The LINNEAN SOCIETY of London BOTANICAL Journal of the Linnean Society

Botanical Journal of the Linnean Society, 2015, **179**, 487–498. With figures

Cytotype distribution and phylogeography of *Hieracium intybaceum* (Asteraceae)

JAROSLAV ZAHRADNÍČEK[1,2]* and JINDŘICH CHRTEK[1,2]

[1]*Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, 128 01 Prague, Czech Republic*
[2]*Institute of Botany, Academy of Sciences of the Czech Republic, Zámek 1, 252 43 Průhonice, Czech Republic*

Received 20 September 2013; revised 30 July 2015; accepted for publication 8 August 2015

Using flow cytometry and amplified fragment length polymorphism (AFLP), we explored the cytogeography and phylogeography of *Hieracium intybaceum*, a silicicolous species distributed in the Alps and spatially isolated in the Vosges Mountains and the Schwarzwald Mountains. We detected two ploidies, diploid and tetraploid, but no triploid or mixed-ploidy populations. Whereas diploids are sexual and distributed all across the Alps, tetraploids are apomictic and seem to be confined to the western Alps and the Vosges. We detected a low level of genetic variation. Bayesian clustering identified four clusters/genetic groups, which are partly congruent with the ploidal pattern. The first two groups consisting exclusively of diploids dominate the whole distribution range in the Alps and show east–west geographical separation with a diffuse borderline running from eastern Switzerland to the eastern part of North Tyrol. The third genetic group lacks a defined geographical range and includes diploid and tetraploid plants. The last genetic group comprises tetraploid plants in the French Alps and the Vosges. We suppose that diploids colonized the deglaciated areas from source populations most likely located mainly in the southern part of the recent distribution range and occasionally also in the western Alps. Gene flow and further differentiation likely took place. Apomictic tetraploids most likely originated in the western Alps or in the refugium at the south-western foot of the Alps. Their rather limited geographical range (partly contrasting with the theory of geographical parthenogenesis) can be explained by their rather recent origin. © 2015 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2015, **179**, 487–498.

ADDITIONAL KEYWORDS: AFLP – Alps – geographical parthenogenesis – glacial refugia – polyploidy.

INTRODUCTION

Quaternary climatic changes have fundamentally influenced the present-day distribution and genetic variation of high-mountain species, which had to endure induced migrations and repeated range fragmentation, undergo diversification in isolated refugial areas and recolonize previously glaciated areas, or face local extinction (Hewitt, 1996; Bennett, 1997; Comes & Kadereit, 1998; Tribsch & Schönswetter, 2003).

The Alps and surrounding areas are a suitable model system for studies focused on evolutionary processes influenced by glacial cycles and on locating glacial refugia. Numerous phylogeographic studies carried out in the Alps over the last decades, coupled with previously acquired biogeographical, geological and palaeoenvironmental data, have allowed some more general trends and patterns to be postulated. They have shown that alleles and species sharing a common history of glaciation faced more or less the same distinctive alpine topography during their range retractions and expansions, resulting in two principal breaking zones. One zone is in the western Alps, in the area of the valleys of Aosta and Valais, and the other at the transition between the western and eastern part of the Alps between Lake Garda and Innsbruck (e.g. Thiel-Egenter *et al.*, 2011). Species have also undergone parallel processes in peripheral refugia on both siliceous (mainly in the eastern Alps and along the south and west border of the

*Corresponding author. E-mail: j.zahradnicek@gmail.com

range) and calcareous bedrocks (mainly in the southern and eastern Alps; Brockmann-Jerosch & Brockmann-Jerosch, 1926; Merxmüller, 1952; Stehlik, 2000; Tribsch & Schönswetter, 2003; Schönswetter *et al*., 2005; Schönswetter & Tribsch, 2005; Burnier *et al*., 2009; Slovák *et al*., 2012) or in combined peripheral refugia and nunataks, albeit potentially at different time periods (Schneeweiss & Schönswetter, 2011; García *et al*., 2012). Pleistocene climatic changes also most likely promoted polyploid speciation (Stebbins, 1984, 1985; Brochmann *et al*., 2004; Dixon *et al*., 2009; Casazza *et al*., 2012).

Glacial advancement fragmented once continuous diploid populations that then differentiated in isolation. Partially differentiated populations may have come into contact again during glacial retreats, occasionally giving rise to hybrids that were then stabilized by polyploidization (secondary contact model, Stebbins, 1984, 1985). Alternatively, polyploids may have originated in diploid populations via unreduced gametes (primary contact model, triploid bridge). Formerly glaciated areas were then recolonized by polyploids, which today are dominant in central parts of the Alps (e.g. Stebbins, 1950; Parisod & Besnard, 2007; Schönswetter, 2008). Enhanced frequency of polyploids at higher elevations and latitudes can also be explained by the fact that cold temperatures can trigger the formation of unreduced gametes (Ramsey & Schemske, 1998).

Polyploidization and hybridization influenced by range fluctuations of species in the Pleistocene have also been thought to promote genetic and epigenetic changes that caused shifts to gametophytic apomixis (e.g. Carman, 1997; Grimanelli *et al*., 2001; Hörandl, 2009a, b). Apomictic taxa have generally larger geographical ranges that are often shifted to higher elevations and altitudes (often in formerly glaciated areas) compared to related sexuals (geographical parthenogenesis; Vandel, 1928; Hörandl, 2008; Cosendai & Hörandl, 2010; Hörandl, 2011). This can be explained, among other things, by the advantage of pollinator-independent uniparental reproduction for colonization, shortening of the reproductive pathway and faster seed development in shorter vegetation periods, higher abundance of seeds, more efficient niche exploitation and greater physiological and ecological tolerance due to multiple gene copies (Hörandl, 2006; Hörandl, Cosendai & Temsch, 2008). In the Alps, the highest percentage of apomicts seems to grow in the moderate subalpine and alpine grassland zones (Hörandl, 2011; Hörandl *et al*., 2011). High-elevation plants, in contrast, do not show a pronounced tendency towards polyploidy, so gametophytic apomixis turns out to be rare in the subnival and nival zones (Hörandl *et al*., 2011). Recent studies have supposed occasional re-stabilization of meiotic pathways in some apomictic lineages, followed by reversal to sexuality and allopatric speciation of newly formed sexual populations (expanded transition theory; Hörandl & Hojsgaard, 2012; Hojsgaard *et al*., 2014).

We selected *Hieracium intybaceum* L., a perennial herb scattered to locally common on siliceous bedrock in the subalpine and alpine belts of the Alps, and spatially isolated in the Vosges and the Schwarzwald Mountains, as our model species to examine the pattern of genetic variation in relation to karyological differentiation. Three ploidies have been reported so far, namely diploids ($2n = 18$), triploids ($2n = 27$) and tetraploids ($2n = 36$; Favarger, 1997; Chrtek *et al*., 2007). The mode of reproduction has not hitherto been examined. *Hieracium* polyploids have been proved to be exclusively agamospermous; conversely, *H. intybaceum* is a sister group to the rest of subtribe Hieraciinae (and thus could be placed in the separate genus *Schlagintweitia* Griseb.), but differences from this general pattern cannot be excluded. *Hieracium intybaceum* is an entomophilous species. Its achenes are relatively large and well adapted for dispersal over long distances.

In this study, we addressed the following questions:

(1) Is there any geographical pattern of genetic variation in *Hieracium intybaceum*?
(2) If there is a pattern, does it reflect cytotype and breeding system differentiation and morphological variation?
(3) Does this pattern agree with the concept of geographical parthenogenesis?
(4) What is the mode of reproduction of diploid and tetraploid cytotypes, and are tetraploids agamospermous as in other *Hieracium* spp. or sexual?

MATERIAL AND METHODS

Sampling

Forty-three populations of *Hieracium intybaceum* were collected throughout the Alps and in the Vosges (Table 1) in 2009–2011. Leaf material was taken from at least ten individuals per population and immediately stored in silica gel for further molecular analyses. Five plants from the same sets of individuals were transferred to the experimental garden of the Institute of Botany AS CR in Průhonice (49°59′41″N, 14°34′00″E). Herbarium specimens (a further five plants from the same sets) are deposited in the herbarium of the Faculty of Science, Charles University in Prague (vouchers are without identification number). Small pieces of fresh leaves of 50 plants per population (except for small populations BER, BAL, COL and HOH) were used for cytotype screening using flow cytometry.

Table 1. Details of 43 populations of *H. intybaceum* studied; * – population not used in molecular analyses; DW – rarity index (frequency down-weighted marker values)

Number of locality	Code	Locality name	Latitude	Longitude	Elevation	No. of plants	Ploidy	DW
Int0210	STU	Austria, Salzburg	47.08135	12.492167	1882	4	2*x*	0.588
Int0309	SAB	Austria, Steiermark	47.08135	14.577017	2060	3	2*x*	0.901
Int0310	KON	Austria, Salzburg	47.258861	12.105417	2014	3	2*x*	0.704
Int0410	JAU	Italy, Südtirol	46.837667	11.321028	2186	3	2*x*	1.086
Int0510	SEM	Italy, Südtirol	46.681472	11.712389	1939	3	2*x*	0.815
Int1009	TUR	Austria, Steiermark	46.914333	13.854833	2186	3	2*x*	0.671
Int1010	MOL	Italy, Lombardy	46.542389	10.422889	2477	3	2*x*	0.795
Int1109	OSS	Austria, Kärnten	46.693567	13.921167	1820	5	2*x*	0.769
Int1110	STE	Italy, Lombardy	46.53275	10.477194	2224	4	2*x*	1.030
Int1309	REI	Austria, Kärnten	46.904933	13.338967	2243	4	2*x*	0.693
Int1310	LAG	Switzerland, Ticino	46.47875	8.589111	2214	4	2*x*	0.680
Int1609	STR	Austria, Kärnten	46.5967	13.118367	1543	4	2*x*	0.579
Int1610	FIE	Switzerland, Ticino	46.536417	8.567833	1968	5	2*x*	0.724
Int1709	LAM	Austria, Kärnten	46.631783	12.935483	1598	5	2*x*	0.788
Int1809	EGG	Austria, Ost-tirol	46.880833	12.776233	2114	5	2*x*	0.725
Int1810	AND	Switzerland, Uri	46.656444	8.642611	2013	4	2*x*	1.048
Int1909	STA	Italy, Trento	46.889983	12.19655	1949	6	2*x*	0.978
Int1910	GLA	Switzerland, Uri	46.572694	8.395444	2128	2	2*x*	0.645
Int2009	PEN	Italy, Trento	46.818483	11.434633	2246	5	2*x*	0.864
Int2010	BRU	Switzerland, Valais	46.374694	7.977111	2125	4	2*x*	0.722
Int2109	VER	Italy, Trento	46.732817	10.815667	1761	5	2*x*	1.068
Int2209	ROM	Italy, Trento	46.885433	11.118467	1936	5	2*x*	0.875
Int2210	JAM	Austria, Tirol	46.926667	10.176944	1805	4	2*x*	0.871
Int2309	VEN	Austria, Tirol	46.8785	10.932183	2043	3	2*x*	0.718
Int2310	LUC	Austria, Tirol	47.019778	12.694444	2070	5	2*x*	0.699
Int2410	PAN	Austria, Tirol	46.998333	12.598333	2205	6	2*x*	0.842
Int2509	PEE	Austria, Tirol	47.156117	11.5634	1858	5	2*x*	0.754
Int2510	BER	Italy, Valle de Aosta	45.680083	6.885333	2219	6	2*x*	1.071
Int2609	GUR	Austria, Tirol	46.860528	11.024278	2065	4	2*x*	0.749
Int2610	COI	France, Savoy	45.66275	6.870472	2082	5	2*x*	0.673
Int2709	RET	Austria, Tirol	46.470194	11.024278	2035	4	2*x*	0.619
Int2710	ROS	France, Savoy	45.685528	6.696139	2145	4	2*x*	0.604
Int2809	MOR	Switzerland, Graubünden	46.470194	9.757472	2145	5	2*x*	0.890
Int2909	FLU	Switzerland, Graubünden	46.744361	9.982389	2155	4	2*x*	0.929
Int2910	CEZ	France, Hautes-Alpes	44.925222	6.409611	2112	7	4*x*	0.764
Int3410	BRO	France, Isére	45.03525	5.888972	2126	6	4*x*	0.940
Int3810	BAL	France, Haute-Savoie	46.025333	6.9585	2047	5	4*x*	0.902
Int38102	COL	France, Isére	45.387669	6.135403	1650	5	2*x*	0.708
Int4010	HOH	France, Vosges	48.041722	7.013111	1279	7	4*x*	0.856
Int4111	THO	France, Savoie	45.2893	6.588733	2475	2	2*x*	0.839
Int4911	CAS	France, Hautes-Alpes	44.98085	6.4352	2164	4	4*x*	0.976
Int3009*	PYR	Austria, Tirol	47.131333	10.209972	1787		2*x*	
Int2911*	TER	France, Haute-Savoie	45.949333	6.850633	2203		2*x*	

Ploidy and mode of reproduction

DNA ploidy (Suda *et al*., 2006) was determined by flow cytometry using a Partec PA II device equipped with an HBO mercury arc lamp. Sample preparation followed the two-step procedure using Otto's buffers I and II (Doležel, Greilhuber & Suda, 2007). Pooled samples for faster estimation were prepared from intact leaf tissue. Each sample, with an appropriate quantity of the internal reference standard *Bellis perenis* L. (2C = 3.37 pg, Schönswetter *et al*., 2007), was chopped with a new sharp razor blade in a Petri dish containing 1 mL ice-cold Otto I buffer (0.1 M citric acid, 0.5% Tween 20) as the nuclear isolating

solution. The suspension was filtered through a nylon mesh (loop size 42 μm), incubated for 10 min at room temperature and mixed with 1 mL Otto II buffer (0.4 M Na_2HPO_4) supplemented with AT-selective DAPI (4′,6-diamidino-2-phenylindole) as the fluorochrome (at the concentration of 4 mg mL^{-1}), RNase IIA (50 mg mL^{-1}) and β-mercaptoethanol (2 mg mL^{-1}). Samples were stained for 5 min at room temperature prior to cytometry. Three thousand nuclei were analysed for each sample. Only histograms with coefficients of variation (CVs) for the G0/G1 < 5% were accepted.

The mode of reproduction was tested using the plants transferred to the experimental garden following Gadella (1987) and Krahulcová & Krahulec (1999). In diploids (in which sexual reproduction was expected), capitula were bagged and pollinated by hand (test for autogamy); results were compared with control capitula from the same plant in open pollination treatment. In tetraploids (in which agamospermy was expected), the upper part of the capitulum was cut off at the bud stage, capitula were bagged to prevent lost of achenes by wind and full (developed) achenes were counted.

AFLP data acquisition and scoring

Total genomic DNA was extracted using the Invisorb Spin Plant Mini Kit (Invitek) following the manufacturer's instructions using 0.5 g of silica-dried leaf tissue. The concentration of DNA in the first elution was measured using a NanoDrop spectrophotometer, and the DNA was diluted to the initial concentration of 50 ng μL^{-1}. One hundred and eighty-two individuals (two to seven per population, Table 1) were analysed for AFLPs using the AFLP Core Reagent Kit I (Invitrogen) and AFLP Pre-Amp Primer Mix I (Invitrogen). The whole procedure (restriction, ligation, pre-amplification and selective amplification) followed Rejzková *et al*. (2008) with the following modifications: ca. 130 ng of genomic DNA were digested for 5 h and ligation was extended to 12 h. Pre-amplification and selective amplification was done using JumpStart RedTag Polymerase (Sigma), and the same volume of 10× buffer was used with JumpStart RedTag Polymerase (Sigma). The reaction conditions were an initial step of 2 min at 94 °C and 2 min at 72 °C followed by 20 cycles of 30 s at 94 °C, 30 s at 56 °C, 2 min at 72 °C with final extension at 60 °C for 30 min. A 10 × diluted (in ddH_2O) product of pre-amplification was used for the final, selective amplification.

Based on preliminary tests, three primer combinations were used for selective amplification: *Eco*RI-AAC + *Mse*I-CTA, *Eco*RI-ACG + *Mse*I-CAC and *Eco*RI-ACA + *Mse*I-CTG (Applied Biosystems). The reaction conditions were an initial step of 2 min at 94 °C, 30 s at 65 °C and 2 min at 72 °C followed by eight cycles of 30 s at 94 °C, 30 s at 64 °C (reduced by 1 °C per cycle) and 2 min at 72 °C, which were followed by 23 cycles of 30 s at 94 °C, 30 s at 56 °C and 2 min at 72 °C, then the final extension of 30 min at 60 °C and a hold period at 10 °C. For each sample, 1 μL of each 6-FAM, HEX and NED-labelled products of selective amplification was purified and precipitated using ethanol/sodium acetate precipitation. The precipitate was resuspended in 0.25 μL GeneScan-ROX-500 size standard in 10 μL formamide and denatured for 3 min at 95 °C. Fragments were analysed on a 3100 Avant Genetic Analyzer (Laboratory of Sequencing, Faculty of Science, Charles University in Prague) and scored with GeneMarker v 1.8 (www.SoftGenetics.com). Twelve per cent of all samples were re-analysed from the first step of AFLP procedure (restriction) under normal conditions to test the reproducibility of the method and estimate the average error rate (Bonin *et al*., 2004). Fragments between 75 and 500 bp long were scored.

AFLP data analysis

Frequency down-weighted marker values (DW; Schönswetter & Tribsch, 2005) per population were calculated using the AFLPdat script (Ehrich, 2006) in R. Higher values are thought to indicate populations that have been isolated for a longer time and low values populations that diverged rather recently. The same script was used to create input files for Structure and Arlequin.

Partitioning of genetic variation within and among populations and cytotypes was tested for by analysis of molecular variance (AMOVA) using Arlequin 3.1 (Excoffier, Laval & Schneider, 2005). We performed two AMOVA analyses using two (within and among populations) and three (within and among populations, and between cytotypes) hierarchical levels. Population pairwise PhiPT (an F_{ST} analogue) values were computed and Mantel test (pairwise PhiPT values and geographical distances) was performed in GenAlEx 6.5 (Peakall & Smouse, 2012).

The genetic structure was evaluated using three different approaches:

(1) The first was a non-model-based approach; non-hierarchical K-means clustering (Hartigan & Wong, 1979) was chosen because of the presence of two ploidies using a script by Arrigo *et al*. (2010) in R. This approach assigns individuals to a defined number of genetic groups in order to maximize intergroup variance. This technique was successfully applied in the analysis of genetic structure of the AFLP datasets in polyploid complexes (Burnier *et al*., 2009; Arrigo *et al*., 2010).

(2) The second approach taken used model-based Bayesian clustering implemented in Structure 2.2 (Pritchard, Stephens & Donnelly, 2000; Falush, Stephens & Pritchard, 2007). This approach is less appropriate for detecting admixture between cytotypes because it is based on the ideal Hardy–Weinberg population model minimizing linkage disequilibrium.
(3) Principal coordinate analysis (PCoA, Jaccard's similarity coefficients) performed using PAST (Hammer, Harper & Ryan, 2001) was the last approach. K-means clustering was conducted in 100 000 independent runs (i.e. starting from random points) for each assumed value of K clusters (ranging from 1 to 10), and the intergroup inertia of each run was recorded.

We followed the procedure described by Evanno, Regnaut & Goudet (2005) to select the most likely number of groups using intergroup inertia as a proxy of clustering accuracy. In the Structure analysis, the number of clusters was estimated using 10^6 iterations with a burn-in period of 10^5 iterations under an admixture model. Ten replicates for each K were analysed from K = 1 to K = 10; the outputs were processed following the approach of Evanno *et al*. (2005) implemented in Structure Harvester to determine the most likely groups of clusters (Earl & Vonholdt, 2012). The mixed-ploidy dataset was encoded into the 'polyploid' format used by Stock *et al*. (2010). For the purpose of the Structure analysis (a complete dataset), all individuals were coded as having four chromosome sets (corresponding to the highest ploidy in the dataset, i.e. 4*x*). The two chromosome sets not present in diploid individuals were marked as missing data (–9). This model was compared with the usual coding for diploid plants. Both these procedures were also applied to subsets of diploids and tetraploids (triploids were not found). The nonhierarchical K-means clustering analysis and Bayesian clustering analysis were performed for the mixed ploidy dataset ('complete dataset analysis') and for two ploidy subsets of diploids and tetraploids separately ('nested analysis').

RESULTS

Cytotype screening

We ascertained the ploidy of 2270 plants from 43 populations. Two ploidies, diploid and tetraploid, were detected. Whereas diploids are distributed all across the Alps (38 populations), tetraploids seem to be confined to the western Alps (four populations analysed come from the Savoy Alps) and the Vosges (one population). We found no triploids and no mixed-ploidy populations.

Mode of reproduction

Only empty achenes were found in bagged capitula of diploid plants (eight plants, 12 capitula). Untreated (open pollination) capitula, in contrast, produced nearly 100% of full achenes, showing allogamy in sexual diploids. A high proportion of full achenes (range 86.9–92.0%, mean 89.3%) confirming agamospermy was found in tetraploid plants with emasculated capitula (five plants, one capitulum per plant).

DNA analysis

Using two of the three AFLP primer combinations (*Eco*RI-ACG + *Mse*I-CAC and *Eco*RI-ACA + *Mse*I-CTG) together generated 152 clear polymorphic fragments. The third primer combination (*Eco*RI-AAC + *Mse*I-CTA) was not included in the final analysis because it generally yielded low-quality results. Based on 22 replicates (12% of samples), the reproducibility of the dataset was 94.5%.

The rarity index (DW) varied from 0.58 in population STR (southern Austria) to 1.068 in population JAU (northern Italy; Table 1). Populations with higher DW indices are located mainly in the southern part of the geographical range (Fig. 1). The DW values of diploid (mean 0.796) and tetraploid (mean 0.888) populations are not significantly different (two-tailed *t*-test, d.f. = 38, t = 1.369, $P = 0.178$).

The AFLP analysis revealed low variability. We found no private markers for the individual populations. After dividing the whole dataset into two groups based on ploidy, diploid populations had eight private markers and tetraploid populations only one private marker. In the AMOVA with only two levels of variation (i.e. within and among populations), 27% of the overall genetic variation was found among populations. The population pairwise PhiPT values showed the greatest differentiation between ROS and BER (PhiPT = 0.667), CEZ and ROS (0.637) and CEZ and BER (0.634) with mean PhiPT = 0.258. In our nested AMOVA analyses, the variation between cytotypes accounted for 20% of the overall variation. In the subset of diploid populations, 25% of variation was detected among populations; the highest differentiation was found between ROS and BER (PhiPT = 0.667) with mean PhiPT = 0.246. Lower inter-population differentiation was found in the subset of tetraploid populations (12% of the variation can be explained by inter-population variation); the highest differentiation was found between CAS and CEZ (PhiPT = 0.280) with mean PhiPT = 0.123.

In the among-population AMOVA of the diploid sub-dataset, populations ROS from south-eastern France and the easternmost population SAB were found to be most distinct, differing by 27% and 15% from the nearest population, respectively. In the

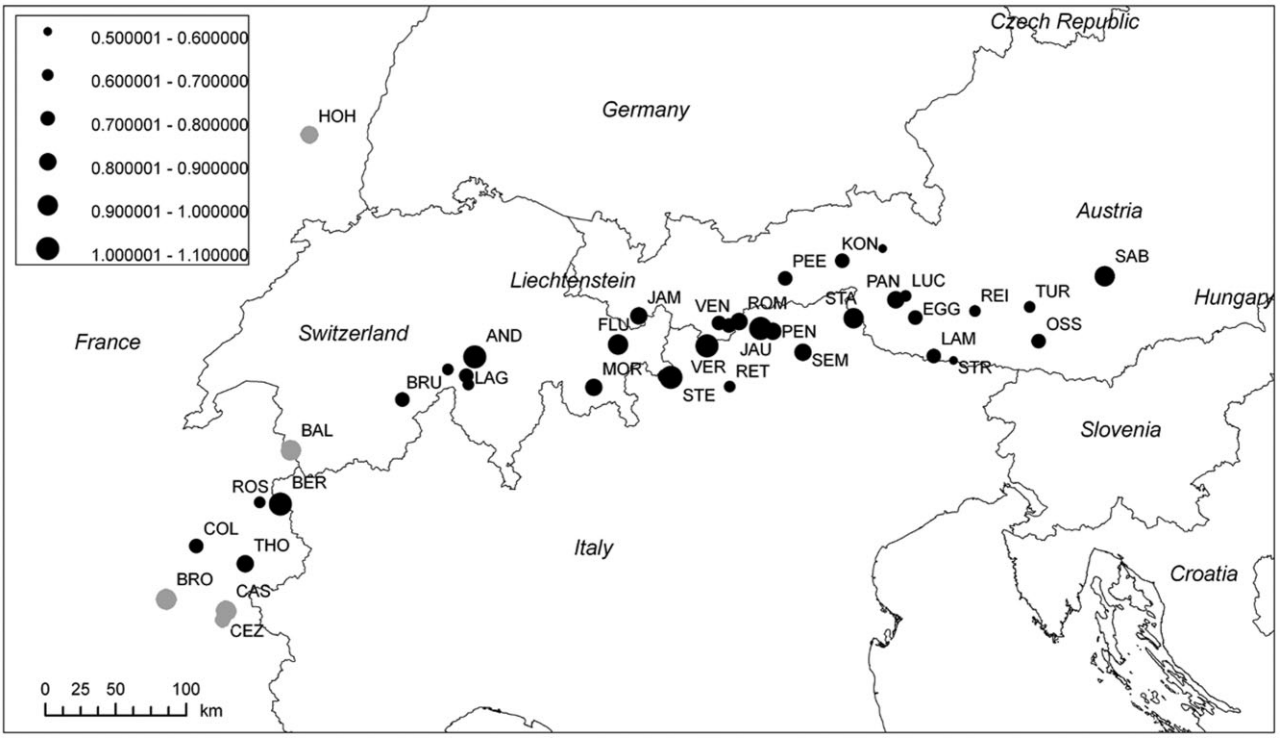

Figure 1. Cytotype distribution and rarity index of 41 populations of *H. intybaceum*. The two symbols present ploidy groups: circles – diploid populations, grey circles – tetraploid populations. The size of each symbol is proportional to the value of the rarity index.

tetraploid sub-dataset, population BAL from south-eastern France had the greatest genetic distance of 10% from the rest of tetraploid populations.

The Mantel test computed for the whole dataset revealed a low but significant correlation between genetic and geographical distances ($r_m = 0.246$, $P = 0.001$), indicating an isolation-by-distance pattern. A positive correlation was found also for diploids ($r_m = 0.193$, $P = 0.010$).

Nonhierarchical K-means clustering resulted in an optimal grouping of the complete dataset into two or four groups (Supporting Information Fig. S1). However, four groups reflect the observed cytological and geographical differentiation much better (Fig. 2A). Individuals assigned to the first genetic cluster K1 were found in diploid populations of central and eastern Alps. The individuals belonging to the genetic cluster K2 were detected in diploid populations throughout the entire distributional range of the species. Individuals falling in genetic cluster K3 were detected mainly in diploid and rarely in tetraploid populations encompassing the entire distributional range. The remnant tetraploid individuals fallen in the genetic cluster K4. Populations were mainly constituted of individuals belonging to one or two genetic clusters; rarely they shared three clusters. The K-means analysis was applied separately to the diploid and tetraploid datasets (nested analysis). It resulted in an optimal grouping into three groups for diploids and two groups for tetraploids, the latter without any geographical pattern. The classification of most individuals (except for two plants from SEM and MOL) by the nested analysis is congruent with the analysis of the complete dataset.

The Bayesian analysis of the complete dataset using Structure resulted in the best partitioning into either two (similarity coefficient 0.90) or five (similarity coefficient 0.98) groups (Supporting Information Fig. S2). In the latter case, tetraploids fall into two groups (the first one exclusively tetraploid, the second one with a small admixture of diploids) and the remaining three groups are formed exclusively by diploid plants. The first two diploid groups show a geographical pattern (east, west), whereas the third one comprises plants across the whole geographical range without any pattern. This contradicts the number of clusters suggested by the K-means clustering (four). Nested Bayesian analysis resulted in three groups of diploids and two groups of tetraploids, the latter groups without any geographical pattern, which is in agreement with results of nested K-means analysis. Classifications of plants in Bayesian

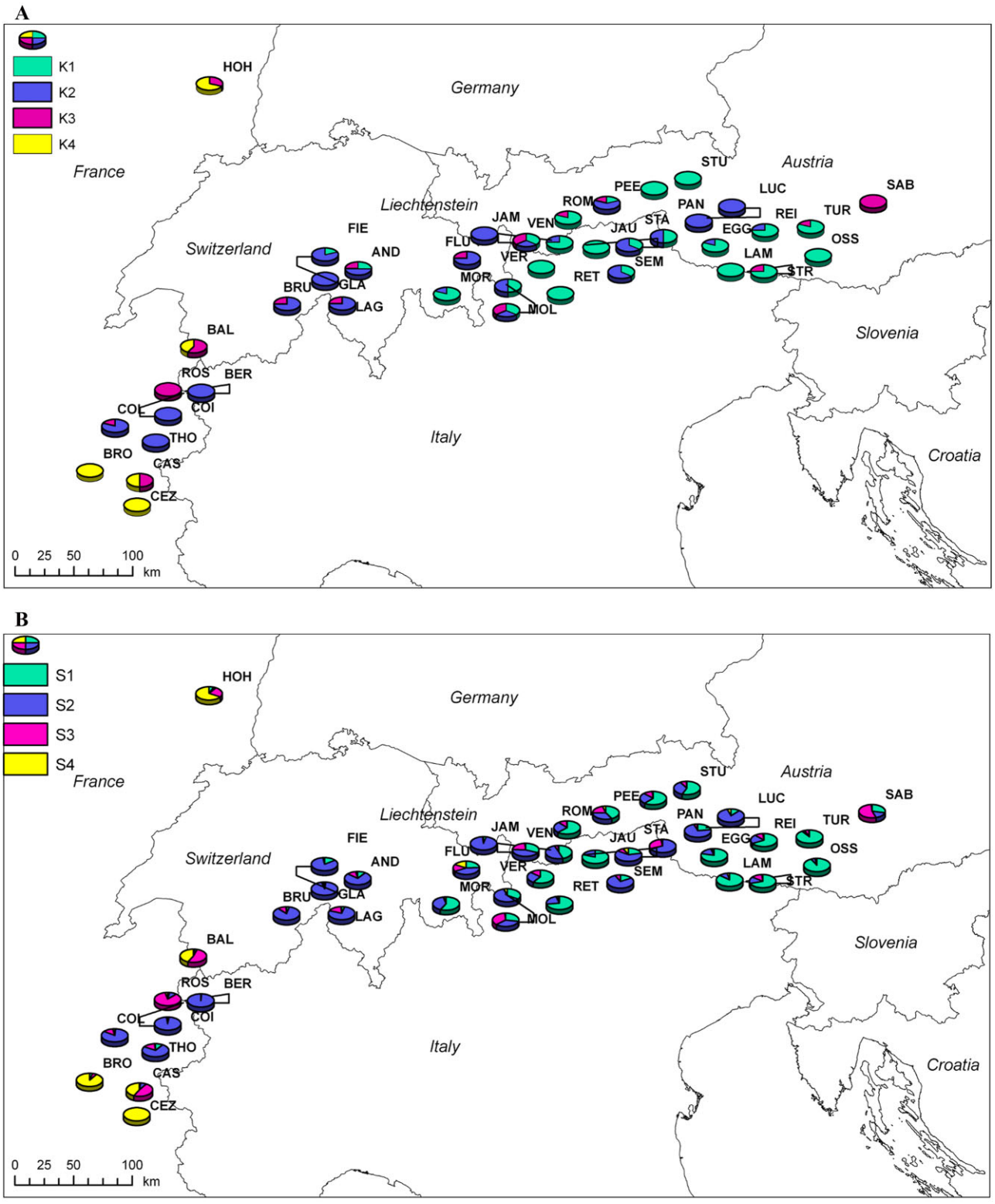

Figure 2. Phylogeographic pattern of 41 populations of *H. intybaceum* analysed. The groups are based on (A) K-means clustering; and (B) Bayesian clustering. The graphs show the proportion of individuals classified into clusters K1–K4 in the case of K-means clustering and the mean proportional presence of clusters S1–S4 in each population in the case of Bayesian clustering.

Table 2. Contingency table of numbers of individuals comparing the grouping based on Bayesian clustering (S1–S4) and K-means clustering (K1–K4)

	K1	K2	K3	K4
S1	59	2		
S2	8	63	2	
S3			23	
S4		1		22

analysis of the complete dataset (five groups) on one hand and separated sets of diploids and tetraploids (nested analysis) on the other differ considerably. Partitioning of the complete dataset into four groups, however, shows nearly the same classification of plants as the nested analyses. With these facts in mind, we decided to consider Structure clusters S1–S4 and discuss this classification further. The Structure results are (except for 13 individuals) congruent with the K-means analysis of the complete dataset (Table 2). Discrepancies between results of K-means and Bayesian clustering can be explained by the clustering method adopted; the latter takes into account possible admixture, i.e. it allows for the possibility that individuals may have mixed ancestry.

Principal coordinate analysis (PCoA, Fig. 3) confirmed the rather weak structuring of the dataset detected by the previous analyses. Similarly to K-means clustering, it divided the dataset into four groups. The first three axes explained 18.6% of the total variation. The first axis (explaining 8.7% of the total variation) separated diploid clusters (K1, K2) from tetraploids (K4). The second axis (explaining 5.5% of the total variation) separated the 'eastern diploid group' (K1) from the 'western diploid group' (K2). The third axis (explaining 5.5% of the total variation) separated the cluster K3 from the rest of the groups. Structure genetic groups (clusters S1–S4) are largely congruent with K-means clusters K1–K4 and thus also match the results of PCoA.

DISCUSSION

Cytogeography of *Hieracium intybaceum*

The present data corroborate and add details to the geographical pattern proposed by Favarger (1997), who reported tetraploids from south-western Switzerland and diploids from Switzerland, Austria and Italy. Sexual diploids occur throughout the distribution area in the Alps, whereas apomictic tetraploids are confined to the western Alps (reaching eastwards to the canton of Wallis in south-western Switzerland) and the Vosges.

We did not confirm the occurrence of triploids. Published triploid counts for *H. intybaceum* (Rosenberg, 1927; Larsen, 1954) seem doubtful even though triploid plants can generally appear at low frequencies in diploid populations. The discrepancy between our results and previous reports of triploids can be explained by misidentifications with *H. pallidiflorum* Hausm. This triploid species is morphologically similar to *H. intybaceum* and often gets confused with it in herbarium material. Alternatively and less likely, triploids of *H. intybaceum* may not have been detected due low frequencies in populations. Moreover, the geographical distribution of diploid and tetraploid populations of *H. intybaceum* does not correspond with the model of geographical parthenogenesis (van Dijk, 2003; Hörandl, 2008; Cosendai & Hörandl, 2010), a pattern in which apomictic polyploids have larger distribution ranges shifted to higher latitudes and elevations compared to related sexual diploids. Diploids of *H. intybaceum* are widely distributed across the entire range of the species (most likely except in the Vosges), mostly in previously strongly glaciated areas. Tetraploid apomictic populations, in contrast, are confined to a rather small geographical area. This pattern also differs from the trends found in other *Hieracium* spp. with ploidal differentiation in the Alps, e.g. in *H. prenanthoides* Vill. and *H. tomentosum* L. (diploids at the SW edge of the Alps; Chrtek *et al*., 2007, 2009; J. Chrtek and P. Mráz, unpubl.). We did not detect mixed-ploidy populations.

Genetic and ploidal pattern

We detected a rather low level of genetic variation in *H. intybaceum*. The relatively low contribution of the among-population component and low genetic distances (Fig. 3) can indicate gene flow (post-glacial genetic erosion) and/or rather recent divergence, especially in clusters S1 and S2. We preferred the result obtained by Bayesian clustering which identified four clusters, three of them more or less corresponding with the observed ploidal differentiation. The first two genetic groups (clusters S1 and S2) dominate across the whole distribution range and show east-west geographical separation (Fig. 2B).

The observed genetic and ploidal pattern can be explained as follows. Diploids colonized the deglaciated areas from source populations most likely located mainly in the southern part of the distribution range, occasionally also in the western Alps (see the highest DW values). Partially differentiated populations/lineages came into contact and consequently most likely further differentiation occurred. Apomictic tetraploids most likely originated in the western Alps or in the refugium at the south-western foot of the Alps. Their rather limited recent geographical range (partly

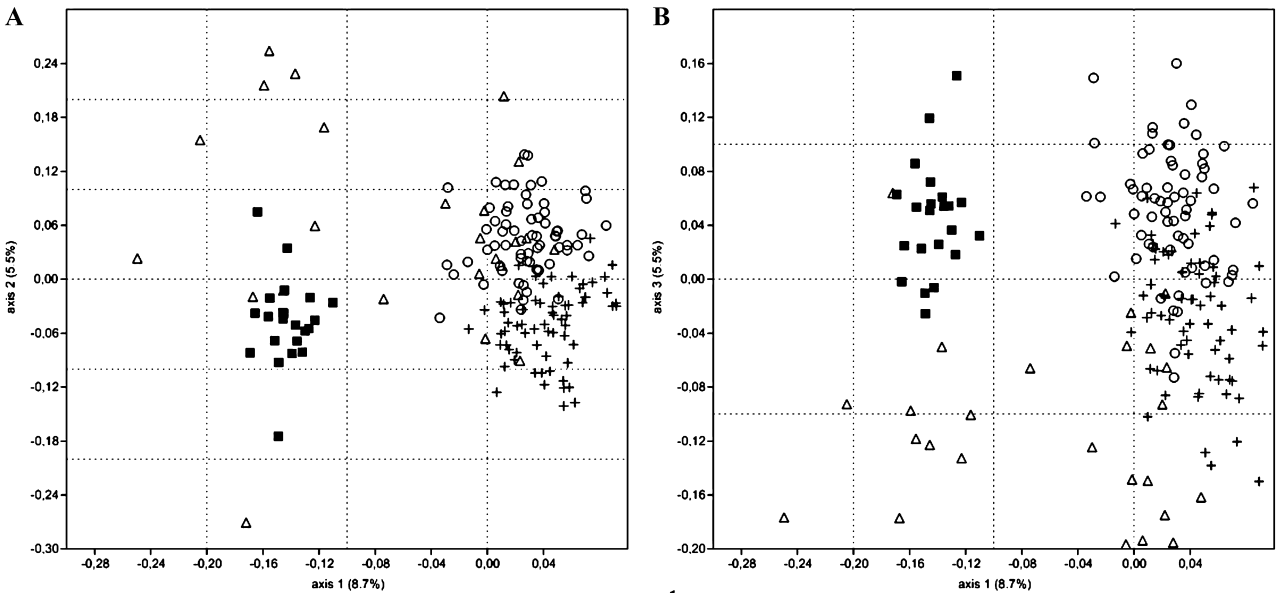

Figure 3. Principal coordinate analysis based on Jaccard similarity among AFLP phenotypes of *Hieracium intybaceum*. (A) axis 1 and axis 2; (B) axis 1 and axis 3. The four symbols represent groups produced by K-means analysis, which are congruent with those produced by Bayesian clustering; crosses – K1, white circles – K2, white triangles – K3, black squares – K4.

contrasting with the theory of geographical parthenogenesis) can be explained in two ways: (1) they originated rather recently but their occurrence coupled with a most likely complete absence of diploids in the Vosges should be taken into consideration, a more plausible explanation; or (2) they had/have a lower fitness than diploid plants. However, the latter is in a contradiction with putative general advantage of uniparental reproduction for colonization (plants reproduce independently of foreign pollen and do not suffer from low activity of pollinators in extreme habitats; van Dijk, 2003; Hörandl, 2009a). We did not carry out a detailed comparison of the fitness of diploids and tetraploids; however, seed production in tetraploids, for example, has proved to be at a similarly high level as in other successful apomictic hawkweeds (J. Zahradníček & J. Chrtek, unpubl.). Thus, the latter hypothesis seems to be much less plausible.

The diffuse borderline between clusters S1 and S2 runs from eastern Switzerland to the eastern part of North Tyrol. The geographical pattern in this genetic group is nearly, but not completely, identical to that detected in a set of silicicolous alpine species exhibiting differentiation along the main break zone (the transition between the western and eastern Alps), which runs roughly from Lake Garda to Innsbruck (Brenner zone; Thiel-Egenter *et al*., 2011).

The third genetic group (cluster S3) lacks a defined geographical range; it comprises mainly plants from two diploid populations from the opposite (west and east) margins of the distribution range and a few plants from both diploid and tetraploid populations across the distribution range in the Alps. Principal component analysis separated only one population from the south-western Alps (ROS) and tetraploid individuals along the first and second axis, respectively; other diploid individuals of the third genetic group are distributed in populations assigned to clusters S1 and S2. We therefore regard part of the distribution of genetic group 3 as an artefact of conflicts between approaches and the already mentioned low variability in the dataset. That at least one individual in each population can be assigned to cluster S3 confirms this claim.

The fourth genetic group (cluster S4) is well defined and includes only plants from tetraploid populations in the French Alps and the Vosges, and is genetically closely related to cluster S3 (most closely to population ROS). Geographically, its eastern borderline (Aosta valley and Wallis) is congruent with the main break zone detected in many alpine plants (e.g. Thiel-Egenter *et al*., 2011). The high genetic differentiation between the ROS population (cluster S3) and tetraploids (cluster S4) on the one hand and other diploids (clusters S1, S2) on the other might indicate long-term separation of these groups. We assume that genetic group 3 can be a source group of polyploid populations due to the considerably high presence of tetraploid plants (23%) in this genetic group and the genetic proximity between genetic group 3 and tetraploids.

Although certain break zones in *H. intybaceum* are more or less similar to those identified in other alpine plants, its genetic differentiation is generally lower. Four rather clearly defined groups of populations (SW, W, C and E) were found, for example, in *Ranunculus glacialis* L. (Schönswetter *et al*., 2004) and *Androsace alpina* (L.) Lam. (Schönswetter, Tribsch & Niklfeld, 2003). In both these species, the two western groups (SW and W) were genetically isolated both from each other and from the two eastern groups (C and E). Our genetic group S4 at least slightly corresponds to the SW group, S1 to the W groups and S2 approximately overlaps with the C and E groups. However, we did not detect a strong genetic divergence between the W group on the one hand and C and E groups on the other like in both *Ranunculus glacialis* and *Androsace alpina*.

To sum up, we found two cytotypes of *H. intybaceum*, namely sexual diploids prevailing throughout the Alps, and apomictic tetraploids confined to the western Alps and the Vosges. The majority of tetraploid plants forms one genetic group; diploids can be divided into three groups, but the overall genetic variation is low, most likely indicating gene flow and/or recent differentiation.

ACKNOWLEDGEMENTS

We are grateful to Filip Kolář and Pavel Trávníček for statistical advice, Věroslava Hadincová for preparing the maps and Lenka Flašková for a kind introduction to the techniques of AFLP. For help with the collection of samples, we are indebted to Patrik Mráz, Petra Zahradníčková, Zuzana Dočkalová, Frederick Rooks, Hana Chudáčková and Zita Červenková. Frederick Rooks improved the English. The study was funded by the Grant Agency of Charles University in Prague (project No. 72309/2009) and partly also by the Czech Science Foundation (grant No. P506/10/1363) and the long-term research development project RVO 67985939.

REFERENCES

Arrigo N, Felber F, Parisod C, Buerki S, Alvarez N, David J, Guadagnuolo R. 2010. Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat. *New Phytologist* **187:** 1170–1180.

Bennett KD. 1997. *Evolution and ecology. The pace of life*. Cambridge: Cambridge University Press.

Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P. 2004. How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13:** 3261–3273.

Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen A-C, Even R. 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society* **82:** 521–536.

Brockmann-Jerosch H, Brockmann-Jerosch M. 1926. Die Geschichte der Schweizerischen Alpenflora. In: Schröter C, ed. *Das Pflanzenleben der Alpen*. Zurich: Raustein, 1110–1215.

Burnier J, Buerki S, Arrigo N, Kupfer P, Alvarez N. 2009. Genetic structure and evolution of alpine polyploid complexes: *Ranunculus kuepferi* (Ranunculaceae) as a case study. *Molecular Ecology* **18:** 3730–3744.

Carman JG. 1997. Asynchronous expression of duplicate genes in angiosperms can cause apomixis, bispory, tetraspory, and polyembryony. *Biological Journal of the Linnean Society* **61:** 51–94.

Casazza G, Granato L, Minuto L, Conti E. 2012. Polyploid evolution and Pleistocene glacial cycles: a case study from the alpine primrose *Primula marginata* (Primulaceae). *BMC Evolutionary Biology* **12:** 56.

Chrtek J Jr, Zahradníček J, Krak K, Fehrer J. 2009. Genome size in *Hieracium* subgenus *Hieracium* (Asteraceae) is strongly correlated with major phylogenetic groups. *Annals of Botany* **104:** 161–178.

Chrtek J, Mráz P, Zahradníček J, Mateo G, Szeląg Z. 2007. Chromosome numbers and DNA ploidy levels of selected species of *Hieracium s.str*. (Asteraceae). *Folia Geobotanica* **42:** 411–430.

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

Cosendai AC, Hörandl E. 2010. Cytotype stability, facultative apomixis and geographic parthenogenesis in *Ranunculus kuepferi* (Ranunculaceae). *Annals of Botany* **105:** 457–470.

van Dijk PJ. 2003. Ecological and evolutionary opportunities of apomixis: insights from *Taraxacum* and *Chondrilla*. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **358:** 1113–1120.

Dixon CJ, Schönswetter P, Suda J, Wiedermann MM, Schneeweiss GM. 2009. Reciprocal Pleistocene origin and postglacial range formation of an allopolyploid and its sympatric ancestors (*Androsace adfinis* group, Primulaceae). *Molecular Phylogenetics and Evolution* **50:** 74–83.

Doležel J, Greilhuber J, Suda J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2:** 2233–2244.

Earl DA, Vonholdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:** 359–361.

Ehrich D. 2006. AFLPDAT: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes* **6:** 603–604.

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

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **1:** 47–50.

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

Favarger C. 1997. Notes de caryologie alpienne VI. *Bulletin de la Société Neuchâteloise, Sciences Naturelles* **120:** 19–33.

Gadella TWJ. 1987. Sexual tetraploid and apomictic pentaploid populations of *Hieracium pilosella* (Compositae). *Plant Systematics and Evolution* **157:** 219–246.

García PE, Winkler M, Flatscher R, Sonnleitner M, Krejčíková J, Suda J, Hülber K, Schneeweiss GM, Schönswetter P. 2012. Extensive range persistence in peripheral and interior refugia characterizes Pleistocene range dynamics in a widespread alpine plant species (*Senecio carniolicus*, Asteraceae). *Molecular Ecology* **21:** 1255–1270.

Grimanelli D, Leblanc O, Perotti E, Grossniklaus U. 2001. Developmental genetics in gametophytic apomixis. *Trends in Genetics* **17:** 597–604.

Hammer O, Harper DAT, Ryan PD. 2001. PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4:** 1–12.

Hartigan JA, Wong MA. 1979. A K-means clustering algorithm. *Applied Statistics* **28:** 100–108.

Hewitt GM. 1996. Some genetic consequences of ice ages and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58:** 247–276.

Hojsgaard D, Klatt S, Baier R, Carman JG, Hörandl E. 2014. Taxonomy and biogeography of apomixis in angiosperms and associated biodiversity characteristics. *Critical Reviews in Plant Sciences* **33:** 417–427.

Hörandl E. 2006. The complex causality of geographical parthenogenesis. *New Phytologist* **171:** 525–538.

Hörandl E. 2008. The complex causality of geographical parthenogenesis. *New Phytologist* **171:** 525–538.

Hörandl E. 2009a. Geographical parthenogenesis: opportunities for asexuality. In: Schön I, Martens K, Van Dijk PJ, eds. *Lost sex*. Dordrecht: Springer, 161–186.

Hörandl E. 2009b. A combinational theory for maintenance of sex. *Heredity* **103:** 445–457.

Hörandl E. 2011. Evolution and biogeography of alpine apomictic plants. *Taxon* **60:** 390–402.

Hörandl E, Cosendai A-C, Temsch EM. 2008. Understanding the geographic distribution of apomictic plants: a case for a pluralistic approach. *Plant Ecology and Diversity* **1:** 309–320.

Hörandl E, Dobeš C, Suda J, Vít P, Urfus T, Temsch EM, Cosendai A-C, Wagner J, Ladinig U. 2011. Apomixis is not prevalent in subnival to nival plants of the European Alps. *Annals of Botany* **108:** 381–390.

Hörandl E, Hojsgaard D. 2012. The evolution of apomixis in angiosperms. A reappraisal. *Plant Biosystems* **146:** 681–693.

Krahulcová A, Krahulec F. 1999. Chromosome numbers and reproductive systems in selected representatives of *Hieracium* subgen. *Pilosella* in the Krkonoše Mts (the Sudeten Mts). *Preslia* **71:** 217–234.

Larsen K. 1954. Chromosome numbers of some European flowering plants. *Botanisk Tidsskrift* **50:** 163–174.

Merxmüller H. 1952. Untersuchungen zur Sippengliederung und Arealbildung in den Alpen, I. *Jahrbuch des Vereins zum Schutze der Alpenpflanzen und Tiere* **17:** 96–133.

Parisod C, Besnard G. 2007. Glacial *in situ* survival in the western Alps and polytopic autopolyploidy in *Biscutella laevigata* L. (Brassicaceae). *Molecular Ecology* **16:** 2755–2767.

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

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

Ramsey J, Schemske DW. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* **29:** 467–501.

Rejzková E, Fér T, Vojta J, Marhold K. 2008. Phylogeography of the forest herb *Carex pilosa* (Cyperaceae). *Botanical Journal of the Linnean Society* **158:** 115–130.

Rosenberg O. 1927. Die semiheterotypische Teilung und ihre Bedeutung für die Entstehung verdoppelter Chromosomenzahlen. *Hereditas* **8:** 305–338.

Schneeweiss GM, Schönswetter P. 2011. A re-appraisal of nunatak survival in arctic-alpine phylogeography. *Molecular Ecology* **20:** 190–192.

Schönswetter P. 2008. Polyploid evolution and ecological differentiation in *Senecio carniolicus* (Asteraceae). Book of abstracts, p. 23. *Xth Symposium of the International Organization of Plant Biosystematics*, 2–4 July 2008, Štrbské Pleso, Slovakia.

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.

Schönswetter P, Suda J, Popp M, Weiss-Schneeweiss H, Brochmann C. 2007. Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. *Molecular Phylogenetics and Evolution* **42:** 92–103.

Schönswetter P, Tribsch A. 2005. Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L. (Apiaceae). *Taxon* **54:** 725–732.

Schönswetter P, Tribsch A, Niklfeld H. 2003. Phylogeography of the high alpine cushion-plant *Androsace alpina* (Primulaceae) in the European Alps. *Plant Biology* **5:** 623–630.

Schönswetter P, Tribsch A, Stehlik I, Niklfeld H. 2004. Glacial history of high alpine *Ranunculus glacialis* (Ranunculaceae) in the European Alps in a comparative phylogeographic context. *Biological Journal of the Linnean Society* **81:** 183–195.

Slovák M, Kučera J, Turis P, Zozomová-Lihová J. 2012. Multiple glacial refugia and postglacial colonization routes inferred for a woodland geophyte, *Cyclamen purpurascens*:

patterns concordant with the Pleistocene history of broad-leaved and coniferous tree species. *Biological Journal of the Linnean Society* **105:** 741–760.

Stebbins GL. 1950. *Variation and evolution in plants*. New York: Columbia University Press.

Stebbins GL. 1984. Polyploidy and the distribution of the Arctic-Alpine flora: new evidence and a new approach. *Botanica Helvetica* **94:** 1–13.

Stebbins GL. 1985. Polyploidy, hybridization, and the invasion of new habitats. *Annals of the Missouri Botanical Garden* **72:** 824–832.

Stehlik I. 2000. Nunataks and peripheral refugia for alpine plants during the Quaternary glaciation in the middle part of the Alps. *Botanica Helvetica* **110:** 25–30.

Stock M, Ustinova J, Lamatsch DK, Schartl M, Perrin N, Moritz C. 2010. A vertebrate reproductive system involving three ploidy levels: hybrid origin of triploids in a contact zone of diploid and tetraploid palearctic green toads (*Bufo viridis* subgroup). *Evolution* **64:** 944–959.

Suda J, Krahulcová A, Trávníček P, Krahulec F. 2006. Ploidy level versus DNA ploidy level: an appeal for consistent terminology. *Taxon* **55:** 447–450.

Thiel-Egenter C, Alvarez N, Holderegger R, Tribsch A, Englisch T, Wohlgemuth T, Colli L, Gaudeul M, Gielly L, Jogan N, Linder HP, Negrini R, Niklfeld H, Pellecchia M, Rioux D, Schönswetter P, Taberlet P, van Loo M, Winkler M, IntraBioDiv Consortium, Gugerli F. 2011. Break zones in the distributions of alleles and species in alpine plants. *Journal of Biogeography* **38:** 772–782.

Tribsch A, Schönswetter P. 2003. Patterns of endemism and comparative phylogeography confirm palaeoenvironmental evidence for Pleistocene refugia in the eastern Alps. *Taxon* **52:** 477–497.

Vandel A. 1928. La parthénogénèse geographique. Contribution à l'étude biologique et cytologique de la parthénogénèse naturelle. *Bulletin Biologique de la France et de la Belgique* **62:** 164–281.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Summary of K-means analysis based on AFLP data. Delta K values were counted following Evanno *et al*. (2005) and are showed for whole dataset (A), diploid dataset (B) and tetraploid dataset (C).

Figure S2. Summary of Structure analysis based on AFLP data. Delta K-values were counted following Evanno *et al*. (2005) and are showed for whole dataset (A), diploid dataset (B) and tetraploid dataset (C). Values of ln probability of the data for each number of groups (K) are plotted against the K-value for whole dataset (D), diploid dataset (E) and tetraploid dataset (F).