UNIVERSITAT DE BARCELONA
FACULTAT DE BIOLOGIA
Departament de Biologia Vegetal, Unitat de Botànica

Programa de Doctorat: Biologia Vegetal
Bienni: 2002-2004

INSTITUT BOTÀNIC DE BARCELONA (CSIC-ICUB)

EVOLUCIÓ, FILOGÈNIA I SISTEMÀTICA DEL COMPLEX *ARCTIUM-COUSINIA*

Memòria presentada per Sara López Viñallonga per a optar al títol de Doctor per la Universitat de Barcelona

Amb el vist-i-plau dels directors de tesi:

Dra. Núria Garcia Jacas   Dr. Alfonso Susanna de la Serna

i del tutor de tesi:

Dr. Ramon M. Masalles i Saumell

Sara López Viñallonga
Barcelona, 2009

Genetic structure and diversity in *Arctium minus* (Compositae): effects of historical climate change and life history

López-Vinyallonga, S.[1*], P. S. Soltis[2], A. Susanna[1], D. E. Soltis[3] & N. Garcia-Jacas[1]

[1]*Botanic Institute of Barcelona (CSIC-ICUB), Pg. del Migdia s. n., 08038 Barcelona, Spain*
slopez@ibb.csic.es (author for correspondence)

[2]*Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA*

[3]*Department of Botany, University of Florida, Gainesville, FL 32611, USA*

Abstract

Plant life history is one of the main factors affecting the genetic structure of organisms. In the present work the genetic variation in *Arctium minus* is investigated in order to test the effect of mating system (facultative self-pollination), dispersal strategy (efficient seed dispersal by epizoochory) and life cycle (biannuality) on its genetic structure. The analysis of eight microsatellite loci and 14 populations from most of its worldwide distribution has been used for this purpose. The observed patterns of genetic variation are consistent with *A. minus* life style. Low gene diversity, high inbreeding values and significant homozygote excess were recovered, factors linked with the self-pollination ability of *A. minus* and its short-life cycle. Long distance seed dispersal is suggested as the main cause for three patterns observed: i) most of the genetic variability is found among instead of within populations, ii) absence of isolation by distance across Europe and iii) the lack of influence of Pleistocene climate changes in the European populations. No differences are found between European and American populations. It is suggested that original biogeographic patterns in *A. minus* may have been blurred by human activity.

INTRODUCTION

Nonrandom spatial distribution of genotypes, i.e. spatial genetic structure, is more the rule than an occasional phenomenon in the natural populations of plant species (i.e. McCauley 1997; Balloux & Lugon-Moulin 2002; Palmé & Vendramin 2002; Rendell & Ennos 2002; Leblois *et al.* 2006). A variety of agents or processes such as environmental and historical factors and life histories can affect the partitioning of genetic variability among plant populations. Regarding life history, plant breeding system, modes of seed dispersal, and length of life cycle are important nonexclusive factors affecting the spatial genetic structure in wild populations (Austerlitz *et al.* 2000;

Nordborg 2000; Charlesworth 2003; Clauss & Mitchell-Olds 2006). Differences in self-fertilization rates are among the main factors affecting the genetic diversity. As a trend, lower genetic variability is expected for inbreeding than outcrossing plants, and in inbreeding species the variability is found between populations rather than within them (*cf.* Charlesworth 2003).

Different modes of seed dispersal provide diverse capacities for the species to spread their propagules to long distances and therefore determine both the distribution of plant populations and the species range. Likewise, dispersal distances affect the spatial genetic structure of the species, since longer dispersal distances yield more separate populations and stronger founder effects (long dispersions are rare events in which few colonizers are involved, and the gene flow with other populations is low or it does not exist). However, long dispersions can also increase the gene flow between previously isolated populations (Nichols & Hewitt 1994; Petit *et al*. 2003).

The effect of colonization processes on genetic diversity and population structure also depends on the life cycle of the species, having a lesser impact in long life-cycle species than in short-lived plants. Long life-cycle species have a longer juvenile phase and therefore a delayed first reproduction which allows a large increase in the number of initial founders (by immigration) of a given population before reproduction begins (Austerlitz *et al*. 2000).

In this context, *Arctium minus* (Hill) Bernh. provides a useful system to analyze the effect of the different life history factors on the spatial structure of the genetic diversity, since this species combines a mixed mating system with short-life cycle and the ability of long distance dispersal as we shortly comment.

Arctium minus (Lesser Burdock) is a diploid herbaceous plant, native to Eurasia. It also grows in the northernmost Africa, where it is rare, and it is widespread and often naturalized as a weed in many parts of South and North America, where is considered as an invasive species in many areas of US and Canada. In spite of its wide distribution, *A. minus* populations are patchily distributed due to its ecological requirements. This species has a broad ecological range with regard to climate, altitude and soil type although it prefers moderate to high levels of moisture as well as high soil nitrogen content. As a consequence, populations of *A. minus* are restricted to nitrified-mesic places like open and disturbed woods, disturbed areas and pastures, abandoned fields and stream banks, and even roadsides.

Lesser burdock has a mixed mating system with predominant allogamy, although it can be self-pollinated if insect pollination fails (Gross *et al.* 1980; Fenner *et al.* 2002). According to Gross *et al.* (1980) high selfing levels lead to an increase of achene abortion. In addition, hybridization between sympatric *Arctium* species has been largely described (*cf.* Repplinger *et al.* 2007). However, Repplinger *et al.* (2007) have suggested the existence of pre- and postzygotic isolation mechanisms to explain the few hybrids detected in mixed populations of *A. minus*, *A. lappa* L. and *A. tomentosum* Mill.

Seed dispersal is epizoochorous by the attachment of burrs to the coats of animals and therefore achenes are dispersed altogether as a whole. Between 29 and 47 seeds are produced per head, but more than 60% of the seeds are predated before dispersion (Hawthorn & Hayne 1978). In contrast, the germination success of this species is of 90% (Gross *et al.* 1980).

Long-distance dispersion of lesser burdock burrs of up to 700 m (166 m in average) has been reported (Couvreur *et al.* 2008). The main vectors of dispersion of this species are medium-sized or big mammals, both wild and domestic, although some birds and small mammals have been reported as well. In addition, it is very likely that the dispersion of *A. minus* has been favoured by humans. Actually, lesser burdock is hypothesized to be introduced to North America by early English and French settlers in the mid 17th century (Gross *et al.* 1980).

Regarding life cycle, *A. minus* is a biennial monocarpic species growing vegetatively and then dying after flowering during the second year. However, several authors have reported that under unfavourable conditions many plants may take four or more years to flower or even not flower (Gross *et al.* 1980; Gross & Werner 1983). Moreover, Rollo *et al.* (1984) observed that this species colonizes habitats in early stages of succession but over the following 3 years it declines rapidly and at the 4th year it becomes rare, suggesting an ephemeral character for lesser burdock populations.

Nothing is known about the genetic diversity of *A. minus* and how this genetic diversity is structured among its populations.

The current distribution of *A. minus* throughout Eurasia, including the glacial refugia of South Europe, North Africa and Middle East (i.e. Turkey), indicate that this species likely endured during the glacial and interglacial periods occurred in the Pleistocene. As many other species now widespread in Europe, lesser burdock could recolonize the continent from southern glacial refugia during the present interglacial

period, and thus its present range may be consequence of glacial advance and retreat. This migration pattern across Europe during Pleistocene has been detected in many plants using molecular markers, including microsatellites (e.g. Rendell & Ennos 2002; Grivet & Petit 2003; Sharbel *et al.* 2006), and a common pattern for genetic diversity partition has been established (Petit *et al.* 2003). Thus populations in different glacial refugia, as a consequence of prolonged isolation, should be highly divergent. Moreover, the intrapopulation diversity should decline away from refugia as a consequence of successive founder events during European recolonization. However, this cline can become blurred if populations contact during the recolonization period.

In the present work we analyze the spatial model of genetic variation within and among 11 populations from the native range of *A. minus*, namely Europe, as well as from the limits of its distributional area (Turkey, Morocco and Norway) using eight microsatellite loci (López-Vinyallonga *et al.* in prep). Three populations from Americas are also analyzed in order to compare them with those of the native range of this species. In this study we try to evaluate the effect on the genetic structure of the mating system, seed dispersal strategy and life cycle of *A. minus*. We test if the current spatial genetic structure of lesser burdock across its natural range fits to the expected pattern for one species which has underwent the past Pleistocene climatic changes or, in contrast, that pattern has been masked by the effect of the life history factors. If the latter occurs, the spatial genetic structure should fit that expected for a species with long-distance seed dispersion, short life cycle and mixed mating system.

MATERIAL AND METHODS

Sampling

We sampled 14 populations of *Arctium minus* representing its distribution area (Table 2). Eleven populations represented the natural range of the species (Table 2; Fig. 1): nine from Europe, one from Morocco and one from Turkey. Populations from Morocco, Turkey, South of Italy and the Iberian Peninsula were selected in order to include populations from places which were important refugia during Pleistocene climatic changes. Five populations were included from Centre and North Europe, where the diversification centre of *Arctium* is thought (Tscherneva 1962; Duistermaat 1996), and

one population from Norway was used to represent the northernmost edge of *A. minus* (where is a rare species). Finally, three populations from America (two from North America and one from South America; Table 2) were included in the study to compare them with the populations from the natural range of the species.
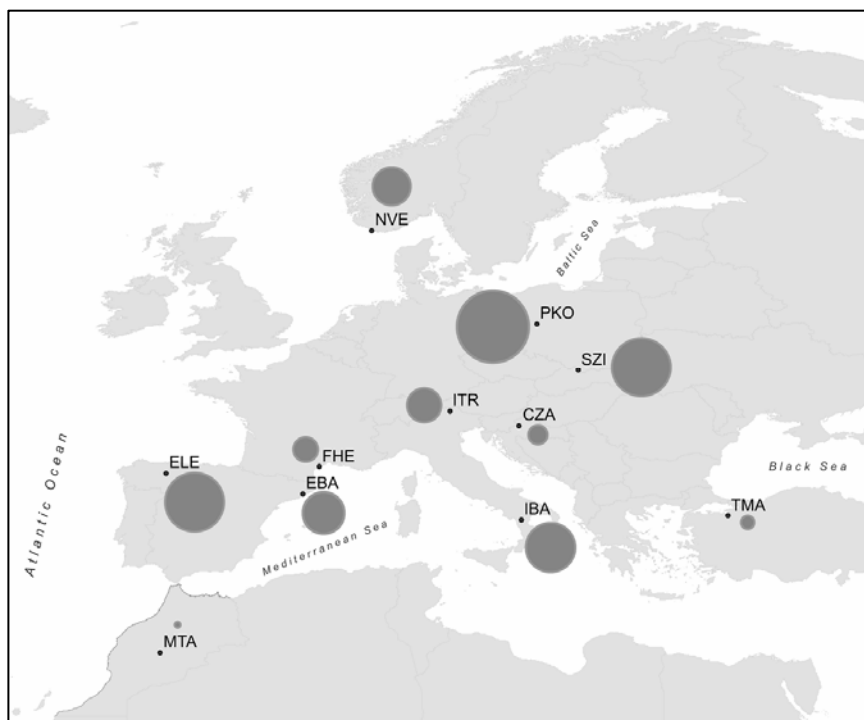

Fig. 1. Native area of distribution of *A. minus*. Circles represent H_E (expected heterozygosity under Hardy-Weinberg equilibrium) values observed for each population. CZA: Croatia, EBA and ELE: Spain, FHE: France, IBA and ITR: Italy, MTA: Morocco, NVE: Norway, PKO: Poland, SZI: Slovakia, TMA: Turkey.

Thirty individuals per population (368 individuals in total) were sampled, except for those populations with lower number of plants (CAN, MTA, NVE, and USA), for which all individuals in the population were sampled (Table 2).

DNA isolation and Microsatellite loci

Genomic DNA was extracted from dried leaf tissue of up to 30 specimens, collected from each wild population, using the NucleoSpin® Plant Kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany).

Eight microsatellite loci, showing different polymorphism levels, were used in this work (Table 1). Microsatellite loci were amplified using FAM, NED, PET and VIC fluorescently labelled forward primers as explained in López-Vinyallonga *et al*. (in

pres.). Genotyping was performed on an ABI 3730xl DNA Analyzer (Applied Biosystems) using LIZ600 size standard at the Interdisciplinary Center for Biotechnology Research (ICBR) facility at the University of Florida. Fragment analysis was performed with GENEMARKER 1.5 (SoftGenetics, LLC) software.

Table 1. Configuration of 8 micosatellite loci used in this work. Optimized annealing temperature (T_a), repeat motifs, size ranges of PCR products, number of alleles observed (*NA*).

Locus	T_a (°C)	Repeat Motif	Size Range (bp)	*NA*
Am30	52	$(CA)_8$	180-189	3
Am31	52	$(CA)_9$	182-272	9
Am32	52	$(CA)_8$	198-226	8
Am33	52	$(CA)_9$	197-223	10
Am34	52	$(GA)_9$	158-185	7
Am35	50	$(TATG)_6(TG)_4[(TATG)_2(TG)_4]_4$	152-238	16
Am37	50	$(TC)_9$	171-179	4
Am39	52	$(CA)_3CGC(CA)_5$	252-256	2

Data analysis

For each population the mean number of alleles per locus (*NA*), the observed heterozygosity (H_O), the unbiased expected heterozygosity (H_E; Nei 1978), and the fixation index (F_{IS}; Weir & Cockerham 1984) were calculated for each locus using GENETIX software (Belkhir *et al.* 1996-2004). Significant deviations from Hardy-Weinberg equilibrium expectations were assessed in each population (GENETIX; 10,000 permutations of F_{IS}). In order to test for significant differences in H_E and F_{IS} between populations a Mann-Whitney *U*-test was used for each pairwise comparison. This test was also implemented between American and Eurasian-North African populations.

Nei's (1973) population structure statistics (unbiased for sample size; Nei & Chesser 1983) were calculated for all 14 *A. minus* populations as a whole and for the 11 populations of the natural range of the species. The differentiation between pairs of populations was quantified with θ (F_{ST}; Weir & Cockerham 1984) using GENETIX, and a permutation test of the pairwaise θ differentiation was performed (GENETIX; 10,000 permutations). A chord distance matrix (Dc; Cavalli-Sforza & Edwards 1967) among populations was constructed from allele frequency data using MICROSAT ver. 1.5b (Minch *et al* 1995), from which 1,000 bootstrapped replicate matrices were

computed. Bootstrap analysis was carried out with the software package PHYLIP ver. 3.66 (Felsenstein 2006) using the module NEIGHBOR to compute Neighbor-Joining dendrograms for all bootstrapped matrices, and the module CONSENSE to produce an extended majority-rule consensus tree.

Isolation-by-distance among populations was investigated by computing the correlation between the matrix of pairwise genetic differentiation $[F_{ST}/(1\text{-}F_{ST})]$ and the matrix of the logarithm of geographical distances (Rousset 1997) by applying the Mantel test (1,000 permutations) using GENETIX. The geographical distances between pairs of populations were calculated from the longitudes and latitudes given in Table 2. An analysis of molecular variance (AMOVA), using the program ARLEQUIN version 3.11 (Excoffier *et al.* 2005), was performed to estimate the geographical structure of genetic variation. AMOVA was carried out at different hierarchical levels between the Eurasian and North African populations: 1) among and within populations, without regional grouping, 2) among geographical groups (regions), among populations within regions and among individuals within populations. Regions were defined: a) according to the sampling scheme and b) according to the phylogenetical result. Moreover, an AMOVA analysis was performed with all 14 populations, nested in two regions (America vs. Eurasia-North Africa).

RESULTS

In total, 59 alleles were observed from the eight loci surveyed. The number of alleles per locus ranged from two (Am39) to 16 (Am35), being the mean number of alleles per locus 7.13. All loci were monomorphic in at least one of the studied populations. A high significant homozygote excess overall was detected (F_{IS} = 0.347 ± 0.185; $P < 0.001$). All loci except two (Am32 and Am37) showed significantly positive inbreeding coefficients (data not shown).

At the population level, values of the inbreeding coefficient were significantly greater than zero for seven of the 14 populations, while in four populations (CZA, ITR, MTA and TMA) we recovered $F_{IS} < 0$ suggesting heterozygote excess (Table 2). The Mann-Whitney *U*-test indicated significant differences in F_{IS} for five pairwise comparisons (EBA-IBA, EBA-MTA, MTA-IBA, MTA-PKO and MTA-SZI; $P < 0.05$).

Different levels of polymorphism (at 95%) were found between the populations, being TMA, MTA and ITR the less polymorphic, while the most polymorphic were SZI and ELE. The number of alleles per population ranged from nine (ITR; 1.12 in average) to 26 (PKO; 3.25 in average), and 16 (18 when U.S.A. population was considered) privative alleles were detected in seven populations (Table 2). All eleven Eurasian-North African populations showed low gene diversity (from 0.037 for MTA to 0.444 for PKO; Table 2, Fig. 1), but significant differences ($P < 0.05$; Mann-Whitney U-test) were found for 10 pairwise comparisons (PKO-CZA, PKO-MTA, PKO-TMA, ELE-CZA, ELE-MTA, ELE-TMA, SZI-CZA, SZI-FHE, SZI-MTA, SZI-TMA). Non-significant difference in H_E was found when the U-test was performed between American *versus* Eurasian-North African populations.

Table 2 Sampling localities for 14 populations from across the range of *Arctium minus* and main parameters of genetic diversity within populations. Lat., latitude; Long., longitude; N, number of individuals; K, number of alleles (na; mean number of alleles); Priv., number of privative alleles; P_{95}, proportion of polymorphic loci at 95%; H_O, observed heterozygosity; H_E, unbiased expected heterozygosity (Nei, 1978); F_{IS}, inbreeding coefficient. *$P < 0.05$.

Code	Population	Country	Lat.	Long.	N	K(na)	Priv	P_{95}	H_O	H_E	F_{IS}
ABA	La Plata	Argentina	S34°55'	W57°57'	30	16 (2)	0	0.500	0.049	0.108	0.553*
CAN	-	Canada	-	-	16	12 (1,5)	0	0.375	0.102	0.115	0.118
CZA	Medveščak	Croatia	N45º49'	E15º59'	30	14 (1,75)	0	0.375	0.150	0.113	-0.333
EBA	Canyamars	Spain	N41º35'	E2º24'	30	16 (2)	0	0.625	0.162	0.258	0.380*
ELE	Murias de Paredes	Spain	N42º52'	W6º11'	30	20 (2,5)	1	0.875	0.185	0.363	0.496*
FHE	Plateau de l'Escandorque	France	N43°16'	E3º26'	30	15 (1,87)	0	0.375	0.089	0.154	0.424*
IBA	Prastio	Italy	N39º56'	E16º7'	30	17 (2,12)	7	0.500	0.107	0.306	0.655*
ITR	Kloster Neustift	Italy	N46º44'	E11º38'	30	9 (1,12)	0	0.286	0.243	0.209	-0.167
MTA	Demnate	Morocco	N31º37'	W6º33'	17	10 (1,25)	4	0.286	0.038	0.037	-0.028
NVE	Farsund	Norway	N58°4'	E6º45'	8	15 (1,87)	1	0.500	0.205	0.241	0.156
PKO	Kórnik	Poland	N52º14'	E17º5'	30	26 (3,25)	1	0.750	0.222	0.444	0.505*
SZI	Liesek	Slovakia	N49º21'	E19º40'	30	21 (2,62)	2	0.857	0.287	0.367	0.223
TMA	Bursa	Turkey	N40º11'	E29º3'	30	13 (1,62)	0	0.125	0.112	0.085	-0.329
UWA	Washington DC	USA	N38°58'	W77° 1'	28	11 (1,37)	2	0.375	0.102	0.183	0.446*

Total gene diversity (H_T) through all loci and populations was 0.58, ranging from 0.167 (Am30) to 0.837 (Am35) (Table 3). The population structure statistics (Table 3) showed that the interpopulation component explains most of the genetic variation detected in *A. minus* ($G_{ST} = 0.599$).

Table 3 Main parameters of gene diversity across loci among populations for *A. minus*. H_T, total gene diversity; H_S, within-population gene diversity; D_{ST}, mean gene diversity among populations; G_{ST}, Nei's coefficient of gene diversity among populations. *$P < 0.001$

	H_T	H_S	D_{ST}	G_{ST}
Am30	0.167	0.048	0.118	0.709
Am31	0.772	0.354	0.417	0.541
Am32	0.736	0.314	0.422	0.573
Am33	0.831	0.332	0.499	0.601
Am34	0.642	0.243	0.399	0.621
Am35	0.837	0.350	0.488	0.583
Am37	0.170	0.132	0.038	0.225
Am39	0.484	0.087	0.397	0.820
Multilocus	0.580	0.233	0.347	0.599

In the pairwise θ analysis all populations showed high significant differentiation from each other ($P < 0.001$), except the NVE and PKO comparison which was non-significant (Table 4). PKO and SZI were the most related populations (excluding the comparison NVE-PKO), while MTA and ITR were the populations most genetically differentiated from the rest of populations.

Table 4. Pairwise comparisons showing differentiation between populations based on θ values (Weir & Cockerham 1984). *P< 0.001. ABA: Argentina, CAN: Canada, CZA: Croatia, EBA and ELE: Spain, FHE: France, IBA and ITR: Italy, MTA: Morocco, NVE: Norway, PKO: Poland, SZI: Slovakia, TMA: Turkey, UWA: USA.

	ABA	CAN	CZA	EBA	ELE	FHE	IBA	ITR	MTA	NVE	PKO	SZI	TMA	UWA
ABA	0.000													
CAN	0.678*	0.000												
CZA	0.772*	0.564*	0.000											
EBA	0.507*	0.598*	0.717*	0.000										
ELE	0.445*	0.436*	0.482*	0.428*	0.000									
FHE	0.479*	0.509*	0.680*	0.484*	0.411*	0.000								
IBA	0.699*	0.689*	0.640*	0.629*	0.425*	0.638*	0.000							
ITR	0.849*	0.860*	0.866*	0.808*	0.723*	0.819*	0.748*	0.000						
MTA	0.889*	0.850*	0.862*	0.775*	0.657*	0.887*	0.778*	0.853*	0.000					
NVE	0.629*	0.347*	0.532*	0.538*	0.381*	0.431*	0.552*	0.807*	0.756*	0.000				
PKO	0.504*	0.386*	0.544*	0.452*	0.348*	0.371*	0.490*	0.518*	0.583*	0.251	0.000			
SZI	0.642*	0.582*	0.657*	0.548*	0.462*	0.593*	0.546*	0.755*	0.707*	0.545*	0.297*	0.000		
TMA	0.833*	0.811*	0.719*	0.768*	0.581*	0.797*	0.632*	0.868*	0.897*	0.782*	0.649*	0.633*	0.000	
UWA	0.729*	0.651*	0.763*	0.653*	0.469*	0.699*	0.669*	0.825*	0.727*	0.563*	0.461*	0.629*	0.823*	0.000

The bootstrap consensus neighbor-joining tree of populations showed four population groups (I-IV; Fig. 2). Only two of them were congruent with the

geographical distribution of the populations gathered. Thus, group I included the Iberian and French populations which are 233 km aside, and group III included the Central European populations from Poland and Slovakia which are separated by 355 km with no evident barrier between them. The Mantel test detected no correlation between increase in genetic differentiation with the log of geographic distance, and therefore an absence of isolation by distance was corroborated across all 11 populations ($r = 0.190$, $P = 0.182$).
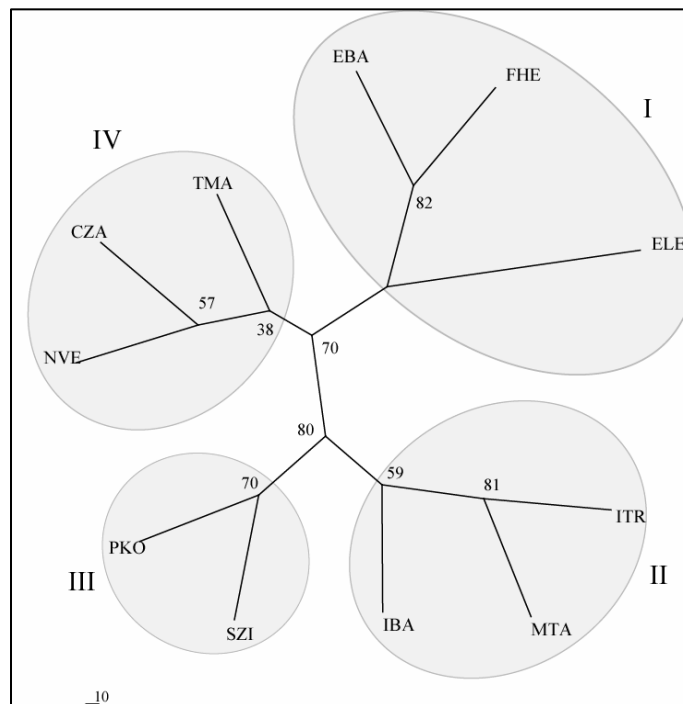
Fig. 2 Extended majority-rule consensus neighbor-joining tree based on bootstrap analysis of the 11 populations from the natural range of *A. minus*. I to IV highlight the groups recovered. CZA: Croatia, EBA and ELE: Spain, FHE: France, IBA and ITR: Italy, MTA: Morocco, NVE: Norway, PKO: Poland, SZI: Slovakia, TMA: Turkey.

Table 5 shows the results for the AMOVA test. When AMOVA was performed among populations without considering higher levels of groupings, the majority of variation was partitioned among populations (60.49%), with a highly significant F_{ST} value (0.605, $P < 0.001$), and with only 39.51% variation within populations. The three-level hierarchical AMOVA considering regional groupings according to the sampling scheme (putative glacial refugia *vs*. Central European populations) yielded in a partition where most of the variation was among populations within regions (62.16%), 39.87% of the variation was within populations, and no variation was found between regions. Only

when we considered the four groups obtained in the neighbor-joining tree as regions, the test detected a significant partition of the variation between regions (17.19%; $F_{CT} = 0.172$, $P < 0.01$). Finally, when the AMOVA analysis was performed with all 14 populations, nested in two regions (America *vs*. Eurasia-North Africa), no variation was detected between both regions.

Table 5. Analisis of Molecular Variance (AMOVA) in *Arctium minus*. Models and populations included: Natural area of distribution (CZA, EBA, ELE, FHE, IBA, ITR, MTA, NVE, PKO, SZI, TMA); Putative refugia *vs*. Central Europe: (EBA, ELE, IBA, MTA, TMA) - (CZA, FHE, ITR, NVE, PKO, SZI); Natural area of distribution partitioned in 4 groups according to Fig. 2 (IBA, ITR, MTA) – (PKO, SZI) – (CZA, NVE, TMA) – (EBA, ELE, FHE); Americas *vs*. Europe (CZA, EBA, ELE, FHE, IBA, ITR, MTA, NVE, PKO, SZI, TMA) – (ABA, CAN, UWA). n.s. no significative.

Model	Partitioning	Variance (%)	F-statistic	P
Natural area of distribution	Among populations	60.49	$F_{ST} = 0.605$	< 0.001
	Within populations	39.51		
Putative refugia vs. Central Europe	Among groups	-2.04	$F_{CT} = -0.020$	n.s.
	Among populations within groups	62.16	$F_{SC} = 0.609$	< 0.001
	Within populations	39.87	$F_{ST} = 0.601$	< 0.001
Natural area of distribution partitioned in 4 groups	Among groups	17.19	$F_{CT} = 0.172$	< 0.01
	Among populations within groups	44.49	$F_{SC} = 0.537$	< 0.001
	Within populations	38.32	$F_{ST} = 0.617$	< 0.001
Americas vs. Europe	Among groups	2.83	$F_{CT} = 0.028$	n.s.
	Among populations within groups	60.80	$F_{SC} = 0.626$	< 0.001
	Within populations	36.37	$F_{ST} = 0.636$	< 0.001

DISCUSSION

Our results show that the spatial genetic structure found for *Arctium minus* in Europe does not conform the expected pattern for Europe recolonization from southern glacial refugia, i.e. (i) decline in genetic diversity northwards from refugied populations, and (ii) strong genetic differentiation between populations in different refugia in contrast with a lesser population differentiation between the recolonized areas (*c.f.* Petit *et al.* 2003). (i) The former pattern is not found since populations of *A. minus* throughout Europe do not show a clear pattern with regard to the genetic diversity; much in contrast, the populations show different levels of genetic diversity independently of the European region considered (Table 2; Fig. 1). Thus, for example, the most diverse

populations are PKO and SZI, while other populations from centre Europe show little diversity (i.e. ITR and CZA). This is also the case for the populations located in the commonly accepted glacial refugia where both moderately diverse (EBA, ELE, IBA) and poor populations (MTA, TMA) are found. (ii) The last pattern is neither found although significant differences in H_E are obtained for comparisons between the Iberian Peninsula (ELE) respect the populations from Morocco (MTA) and Turkey (TMA). This result may suggest the existence of glacial refugia in South Europe, North Africa and Middle East although this pattern has not been detected for any other pairwise comparison between the remaining putative refugia (i.e. IBA, EBA and ELE).

Petit *et al*. (2003) shown that the most simple expected predictions for genetic diversity distribution in recolonizing species can be altered, being the most diverse populations those at intermediate latitudes, instead of the southernmost ones, due to the admixture of divergent lineages from separate refugia. We have found significant differences in H_E for comparisons between Poland (PKO) and Slovakia (SZI) respect the populations from Morocco (MTA) and Turkey (TMA), which agrees with Petit *et al*. (2003), but not respect the populations from the remaining putative refugia, namely Italy and the Iberian Peninsula.

It is worth to mention that many authors have proposed the existence of cryptic northern refugia in Europe, i.e. Hungary (Willis *et al.* 2000), Slovakia (Litynska-Zajac 1995), Belgium (Leroi-Gourhan 1992) and a more controversial one in Norway (Stewart & Lister 2001, Kullman 2001, 2002, Birks *et al.* 2005, Tollefsrud *et al.* 2008). These refugia would have been in areas of sheltered topography that provided suitable stable microclimates during the Last Cold Stage (Stewart & Lister 2001). This would have been the case for *A. minus* since this cold-tolerant species shows its maximum diversity in the populations from Poland and Slovakia, in Central Europe, which likewise have privative alleles.

Regarding population divergence, a strong differentiation was detected between *A. minus* populations, as the population structure statistics and pairwise θ analysis showed. However, as found with H_E, the detected population genetic structure did not agree with the expected pattern under the glacial refugia model. AMOVA analysis indicated that most of the genetic variation was partitioned among populations, independently if refugee or not refugee areas were considered. Moreover, the spatial genetic structure detected does not fit with an isolation by distance model (Mantel's test

not significant), which should be the expected during the Europe recolonization process, as has been shown in other European plants (e.g. Sharbel *et al*. 2000, Schönswetter *et al*. 2003). In spite all these considerations, there are few signs which may suggest that actually *A. minus* adapted to the Pleistocene climatic changes like many other European species. Thus, the presence of seven privative alleles in the population from South Italy and four in the population from Morocco points out to putative refugia in these areas from where the species might radiate northwards.

To sum up, according to our results the current population genetic structure of *A. minus* does not reflect the demographic processes associated with the Pleistocene climatic changes although some signs suggest the likely existence of refugia in North and South Europe. Instead, the detected structure seems to fit well with the life history of this species which combines long-distance seed dispersal with mixed breeding system and short generation time. All three factors have been shown to affect the spatial genetic structure of plant species (Austerlitz *et al*. 2000; Nordborg 2000; Charlesworth 2003; Clauss & Mitchell-Olds 2006), and as a consequence they can blur the genetic signature of colonization and range expansion during Pleistocene. Another factor that may confound genetic patterns is hybridization. This is a frequent phenomenon in *A. minus* and although the pollen produced by hybrid specimens is usually viable, hybrid achenes are often abortive (Duistermaat 1996). Thus hybrid populations are no permanent but there is interespecific genetic flow through pollen grains. And of course we can not ignore human influence since the dispersion of this species is highly correlated with movements of livestock. In addition to this, the lack of information from the fossil record hinders a good reconstruction of the past distribution of *A. minus*.

Range expansion via multiple long-distance dispersal events can result in population genetic structure with little or no isolation by distance. McCauley *et al*. (2003) showed that the best scenario to explain the genetic structure of *Silene vulgaris* in North America (high population differentiation and no isolation by distance) was that in which the range expansion was by long-distance dispersal events. Distances of 166 m in average for seed dispersal in *Arctium minus* have been reported being frequent dispersal distance of 700 m (Couvreur *et al*. 2008), and distances for adhesive-dispersed seeds up to 2.9 km (including *A. minus*; Mouissie *et al*. 2005) and 4.4 km (Cain *et al*. 1998) have also been reported. Therefore, the mode of seed dispersion fits well with the population genetic structure observed in *A. minus*, especially if we consider that this

species behaves as a weed in nature, being adapted to disturbed habitats and having ephemeral populations with frequent events of local extinctions and recolonizations.

Isolated populations, as a consequence of long-distance seed dispersion events, frequently will undergo bottlenecks and founder effects which will cause low within-population diversity (*c.f.* Charlesworth 2003) which is congruent with the low intrapopulation genetic diversity found in *A. minus*. Moreover, this species only produces 29-47 seeds per head, from which more than 60% are predated before dispersion (Hawthorn & Hayne 1978). Therefore, the founder effect in the new colonizing populations must be higher due to few numbers of initial plants. This situation agrees with the high inbreeding coefficient detected in all populations of *A. minus*, since the isolation and the few breeders increase the probability of self-pollination in a species with mixed breeding system. In addition, inbreeding interacts with population isolation, diversity and structure, decreasing within-population diversity while increasing among population differentiation (*c.f.* Charlesworth 2003).

The short-life cycle of *A. minus* may contribute to the high inbreeding values recovered in the present work. Austerlitz *et al.* (2000) stated that short-lived plants arriving to a new site can reproduce the next year and therefore the offspring of these first occupants, which are genetically similar to parental founder individuals, have the opportunity to colonize the whole space. This behaviour leads to a genetically homogeneous population and thus results in more variability among populations than within them.

This pattern where most of the genetic variability accounts between populations, as found in *A. minus*, has been reported for worldwide distributed organisms (Sivasundar & Hey 2003) and autogamous species (Hamrick & Godt 1996), characteristics found in lesser burdock. Low gene flow is been stated as another cause of differentiation among populations and in spite of the highly efficient seed dispersal in *A. minus*, its present patchy distribution may actually restrict gene flow among populations.

Although most of the genetic differentiation in *A. minus* is found among populations, differences are independent from the regions where the populations are located, even if they are separated by important geographic barriers. This is the case for the American populations which are not significantly differentiated from those in Europe, Morocco and Turkey. Furthermore, the populations from Americas are

supposed to date from the mid 17th century (Gross *et al.* 1980) while a fossil *Arctium sp.* achene from Germany was dated to 7-9 mya (Wähnert, pers. comm.). The most likely explanation for these evidences is a combination of recent long distance dispersal events and human influence, both shaping the genetic structure of *A. minus* populations.

ACKNOWLEDGEMENTS

REFERENCES

Austerlitz, F., Mariette, S., Machon, N., Gouyon, P. & Godelle, B. (2000). Effects of colonization processes on genetic diversity, differences between annual plants and tree species. *Genetics* 154: 1309–1321.

Balloux, F. & Lugon–Moulin, N. (2002). The estimation of population differentiation with microsatellire markers. *Molecular Ecology* 11: 155–165.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (1996-2004). *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Birks, H. H., Larsen, E. & Birks, H. J. B. (2005). Did tree-*Betula*, *Pinus* and *Picea*survive the last glaciation along the best coast of Norway? A review of the evidence, in light of Kullman (2002). *Journal of Biogeography* 32: 1461–1471.

Cain, M. L., Damman, H. & Muir, A. (1998). Seed dispersal and the Holocene migration of woodland herbs. *Ecological Monographs* 68: 325–347.

Cavalli-Sforza, L. L. & Edwards, A. W. F. (1967). Phylogenetic analysis, models and estimation procedures. *American Journal of Human Genetics* 19: 233–257.

Charlesworth, D. (2003). Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society B* 358: 1051–1070.

Clauss, M. J. & Mitchell–Olds, T. (2006). Population genetic structure of *Arabidopsis lyrata* in Europe. *Molecular Ecology* 15: 2753–2766.

Couvreur, M., Verheyen, K., Vellend, M., Lamoot, I., Cosyns, E., Hoffmann, M. & Hermy, M. (2008). Epizoochory by large herbivores, merging data with models. *Basic and Applied Ecology* 9: 204–212.

Duistermaat, H. (1996). Monograph of *Arctium* L (Asteraceae). Generic delimitation (including *Cousinia* Cass. p. p.), revision of the species, pollen morphology and hybrids. *Gorteria* 3: 1–143.

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

Felsenstein, J. (2006). *PHYLIP (Phylogeny Inference Package) Version 3.66*, Distributed by the author. Department of Genome Sciences, University of Washington, Seattle. <http,//evolution.genetics.washington.edu/phylip.html>.

Fenner, M., Cresswell, J. E., Hurley, R. A. & Baldwin, T. (2002). Relationship between capitulum size and pre-dispersal seed predation by insect larvae in common Asteraceae. *Oecologia* 130: 72–77.

Grivet, D. & Petit, R. J. (2003). Chloroplast DNA phylogeography of the hornbeam in Europe, evidence for a bottleneck at the outset of postglacial colonization. *Conservation Genetics* 4: 47–56.

Gross, R. S. & Werner, P. A. (1983). Probabilities of survival and reproduction relative to rosette size in the common burdock (*Arctium minus*, Compositae). *American Midland Naturalist* 109: 184–193.

Gross, R. S., Werner, P. A. & Hawthorn, W. R. (1980). The Biology of Canadian Weeds. 38. *Arctium minus* (Hill) Bernh, *A. lappa* L. *Canadian Journal of Plant Science* 60: 621−634.

Hamrick, J. L. & Godt, M. J. W. (1996). *Conservation genetics of endemic plant species*. Pp. 281–304. In: Avise, J. C. & Hamrick, J. L. (eds), *Conservation Genetics*. Chapman and Hall, New York.

Kullman, L. (2001). Immigration of *Picea abies* into North-Central Sweden. New evidence of regional expansion and tree-limit evolution. *Nordic Journal of Botany* 21: 39–54.

Kullman, L. (2002). Boreal tree taxa in the central Scandes during the Late-Glacial: implications for Late-Quaternary forest history. *Journal of Biogeography* 29: 1117–1124.

Leblois, R., Estoup, A. & Streiff, R. (2006). Genetics of recent habitat colonization and reduction in population size, does isolation by distance matter? *Molecular Ecology* 15: 601–3615.

Leroi-Gourhan, A. (1992). Bois et pollens: étude complémentaire. *Bulletin de la Société Botanique de France* 139: 273–280.

Litynska-Zajac, M. (1995). *Anthracological analysis*. Pp. 47–79. In: Hromada, J. & Kozolowski, J. (eds.), *Complex of Upper Palaeolithic sites near Moravany, Western Slovakia*, Jagellonian University Press.

McCauley, D. E. (1997). The relative contributions of seed and pollen movement to the local genetic structure of *Silene alba*. *Journal of Heredity* 88: 257–263.

McCauley, D. E., Smith, R. A., Lisenby, J. D. & Hsieh, C. (2003). The hierarchical spatial distribution of chloroplast DNA polymorphism across the introduced range of *Silene vulgaris*. *Molecular Ecology* 12: 3227–3235.

Minch, E., Ruiz–Linares, A., Goldstein, D., Feldman, M. & Cavalli-Sforza, L. L. (1995). *Microsat (Version 1.4d), a computer program for calculating various statistics on microsatellite allele data* <http,//lotka.stanford.edu/microsat/microsat.html>.

Mouissie, A. M., Lengkeek, W. & Van Diggelen, R. (2005). Estimating adhesive seed-dispersal distances, Field experiments and correlated random walks. *Functional Ecology* 19: 478–486.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA* 70: 3321–3323.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.

Nei, M. & Chesser, R. K. (1983). Estimation of fixation indexes and gene diversities. *Annals of Human genetics* 47: 253–259.

Nichols, R. A. & Hewitt, G. M. (1994). The genetic consequences of long distance dispersal during colonization. *Heredity* 72: 312–317.

Nordborg, M. (2000). Linkage disequilibrium, gene trees and selfing, an ancestral recombination graph with partial self-fertilization. *Genetics* 154: 923–929.

Palmé, A. E. & Vendramin, G. (2002). Chloroplast DNA variation, postglacial recolonization and hybridation in hazl, *Corylus avellana*. *Molecular Ecology* 11: 1769–1779.

Petit, R. J., Aguinagalde, I., de Beaulieu, J. L. *et al* (2003). Glacial refugia, hotspots but no melting pots of genetic diversity. *Science* 300: 1563–1565.

Rendell, S. & Ennos, R. A. (2002). Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology* 11: 69–78.

Repplinger, M., Johannesen, J., Seitz, A. & Comes, H. P. (2007). Morphological and molecular evidence for hybridization and introgression in Central European *Arctium* (Asteraceae). *Plant Systematics and Evolution* 268: 75–95.

Rollo, C. D., MacFarlane, J. D. & Smith, B. S. (1984). Electrophoretic and allometric variation in burdock (*Arctium* spp.), hybridization and its ecological implications. *Canadian Journal of Botany* 63: 1255–1261.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* 145: 1219–1228.

Sharbel, T. F., Haubold, B. & Mitchell–Olds, T. (2006). Genetic isolation by distance in *Arabidopsis thaliana*, biogeography and postglacial colonization of Europe. *Molecular Ecology* 9: 2109–2118.

Sivasundar, A. & Hey, J. (2003). Population genetics of *Caenorhabditis elegans*. The paradox of low polymorphism in a widespread species. *Genetics* 163: 147–57.

Stewart, J. R. & Lister, A. M. (2001). Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution* 16 (11): 608–613.

Tollefsrud, M. M., Kissling, R., Gugerli, F. *et al.* (2008). Genetic consequences of glacial survival and postglacial colonization in Norway spruce: combined analysis of mitochondrial DNA and fossil pollen. *Molecular Ecology* 17: 4134–4150.

Tscherneva, O. V. (1962). *Cousinia* Cass. In: Shishkin, B.K. (ed.), *Flora of the U.S.S.R.*, vol 27, pp. 108–357. Leningrad, Akademiya Nauk.

Weir, B. S. & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.

Willis, K. J., Rudner, E. & Sumegi, P. (2000). The full-glacial forests of central and southeastern Europe. *Quaternary Research* 53: 203–213.

SUMARI

7. Compendi de publicacions 45