15.4	0	0.082 ± 0.014	0.055 ± 0.009	6	51.1 ± 23.1	0.77
SE4	8	34.4	3	0.171 ± 0.018	0.113 ± 0.012	7	64.8 ± 24.5	0.33
Southern range margin: $F_{ST} = 0.238$, $\text{mean}(H_e) = 0.160 \pm 0.06$								
IT3	8	42.8	11	0.217 ± 0.019	0.143 ± 0.013	NM	34.9 ± 11.6	NM
SP1	12	44.7	2	0.204 ± 0.018	0.133 ± 0.012	7	95.8 ± 24.6	0.70
SP2*	2	13.0	2	0.079 ± 0.014	0.054 ± 0.010	7	44.0 ± 12.7	$N = 2$
SP3	8	27.0	2	0.139 ± 0.017	0.093 ± 0.011	8	121.1 ± 42.2	0.47
Atlantic group: $F_{ST} = 0.060$, $\text{mean}(H_e) = 0.181 \pm 0.05$								
SP4	9	27.0	1	0.122 ± 0.015	0.079 ± 0.011	8	89.8 ± 16.2	0.62
SP5	10	43.7	2	0.213 ± 0.018	0.140 ± 0.013	7	55.8 ± 20.9	0.42
SP6	8	39.5	1	0.208 ± 0.019	0.140 ± 0.013	5	54.0 ± 15.3	0.20
Eastern range margin: $F_{ST} = 0.128$, $\text{mean}(H_e) = 0.137 \pm 0.03$								
BW1	15	39.5	3	0.165 ± 0.016	0.105 ± 0.011	7	49.9 ± 11.8	0.47
BW2*	5	11.6	0	0.062 ± 0.012	0.042 ± 0.008	6	46.5 ± 16.1	0.11
BW3	15	37.7	0	0.163 ± 0.017	0.106 ± 0.012	4	70.3 ± 33.3	0.46
BW5	14	28.4	0	0.139 ± 0.017	0.092 ± 0.011	7	61.9 ± 19.0	0.58
SA1	8	23.7	0	0.123 ± 0.016	0.082 ± 0.011	6	56.4 ± 11.9	NS
SA2*	4	24.2	0	0.135 ± 0.017	0.092 ± 0.012	5	32.0 ± 4.6	NS
SA3	15	33.0	0	0.145 ± 0.016	0.094 ± 0.011	6	67.1 ± 19.9	0.44
TH1	16	37.7	1	0.160 ± 0.017	0.104 ± 0.012	5	37.6 ± 12.7	0.38

N_s , number of plants sampled per site; H_e , average expected heterozygosity (equivalent to Nei's (1973) gene diversity); I , Shannon information index; %p, percentage of polymorphic loci; PB, number of private bands; S, minimum number of leaves of flowering plants at a site; Mean(FPP), average number of flowers produced per plant at a site; NM, not measured; NS, not significant. *Sample size below 6, so the result has to be treated cautiously. †Excluding two outliers from the analyses increased R^2 to 0.69.

electropherograms for presence (1) or absence (0) of peaks. Analyses were confined to primer combinations that generated clearly readable profiles (Bonin *et al.*, 2007). Manual scoring was repeated twice to estimate the effects of genotyping error. Double-reading of plants was repeated with an emphasis on individual identity in the case of inconsistency in individual banding patterns to ensure a high accuracy of band matching. The full AFLP analysis was repeated for 16 samples.

Statistical analysis of the genetic structure

The number of private bands (PB), average heterozygosity (H_e), Shannon information index (I) and percentage of polymorphic loci (%p) were computed for each population, assuming Hardy–Weinberg equilibrium using GENALEX version 6.1 (Peakall & Smouse, 2006). Genetic diversity within and among populations was compared and regional differences in genetic variance were computed using the R 2.7.0 package (www.R-project.org).

To analyse population genetic structure, non-spatial Bayesian clustering was employed using STRUCTURE version 2.2

(Pritchard *et al.*, 2000; Falush *et al.*, 2003). During STRUCTURE simulations, individuals are probabilistically assigned to populations, or jointly to more populations if their genotypes indicate that they are admixtures (have mixed ancestry). An 'admixture' model was run with correlated allele frequencies, 50,000 Markov chain Monte Carlo (MCMC) iterations of burn-in period length to minimize the effects of the starting configuration and 50,000 MCMC replicates after burn-in for accurately assessing parameters. To estimate the K number of ancestral genetic populations and the ancestry membership proportions of each individual in these clusters, the algorithm was run 10 times for each user-defined K value from 2 to 20. The final K was estimated as the largest K value with a significantly higher likelihood than that from $K - 1$ runs (two-sided nonparametric Wilcoxon test, $P < 0.005$; Picó *et al.*, 2008). The true number of K groups was also determined by plotting the mean likelihood of K and the rate of change of the likelihood distribution (mean \pm SD) over 10 runs for each K value (Evanno *et al.*, 2005). Membership fractions of individuals in the inferred clusters were computed and displayed using DISTRUCT (Rosenberg, 2004). ARLEQUIN version 2000

(Schneider *et al.*, 2000) was used to carry out an analysis of molecular variance (AMOVA) to quantify population differentiation within and between regions, testing four models of population groups (Table 3). AFLP-SURV 1.0 (Vekemans *et al.*, 2002) was employed to produce a phylogenetic tree, computing F_{ST} values based on the Bayesian method with non-uniform priors (Zhivotovsky, 1999) as recommended by Bonin *et al.* (2007). NEIGHBOUR and DRAWGRAM (both in the software package PHYLIP version 3.67; Felsenstein, 2005) were then used to draw UPGMA (unweighted pair group method with arithmetic mean) dendrograms for the 1000 bootstrapped matrices.

Isolation-by-distance patterns were investigated across all populations and at the group level to reveal any specific barriers to gene flow that might be present. Genetic distances among population pairs were computed using the Zhivotovsky (1999) method and the square-root method (Lynch & Milligan, 1994). Geographical distances between population pairs were calculated as the shortest distance between two points on a sphere using the online calculator CACTUS2000 (<http://www.cactus2000.de/uk/unit/massgrk.shtml>). IBD version 1.52 (Bohannan, 2002) was employed to plot genetic distances as a function of geographical distances between population pairs, to test for significance in the isolation-by-distance relationship with a Mantel test for matrix correlation (1000 randomizations), and to determine the strength of this relationship by regressing all pairwise genetic similarity values against their corresponding geographical distances using reduced major axis regression.

Geographical trends in genetic and demographic structure were analysed by computing the longitudinal gradient from distances between the longitude of a site and the average longitude of populations FR1 and FR3 (= 3.5), which represent core sites, and the latitudinal gradient from distances between the latitude of a site and the average latitude of populations FR1 and FR3 (= 43.9). We tested for significant trends in genetic traits and demographic traits with increasing

Table 3 Analyses of molecular variance (AMOVA) testing four models of group formation. F statistics were computed from a matrix of Euclidean squared distances between every pair of individuals. While all populations were treated equally in model 1, models 2–4 tested different regional relationships (see Table 1 for assignment of populations to regions). All fixation indices (F_{SC} , among populations and within groups; F_{ST} , among individuals within populations; F_{CT} , among groups) were significant, except one in model 4.

Model	F_{SC}	F_{ST}	F_{CT}
Model 1	–	0.146	–
Model 2	0.137	0.150	0.014
Model 3	0.092	0.129	0.041
Model 4	0.077	0.086	0.010*

Model 1, all populations; model 2, core, Northern, Southern, Atlantic, Eastern; model 3, South France versus South Spain; model 4, South France versus Atlantic group.

*1023 permutations, F_{CT} non-significant, P (random value > observed value) smaller than 0.10.

latitudinal and longitudinal distance from the distribution centre using linear modelling in the R 2.7.0 modelling package. The Wilcoxon test was used to check for significant differences in genetic and demographic traits between geographical regions using R 2.7.0.

Microsatellite analysis

Ten universal plastid markers (see Appendix S2) were screened for sequence variation among populations with proof-reading Taq using the ABI 377 DNA Automated Sequencer (Applied Biosystems). The primer combination *accD* 769F vs. *psaI* 75R (amplifying the *accD*–*psaI* intergenic spacer) was able to produce sufficient variability between populations in form of multiple repeats (C)₈ to (C)₁₁ followed by (T)₈ to (T)₉ between base pairs (bp) 618 and 637. A forward and a reverse primer (MSpsa1_F and MSpsa1_R) were designed for subsequent analysis of length variation at this locus for 211 plants: MSpsa1_F (5'-AAG CAT CCC TCT CTT GAC AA-3') and MSpsa1_R (JOE labelled at 5'-end) (5'-CAA CAA ACA GGG ATT CCT AG-3').

Amplifications were performed in 10- μ L reaction volumes with 0.4 μ L extracted DNA, 0.2 μ L of each of the primers, 0.2 μ L bovine serum albumin (BSA) buffer and 0.2 μ L of polymerase chain reaction (PCR) ready-mix (ABgene PCR Mastermix, 1.5 mM Mg, Abgene Limited, Epsom, UK). After a denaturation step at 94°C for 2 min, PCRs were performed with 30 cycles of 30 s denaturing at 93°C, 60 s of annealing at 48°C, 60 s of extension at 72°C, and a final 8-min extension step at 72°C. The PCR products were purified using QIAquick columns (Qiagen). Amplification success was confirmed on 1% agarose gels stained with ethidium bromide. The PCR products were diluted and run on an ABI Prism 310 Genetic Analyzer mixing 1 μ L of the diluted sample with 10 μ L of a formamide solution (High Dye) and 0.2 μ L of GeneScan™ 500 ROX™ Size Standard (Applied Biosystems). Microsatellite length variation at the MSpsa1 microsatellite locus was analysed to compute the relative proportion of haplotypes present in each population using GENESCAN® Analysis Software version 2.0 and GENOTYPER® Software version 2.1 (Applied Biosystems Inc., Foster City, CA, USA). Distinct peaks were scored as present by visual inspection and the scoring procedure was repeated twice on separate occasions to minimize inconsistencies in scoring.

RESULTS

Geographical structure in demographic traits

N_R and N differed among sites (Table 1). The number of flowers produced per plant varied strongly within and among sites (Table 2). There was a positive, significant influence of plant size on the number of flowers produced per plant at all sites except SA1 and SA2 (Table 2). The ratio N_R/N decreased significantly with increasing latitudinal distance from the distribution centre (Fig. 2d). The mean number of flowers

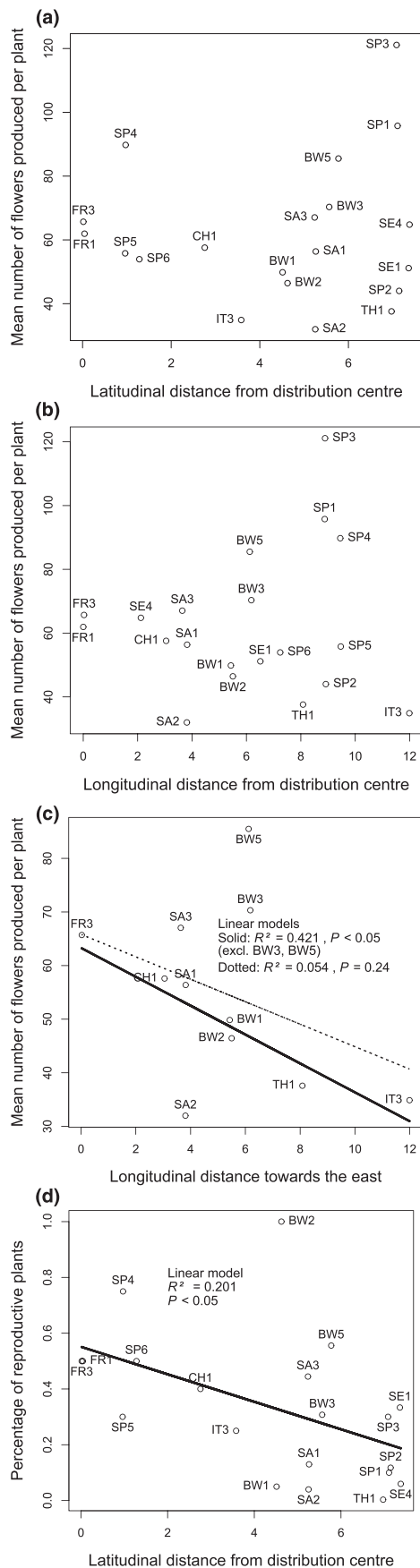


Figure 2 Geographical trends of demographic traits of sampled populations of *Himantoglossum hircinum* in Europe. The range of the mean number of flowers produced per plant increased with increasing latitudinal (a) and longitudinal distance (b) from the distribution centre. The mean number of flowers produced per plant decreased significantly towards the eastern range margin when excluding two outlier populations, BW3 and BW5, from the linear model (c). The ratio N_R/N decreased significantly with increasing latitudinal distance from the distribution centre (d).

produced per plant decreased significantly with increasing longitudinal distance towards the east when excluding two statistical outliers, BW3 and BW5, from the linear model ($P < 0.05$, $R^2 = 0.421$, Fig. 2c). The range of this parameter increased considerably with increased longitudinal distance from the distribution core (Fig. 2a,b). The flowering percentage and size of both populations, BW3 and BW5, were high despite their location at the eastern range margin.

Further significant differences among regions or geographical trends (considering latitudinal or longitudinal distance from distribution centre, either in general, or towards the east/west or the north/south) in demographic traits (considering population size, percentage of reproductive plants, mean number of flowers per plant) were not found.

Molecular analyses

The AFLP data matrix included 215 fragments (mean fragment size: 144 ± 62). Of these, 205 (95.35%) were polymorphic. Fragment size and frequency were not correlated. Genotyping error was negligible for microsatellite analyses, because peaks were clearly readable for all 211 samples and double-reading them resulted in 100% accuracy. Selective amplification repeated for 16 AFLP samples produced identical banding patterns. For 11 samples, all steps were repeated (starting from DNA extraction), and genotyping error was 6% (using only 120 loci that were absolutely clear) and 11% (using all 205 loci).

Geographical genetic structure

Five haplotypes (H1 to H5) were recognized across all populations based on the MSpsa1 locus. Haplotype number (N_{haplo}) differed strongly among sites. The two core populations, FR1 ($N_{\text{haplo}} = 4$) and FR3 ($N_{\text{haplo}} = 5$), had significantly more haplotypes compared with all other populations (Wilcoxon test, $P < 0.05$). Haplotype frequencies differed between geographical areas (Fig. 1). H1 dominated in Atlantic (78.6%) and northern (80.0%) populations, H2 (31.8%) and H3 (54.6%) in southern Spain (50%) and H4 in Italy.

Populations differed strongly in their genetic diversity, with high intraregional variability at the range margins (Table 2). A few populations had no exclusive bands and many had just one or two exclusive bands. The Italian population had 11 private bands. Average genetic diversity was lower in northern (not

significant) and eastern peripheral populations (significant, Wilcoxon test, $P < 0.05$) compared with that in core populations, but was only slightly lower for southern peripheral populations. Atlantic populations exhibited relatively high genetic diversity. When excluding outlying populations from the analyses, geographical trends in genetic diversity became significant (Fig. 3a,b).

All fixation indices were significant except for differentiation among groups in model 4 (Table 3). Molecular variance within populations was much higher than variance among populations, resulting in a fixation index of $F_{ST} = 0.146$ (Model 1). At the regional level, genetic differentiation increased from the geographical range centre to the geographical range margins (Table 2). Genetic differentiation among Atlantic populations

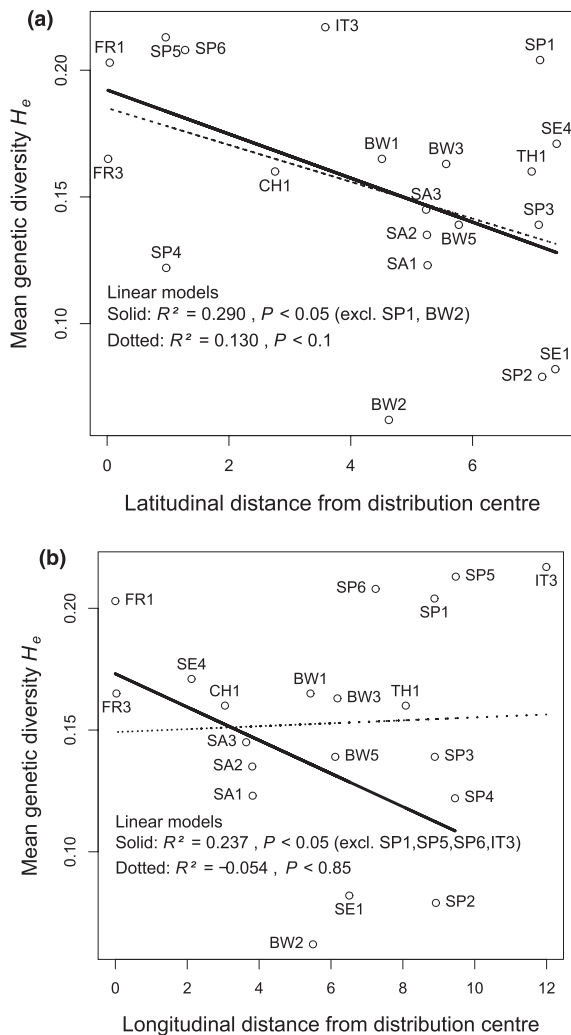


Figure 3 Geographical trends of genetic diversity (H_e) of *Himantoglossum hircinum* populations within its European distribution. (a) Mean genetic diversity plotted against increasing latitudinal distance from the distribution centre. (b) Mean genetic diversity plotted against increasing longitudinal distance from the distribution centre. Dashed lines represent predictions of linear models when including all populations; black lines when excluding statistical outliers.

was low even when grouping them together with populations SE1 and SE4 at the northern range margin ($F_{CT} = 0.028$, not significant). Southern French populations were not strongly differentiated from southern Spanish populations or Atlantic populations (Table 3). Analysed over all populations, genetic distance increased significantly with geographical distance (Lynch & Milligan: $r = 0.27$, $P < 0.05$; Zhivotovsky: $r = 0.31$, $P < 0.01$). No correlations were detected between genetic diversity (e.g. diversity indices, haplotype frequency) and population demographic traits.

The phylogenetic tree did not reveal major groups in Europe. However, most populations at the eastern range margin clustered together except for SA2 and BW2 (Fig. 4). The Italian site was strongly differentiated from all other sites. Six ancestral genetic clusters (C1–C6, Wilcoxon test $P < 0.05$) were inferred with STRUCTURE ('admixture', see Appendix S3). Membership fractions of individuals in the six clusters varied between seven geographical regions (Fig. 5). Cluster three (C3) had high membership proportions in southern Spain and the Atlantic enclave (Fig. 5). Central Germany showed membership fractions that differed markedly from other sites within the eastern range margin (see Appendix S4). Cluster five (C5) was most pronounced in southern France and southern Spain. French and English sites were characterized by high membership fractions in cluster one (C1). The eastern range margin and the Swiss population shared substantial membership fractions in all clusters. Cluster six (C6) presented high proportions at the eastern and northern margins and at the Swiss site (> 59%), intermediate proportions in southern France and the Atlantic group (38% and 39%), but low to very low proportions in southern Spain and Italy (Fig. 5). The Italian site was the only site where cluster two (C2) and cluster four (C4) had high membership fractions.

DISCUSSION

Geographical population structure

Decreased habitat suitability towards the range margins is expected to reduce population size and increase isolation among populations, resulting in the loss of genetic diversity and net flux of genes from central to marginal sites ('swamping'; Kirkpatrick & Barton, 1997). While the reduced genetic diversity and haplotype frequency at many eastern marginal sites of *H. hircinum* matches expectations for the genetic consequences of historical range expansions (Hewitt, 1996), the decrease in genetic diversity towards the range margins of the species is not consistent. Some sites at the southern periphery (IT3, SP1) and within the Atlantic group were among the populations with highest genetic diversity ($H_e > 0.20$). We could detect significant trends in genetic diversity along geographical gradients after excluding outlying populations (Fig. 3a,b). Population history and different processes limiting population growth at different boundaries of a species (Bullock *et al.*, 2000) may be more important for the genetic structure of those outliers. This might also explain

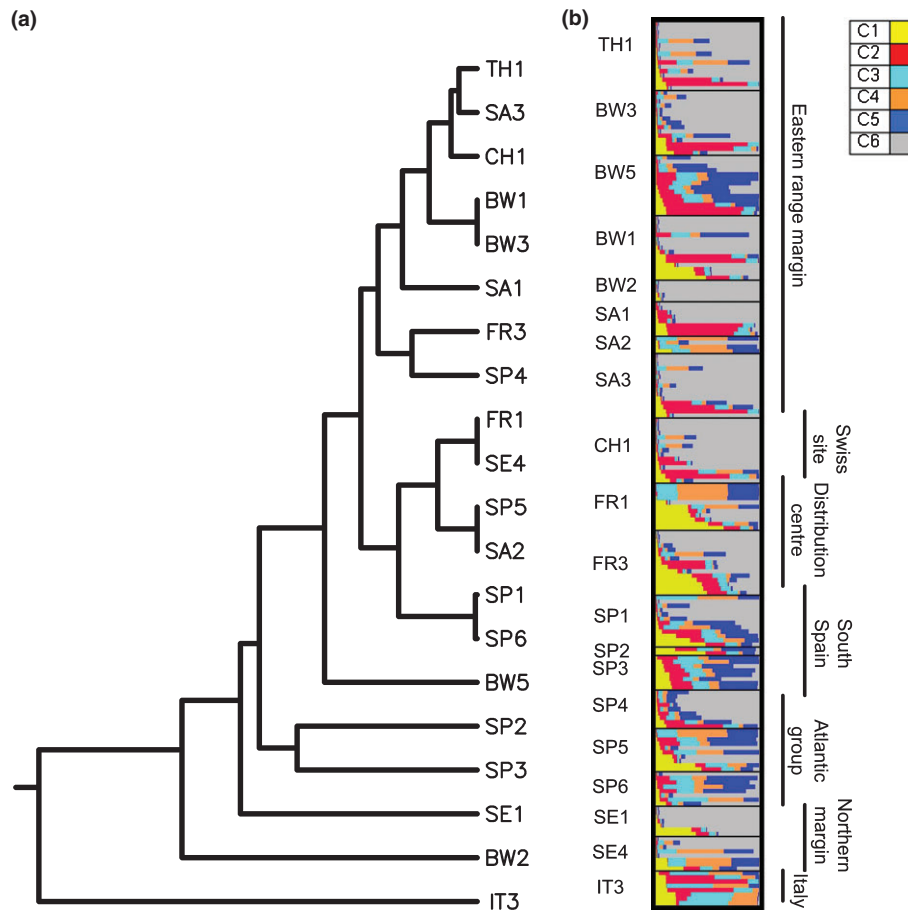


Figure 4 Genetic structure of *Himantoglossum hircinum* within its entire distribution area in Europe. (a) The UPGMA (unweighted pair group method with arithmetic mean) dendrogram was based on 1000 bootstrapped F_{ST} distance matrices. F_{ST} values were calculated from allele frequencies estimated according to the Bayesian method with non-uniform priors (Zhitovitsky, 1999). (b) DISTRUCT plot for 20 *H. hircinum* populations within Europe. The graph is based on the average of 10 STRUCTURE runs ('admixture model', allele frequencies correlated) showing the occurrence of six ancestral genetic clusters (C1–C6) in the sampled populations. Each plant is represented by a line partitioned into six segments corresponding to its membership coefficients in the six inferred clusters. Each cluster is represented by a different colour. Black segments separate plants of different populations (names beside the figure). Populations are classified into seven geographical regions named beside the figure. Plants within each population are arranged according to estimated cluster membership proportions.

the lack of correlation between population genetic and demographic structure at the geographical scale.

Changed habitat suitability (especially with regard to weather) beyond the original range limits allowed for rapid population growth at some eastern and northern marginal sites (TH1, SE4, BW3), which probably prevented loss of diversity. In contrast, populations at the southern range margin exhibit signs of increased stress, including low population sizes (SP2 and IT3) and a small number of flowering plants (IT3, SP1, SP2), because of anomalous weather in the winter and spring of 2007. Spatially variable fruit set (around 30% at FR1 and FR3, 95% in Italy, but only 1.5–20.83% in eastern peripheral sites; M. Pfeifer, unpublished data) may compensate for lower flowering percentages and might explain some of the high diversity observed at IT3 and SP1.

Although the Bayesian method of clustering genotypes did not detect very clear patterns (Fig. 4), significant fixation indices and isolation-by-distance patterns provide evidence for

some genetic structuring of *H. hircinum* on the geographical scale. Among others, intrinsic problems related to AFLP development, such as ascertainment bias or linkage among alleles may have contributed to lower resolution of phylogeography when using STRUCTURE.

In general, the results present a picture of a widespread taxon with low genetic differentiation. Higher genetic differentiation among populations at the northern and eastern range margins (Table 2) can probably be attributed to increased isolation between sites, where the species is often bound to certain habitat types, including golf courses in England and protected dry grasslands on southward-facing slopes in Germany. This is in contrast to the frequent occurrence of the species in France, where it grows in all regions except Finistère (north-western France) and Corsica (Bournérias & Prat, 2005). Lower genetic differentiation within the Core group ($F_{ST} = 0.11$, maximum geographical distance (MD) = 388 km) and the Atlantic group ($F_{ST} = 0.060$, MD = 185 km)

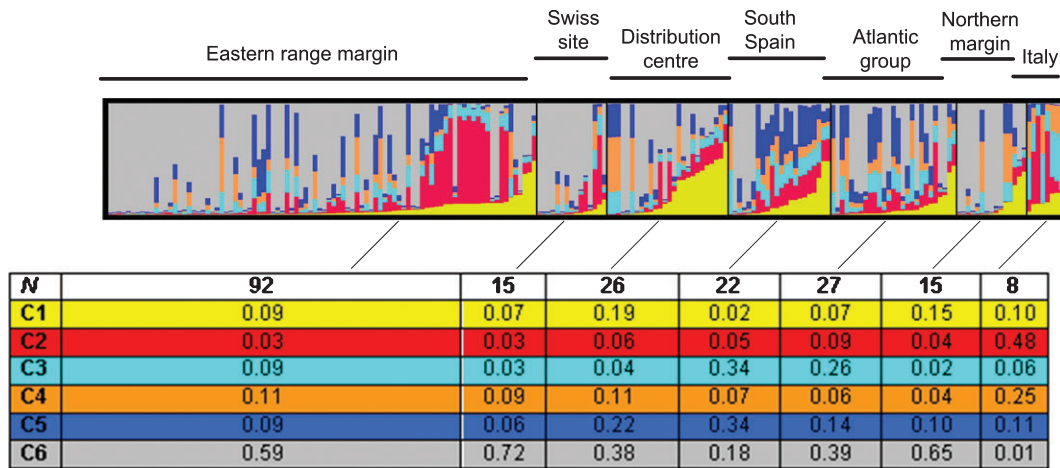


Figure 5 Genetic structure of seven major geographical regions within the European distribution of *Himantoglossum hircinum*. The DISTRUCT plot is based on the average of 10 STRUCTURE runs ('admixture model', allele frequencies correlated) showing the occurrence of six ancestral genetic clusters (C1–C6) in the sampled regions. Each plant is represented by a line partitioned into six segments corresponding to its membership coefficients in the six inferred clusters. Each genetic cluster is represented by a different colour. Black segments separate plants of seven different geographical regions named above the figure. Plants within each region are arranged according to estimated cluster membership proportions. Bottom: the number of individuals (*N*) and the mean membership fractions in the six clusters (C1–C6) in the regions.

can at least be partly explained by low maximum geographical distances. However, genetic differentiation did not increase substantially when grouping Atlantic populations with the two sites in southern England ($F_{ST} = 0.09$), which are both more than 950 km away.

For some populations (SA2, SP2, BW2), only a few samples ($N_s < 6$) could be collected due to low flowering performance in the year of sampling, and thus interpretations have to be treated cautiously (Bonin *et al.*, 2007). However, reanalyses of data after excluding these sites showed no great alteration of results with regard to regional haplotype frequencies, geographical trends in demographic and genetic traits, and genetic differentiation within and among groups. Temporarily varying weather conditions may lead also to erroneous estimates of population fitness in any one year. For example, flowering estimates were made in 2007, but subsequent observations in 2008 revealed large fluctuations in flowering at some sites (e.g. SP3: N_R and individual flower production high in 2007, but low N_R in 2008).

Phylogeographical structure

Isolation of populations in glacial refugia and subsequent recolonization will have left their mark on the structure of modern populations (Thompson, 1999; Skog *et al.*, 2008; Swart *et al.*, 2008). The Mediterranean basin has undergone repeated glaciation episodes since the late Pliocene (c. 2.5 Ma). During the last glaciation, plants of *H. hircinum* probably survived in four Mediterranean refugia (Fig. 1) including southern France, the Atlantic Mediterranean (e.g. southern Spain), the Adriatic–Mediterranean (e.g. southern Italy) and the Pontic–Mediterranean (e.g. Balkan Peninsula) spreading northwards and eastwards via three major colonization routes after the ice retreated (Taberlet *et al.*, 1998). This interpreta-

tion is similar to reports for other Mediterranean species (Demesure *et al.*, 1996; Thompson, 1999) and is supported by regional membership fractions in the six gene pools derived via STRUCTURE simulations (Fig. 5, Appendix S4) and the spatial distribution of Mpsa1 haplotypes (Fig. 1).

Populations of *H. hircinum* originating from the Italian refugium remained confined to this region. This genetic–geographical split coincided with morphological differences (e.g. flower morphology; N. Passalacqua, unpublished data) in comparison with other species such as *Spiranthes romanzoffiana* (Forrest *et al.*, 2004). Southern Italy was a refuge for both Central European and Balkan species during previous glaciations, leading to many relict species clearly showing both distribution patterns, and the presence of species of hybrid origin, the parental species of which are one with European distribution and the other with Balkan distribution (Petit *et al.*, 2002). If haplotype H4 at the Mpsa1 locus, which dominates at the Italian site, occurs in *Himantoglossum adriaticum* (Balkan distribution), we can infer an old contiguous distribution of the two species with gene flow in the contact zone in southern Italy.

The observed loss of genetic diversity of *H. hircinum* during the eastward spread (via central and south-west Germany) from the southern French refugium (all five haplotypes present) is typical for expansion processes and characteristic for species with specific habitat requirements (Demesure *et al.*, 1996). The dominance of haplotype H1 in Atlantic and English populations may indicate northward spread via northern Spain, similar to that hypothesized for the Lusitanian flora in Ireland. However, *H. hircinum* does not occur in Ireland and distances across the sea are large. It is more likely that plants emerging from southern France or Spain migrated via northern France to the south of England in a similar way to the spread of *Quercus* species (Hewitt, 2001).

Populations in northern Spain are likely to have been founded by a mixture of immigrants from France and southern Spain. This 'secondary contact hypothesis' is supported by membership fractions in the STRUCTURE clusters (Fig. 5) and weak east–west differentiation in the Atlantic–Mediterranean element, with southern Spain being more strongly differentiated from southern France than the Atlantic group (Table 3). Modelling using 'a priori population information' in STRUCTURE (results not shown) suggests that some plants in the Atlantic sites SP5 and SP6 have mixed ancestry with high probabilities for these individuals originating from populations FR1 or FR3.

Differentiation in orchids is often low because they produce tiny 'dust' seeds that can float in the air for long distances, favouring gene exchange among populations (Arditti & Ghani, 2000; Fay *et al.*, in press). However, the mean distance travelled by *H. hircinum* seeds was found to be only 1 m (Carey, 1999). Furthermore, pollen dispersal distances appear to be comparatively low, at least in German populations (Kropf & Renner, 2008). However, time since colonization has probably not been sufficient to detect the effects of restricted pollen and seed dispersal, such as inbreeding and increased genetic structuring (Wright, 1978). Some populations, including SE4 at the northern and TH1 at the north-eastern range margin, have been at the sites since 1940 and have therefore been through only approximately seven generations.

Factors affecting geographical genetic structure

Results from the scientific literature are controversial, partly supporting genetic implications of the CMC and partly contradicting it (Eckert *et al.*, 2008). Our study on *H. hircinum* failed to support the genetic implications derived from the CMC with significant evidence. However, we conclude that for those sites where genetic diversity was higher than expected from the CMC, the geographical pattern is probably masked by other factors acting singly or in combination, to affect either demography or population genetic structure directly. We suggest that post-glacial recolonization patterns and more recent range shifts induced by climate change play a major role in the structuring of geographical genetics of *H. hircinum*. High genetic diversity and substantial differentiation of Italian sites are probably the consequences of survival during the ice age in refugia and post-glacial speciation and hybridization processes. The genus *Himantoglossum* is not well-delimited, and hybridization among species (e.g. with *H. adriaticum* in southern Italy) seems to take place frequently (Bernardos *et al.*, 2006). Secondary contact in northern France/northern Spain may have contributed to high diversity in the Atlantic group. Our findings emphasize that to understand geographical structuring of genetic diversity, historical (e.g. glaciation, climate-change-related range shifts) and further effects need to be distinguished from those caused by geographical variation in population demography and dispersion (Channell & Lomolino, 2000).

ACKNOWLEDGEMENTS

Two anonymous referees provided helpful comments on an earlier draft of this manuscript. Peter Steinfeld, Helmut Heimeier, Michael Koltzenburg, Bernard Haynold, Beat Bäumler, Sonia Bernardos, Roberto Gamarra and Eric Walravens supplied information on population locations. Special thanks go to Edward Amiry and Robyn Cowan for their introduction to and help with all molecular genetic analyses. The Royal Botanic Gardens, Kew, wishes to acknowledge the financial support of Natural England. This study is part of the research MODGLOB funded by the German Research Community. M.P. received a travel grant from the European Science Foundation (CONGEN SV/1075) to carry out population surveys and to visit collaborating scientists in Spain.

REFERENCES

- Arditti, J. & Ghani, A.K.A. (2000) Tansley Review No. 110. Numerical and physical properties of orchid seeds and their biological implications. *New Phytologist*, **145**, 367–421.
- Bernardos, S., Santos, M.A., Tyteca, D. & Amich, F. (2006) Phylogenetic relationships of Mediterranean Neottieae and Orchideae (Orchidaceae) inferred from nuclear ribosomal ITS sequences. *Acta Botanica Gallica*, **153**, 153–165.
- Blows, M.W. & Hoffmann, A.A. (2005) A reassessment of limits to evolutionary change. *Ecology*, **86**, 1371–1384.
- Bohonak, A.J. (2002) IBD (Isolation by Distance): a program for analyses of isolation by distance. *Journal of Heredity*, **93**, 153–154.
- Bonin, A., Ehrich, D. & Manel, S. (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, **16**, 3737–3758.
- Bournérias, M. & Prat, D. (2005) *Les orchidees de la France, Belgique et du Luxembourg*. Biotope, Paris.
- Brown, J.H. (1984) On the relationship between abundance and distribution of species. *The American Naturalist*, **124**, 255–279.
- Bullock, J.M., Edwards, R.J., Carey, P.D. & Rose, R.J. (2000) Geographical separation of two *Ulex* species at three spatial scales: does competition limit species' ranges? *Ecography*, **23**, 257–271.
- Carey, P.D. (1999) Changes in the distribution and abundance of *Himantoglossum hircinum* (L.) Sprengel (Orchidaceae) over the last 100 years. *Watsonia*, **22**, 353–364.
- Carey, P.D. & Farrell, L. (2002) *Himantoglossum hircinum* (L.) Sprengel. *Journal of Ecology*, **90**, 206–218.
- Case, T.J. & Taper, M.L. (2000) Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. *The American Naturalist*, **155**, 583–605.
- Channell, R. & Lomolino, M.V. (2000) Dynamic biogeography and conservation of endangered species. *Nature*, **403**, 84–86.
- Demesure, B., Comps, B. & Petit, R.J. (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution*, **50**, 2515–2520.

- Dorken, M.E. & Eckert, C.G. (2001) Severely reduced sexual reproduction in northern populations of a clonal plant *Decodon verticillatus* (Lythraceae). *Journal of Ecology*, **89**, 339–350.
- Doyle, J.J. & Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.
- Eckert, C.G., Samis, K.E. & Lougheed, S.C. (2008) Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170–1188.
- Ellstrand, N.C. & Hoffman, C.A. (1990) Hybridization as an avenue of escape for engineered genes – strategies for risk reduction. *BioScience*, **40**, 438–442.
- 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.
- Falush, D., Stephens, M. & Pritchard, J.K. (2003) Inference of population structure using Multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.
- Fay, M.F., Bone, R., Cook, P., Kahandawala, I., Greensmith, J., Harris, S., Pedersen, H.E., Ingrouille, M.J. & Lexer, C. (in press). Genetic diversity in *Cypripedium calceolus* (Orchidaceae) with a focus on northwestern Europe, as revealed by plastid DNA length polymorphisms. *Annals of Botany*.
- Felsenstein, J. (2005) *PHYLIP (Phylogeny Inference Package) version 3.6*. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle, WA, USA.
- Forrest, A.D., Hollingsworth, M.L., Hollingsworth, P.M., Sydes, C. & Bateman, R.M. (2004) Population genetic structure in European populations of *Spiranthes romanzoffiana* set in the context of other genetic studies on orchids. *Heredity*, **92**, 218–227.
- Giles, B.E. & Goudet, J. (1997) Genetic differentiation in *Silene dioica* metapopulations: estimation of spatiotemporal effects in a successional plant species. *The American Naturalist*, **149**, 507–526.
- Good, R. (1936) On the distribution of lizard orchid (*Himantoglossum hircinum* Koch). *New Phytologist*, **35**, 142–170.
- Hengeveld, R. & Haeck, J. (1982) The distribution of abundance. 1. Measurements. *Journal of Biogeography*, **9**, 303–316.
- Hewitt, G.M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.
- Hewitt, G.M. (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.
- Jump, A.S. & Woodward, F.I. (2003) Seed production and population density decline approaching the range-edge of *Cirsium* species. *New Phytologist*, **160**, 349–358.
- Kirkpatrick, M. & Barton, N.H. (1997) Evolution of a species' range. *The American Naturalist*, **150**, 1–23.
- Kluth, C. & Bruehlheide, H. (2005) Central and peripheral *Hornungia petraea* populations: patterns and dynamics. *Journal of Ecology*, **93**, 584–595.
- Kropf, M. & Renner, S. (2008) Pollinator-mediated selfing in two deceptive orchids and a review of pollinium tracking studies addressing geitonogamy. *Oecologia*, **155**, 497–508.
- Lynch, M. & Milligan, B.G. (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91–99.
- Maurer, B.A. & Taper, M.L. (2002) Connecting geographical distributions with population processes. *Ecology Letters*, **5**, 223–231.
- Mehlman, D.W. (1997) Change in avian abundance across the geographic range in response to environmental change. *Ecological Applications*, **7**, 614–624.
- Nantel, P. & Gagnon, D. (1999) Variability in the dynamics of northern peripheral versus southern populations of two clonal plant species, *Helianthus divaricatus* and *Rhus aromatica*. *Journal of Ecology*, **87**, 748–760.
- Parsons, P. (1991) Evolutionary rates: stress and species boundaries. *Annual Review of Ecology and Systematics*, **22**, 1–18.
- Peakall, R. & Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.
- Petit, R.J., Brewer, S., Bordacs, S., Burg, K., Cheddadi, R., Coart, E., Cottrell, J., Csaikl, U.M., van Dam, B., Deans, J.D., Espinel, S., Fineschi, S., Finkeldey, R., Glaz, I., Goicoechea, P.G., Jensen, J.S., König, A.O., Lowe, A.J., Madsen, S.F., Matyas, G., Munro, R.C., Popescu, F., Sladke, D., Tabbener, H., de Vries, S.G.M., Ziegenhagen, B., de Beaulieu, J.-L. & Kremer, A. (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49–74.
- Pfeifer, M., Wiegand, K., Heinrich, W. & Jetschke, G. (2006a) Long-term demographic fluctuations in an orchid species driven by weather: implications for conservation planning. *Journal of Applied Ecology*, **43**, 313–324.
- Pfeifer, M., Heinrich, W. & Jetschke, G. (2006b) Climate, size and flowering history determine flowering pattern of an orchid. *Botanical Journal of the Linnean Society*, **151**, 511–526.
- Picó, F.X., Méndez-Vigo, B., Martínez-Zapater, J.M. & Alonso-Blanco, C. (2008) Natural genetic variation of *Arabidopsis thaliana* is geographically structured in the Iberian Peninsula. *Genetics*, **180**, 1009–1021.
- Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Rosenberg, N.A. (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.
- Sagarin, R.D. & Gaines, S.D. (2002) The 'abundant center' distribution: to what extent is it a biological rule? *Ecology Letters*, **5**, 137–147.
- Schmitt, T. (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, **4**, 11.

- Schneider, S., Roessli, D. & Excoffier, L. (2000) *ARLEQUIN ver. 2.000: a software for population genetics data analysis*. Distributed by the authors, University of Geneva, Switzerland.
- Skog, A., Zachos, F.E., Rueness, E.K., Feulner, P.G.D., Mysterud, A., Langvatn, R., Lorenzini, R., Hmwe, S.S., Lehoczy, I., Hartl, G.B., Stenseth, N.C. & Jakobsen, K.S. (2008) Phylogeography of red deer in (*Cervus elaphus*) in Europe. *Journal of Biogeography*, **36**, 66–77.
- Swart, B.L., Tolley, K.A. & Matthee, C.A. (2008) Climate change driven speciation in the southern rock agama (*Agama atra*) in the Cape Floristic Region, South Africa. *Journal of Biogeography*, **36**, 78–87.
- 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.
- Thompson, J.D. (1999) Population differentiation in Mediterranean plants: insights into colonization history and the evolution and conservation of endemic species. *Heredity*, **82**, 229–236.
- Vekemans, X., Beauwens, T., Lemaire, M. & Roldan-Ruiz, I. (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139–151.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M. & Zabeau, M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.
- Wright, S. (1978) *Evolution and the genetics of populations*, Vol. 4. *Variability within and among natural populations*. University of Chicago Press, Chicago.
- Zhivotovsky, L. (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, **8**, 907–913.

SUPPORTING INFORMATION

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

Appendix S1 Amplified fragment length polymorphism (AFLP) primer combinations tested to analyse the geographical genetic structure of *Himantoglossum hircinum*.

Appendix S2 Primer combinations used for screening a subset of *Himantoglossum hircinum* plants for microsatellite length variation.

Appendix S3 Non-spatial Bayesian clustering of plants to test for genetically independent groups using STRUCTURE version 2.2.

Appendix S4 DISTRUCT plot for 20 *Himantoglossum hircinum* populations within 10 subregions of its distribution area.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

BIOSKETCH

Marion Pfeifer is a plant ecologist and received her PhD at the Institute of Ecology, University of Jena. The DFG (German Research Community)-funded research project (MODGLOB), which started in January 2007, analyses how population history, demographic variability, environmental fluctuations and genetic processes are interacting over geographical scales to determine the adaptive potential of plant species and their capacity to respond to environmental changes. The aim is to improve predictions of habitat models by incorporating results from population-level analyses into a Bayesian-based habitat model.

Author contributions: M.P., F.X.P. and M.F.F. conceived the ideas; M.P., F.X.P., B.S., N.G.P., F.J. and P.D.C. collected the data; M.P. and M.F.F. carried out the genetic analyses; M.P. analysed the data statistically and led the writing.

Editor: Pauline Ladiges