

OPEN ACCESS – RESEARCH ARTICLE**Phylogeography above the species level for perennial species
in a composite genus**

**Karin Tremetsberger^{1,2,*}, María Ángeles Ortiz¹, Anass Terrab¹,
Francisco Balao¹, Ramón Casimiro-Soriguer^{1,3}, María Talavera¹ and
Salvador Talavera¹**

¹Departamento de Biología Vegetal y Ecología, Universidad de Sevilla, Apdo.
1095, 41080 Seville, Spain.

²Present address: Institute of Botany, Department of Integrative Biology and
Biodiversity Research, University of Natural Resources and Life Sciences,
Gregor Mendel Straße 33, 1180 Vienna, Austria.

³Present address: Departamento de Biología, CASEM, Universidad de Cádiz,
Campus Río San Pedro, 11510 Puerto Real, Spain.

*Corresponding author

Published by Oxford University Press on behalf of the Annals of Botany Company.
This is an Open Access article distributed under the terms of the Creative Commons
Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits
unrestricted reuse, distribution, and reproduction in any medium, provided the original
work is properly cited.

Corresponding author's e-mail address: karin.tremetsberger@boku.ac.at

Running head: Phylogeography in the western Mediterranean

Received: 22 April 2015; Revised: 30 July 2015; Accepted: 11 November 2015

ABSTRACT

In phylogeography, DNA sequence and fingerprint data at the population level are used to infer evolutionary histories of species. Phylogeography above the species level is concerned with the genealogical aspects of divergent lineages. Here, we provide a phylogeographic study to examine the evolutionary history of a western Mediterranean composite, focusing on the perennial species of *Helminthotheca* (Asteraceae, Cichorieae). We used molecular markers (AFLP; ITS and plastid DNA sequences) to infer relationships among populations throughout the distributional range of the group. Interpretation is aided by biogeographic and molecular clock analyses. Four coherent entities are revealed by Bayesian mixture clustering of AFLP data, which correspond to taxa previously recognized at the rank of subspecies. The origin of the group was in western North Africa, from where it expanded across the Strait of Gibraltar to the Iberian Peninsula and across the Strait of Sicily to Sicily. Pleistocene lineage divergence is inferred within western North Africa as well as within the western Iberian region. The existence of the four entities as discrete evolutionary lineages suggests that they should be elevated to the rank of species, yielding *H. aculeata*, *H.*

comosa, *H. maroccana* and *H. spinosa*, whereby the latter two necessitate new combinations.

Keywords: Amplified fragment length polymorphism; *Helminthotheca*; Iberian Peninsula; phylogeography; Strait of Gibraltar; western Mediterranean region; western North Africa.

INTRODUCTION

Intraspecific phylogeography as originally defined by Avise *et al.* (1987) is concerned with the geographic distribution of genealogical lineages at the level of conspecific populations. Originally based upon animal mitochondrial DNA sequence data, phylogeographic studies today make use of a variety of molecular techniques and are based upon DNA sequence and fingerprint data from the mitochondrial, plastid and nuclear genomes. The field of phylogeography today extends from the level of conspecific populations to the supraspecific level (Avise 2000) and targets organisms from the bacteria and eukaryote domains of life as well as viruses. Phylogeography of closely related species spans the boundary between reticulate and divergent relationships, where a partially braided collection of allelic pathways of interbreeding individuals bifurcates into two such collections (Avise 2000). Likewise, genealogical data at the population level of closely related species are especially helpful for explaining distributions resulting from historical and ecological factors.

A region that offered many opportunities for ecogeographic and effective geographic isolation (Sobel *et al.* 2010) due to its variegated geologic and climatic history since the Miocene is the western Mediterranean region. The formation of the Betic-Rif orogen led to the desiccation of the Mediterranean Sea between 6 and 5.3 million years ago (mya; Hsü *et al.* 1973), enabling the migration of animals and plants between the two continents (Bocquet *et al.* 1978). At the end of the Messinian, the Strait of Gibraltar opened and water from the Atlantic Ocean refilled the dry

Mediterranean Basin (Lonergan and White 1997), thus again disconnecting north-western Africa from south-western Europe.

Major climatic events after the opening of the Strait of Gibraltar were the advent of the Mediterranean climate and Pleistocene climatic oscillations. In northern Africa, the appearance of a local or temporary dry season was already recognizable in the floristic composition of the Miocene and this climate trend was accentuated in the Pliocene (Quézel 1978). In the north-western Mediterranean area, a Mediterranean climatic rhythm with summer drought appeared in the Pliocene, at approximately 3.2 mya (Suc 1984). In the Pleistocene, the water level of the Mediterranean Sea underwent great oscillations. During the last glacial maximum (0.02 mya), it was 120 m lower than today, so that the oceanic barrier between north-western Africa and south-western Europe was much reduced in extension, though not completely disrupted (Thiede 1978). The western Mediterranean region was an important refugium for temperate as well as Mediterranean plants and animals during the Pleistocene glacial periods (Médail and Diadema 2009). It is supposed that the changing environmental conditions and appearance/disappearance of biogeographic barriers have massively triggered plant and animal diversification in the Pliocene and Pleistocene (Fiz-Palacios and Valcárcel 2013).

The genus *Helminthotheca* (Asteraceae, Cichorieae, Hypochaeridinae) has a primarily western Mediterranean distribution. It has been separated from *Picris* at the generic level on the basis of the conspicuously enlarged outer involucral bracts and comprises the annual, heterocarpic species *H. echioides* and *H. balansae* and the perennial, homocarpic species *H.*

aculeata, *H. comosa* and *H. glomerata* (*Picris comosa* auct. alg.; Lack 1974, 1975). *Helminthotheca aculeata* has another subspecies, *H. aculeata* subsp. *maroccana*, in addition to the nominal subspecies. Similarly, *H. comosa* also has another subspecies, *H. comosa* subsp. *lusitanica*, in addition to the nominal subspecies.

The most recent common ancestor (MRCA) of extant taxa of *Helminthotheca* can be traced back to the Pliocene or Pleistocene (Tremetsberger *et al.* 2013). It is thus likely that populations have been affected by climatic changes occurring in the Mediterranean region in this epoch, possibly leading to range shifts, disruptions of previously continuous ranges and secondary contacts. This study aims at shedding light on the evolutionary history of the perennial species of *Helminthotheca*. Using molecular markers from the nuclear and plastid genomes, we addressed the following questions: (1) Does genetic grouping correlate with the present taxonomic treatment? (2) Where did perennial species of *Helminthotheca* originate? (3) Along which routes did they spread to occupy their present areas? The inferred evolutionary history is discussed with respect to the evolutionary forces that might have been relevant for speciation.

METHODS

Plant material

We sampled populations belonging to subspecies of *Helminthotheca aculeata* and *H. comosa* in Algeria, Italy (Sicily), Morocco, Portugal and Spain (Table 1, Fig. 1). *Helminthotheca echioides* was included as outgroup. Vouchers of all populations used are deposited in the herbaria SEV and/or WU. The Algerian species *H. balansae* and *H. glomerata* (Quézel and Santa

1963) could not be sampled for this study. In the field, leaves from individual plants were conserved in silica-gel. Distribution ranges of the taxa under study were compiled from survey of material in the herbaria at Barcelona (BC), Geneva (G), Madrid (MA), Montpellier (MPU), Paris (P), Salamanca (SALA) and Seville (SEV). Locations mentioned in Pignatti (1982) and Sauvage (1961) were also used. DIVA-GIS ver. 7.1.7.2 (Hijmans *et al.* 2001) was used to display distribution ranges and sample locations.

AFLP

The protocol for extraction of genomic DNA from silica-gel dried leaves, restriction-ligation (*EcoRI* and *MseI*), preselective and selective amplification, and electrophoresis on a capillary sequencer (3130xl Genetic Analyzer, Applied Biosystems) exactly followed Tremetsberger *et al.* (2009). Three selective primer combinations [*EcoRI*(Fam)-ACA/*MseI*-CAT, *EcoRI*(Vic)-AGG/*MseI*-CTC, and *EcoRI*(Ned)-AAC/*MseI*-CTT] were selected following a primer trial testing twelve selective primer combinations, based on the criteria of interpretability (presence/absence), spread of bands in the region 50-500 bp, and number of bands (approximately 70-90 per primer combination). Raw data were aligned with the GeneScan 500 ROX size standard using ABI PRISM GeneScan ver. 3.71 (© Applied Biosystems). Presence/absence of bands in all individuals and replicates was scored with Genographer ver. 1.6.0 (Benham 2001) after normalizing on total signal in one file. We scored 96 individuals from 25 populations of the perennial species of *Helminthotheca*. Twenty-two individuals were replicated from restriction-ligation and used to calculate the error rate as the ratio between observed number of phenotypic differences and total number of phenotypic comparisons between replicates

(Bonin *et al.* 2004). Three individuals of *H. echioides* were used as outgroup. Only loci observed in the ingroup were scored in the outgroup.

We employed split decomposition (Bandelt and Dress 1992) for constructing a phylogenetic network, which is preferred over a strictly bifurcating tree at the population level, where gene flow is to be expected. FAMD ver. 1.108 (Schlüter and Harris 2006) was used to calculate a matrix of Nei and Li (1979) distances with the length of the restriction enzyme's recognition sequence set to six between all pairs of individuals. This matrix was imported into SplitsTree ver. 4.8 (Huson and Bryant 2006) to construct a NeighborNet network. Splits, whose weight did not exceed a threshold of 0.001, were removed.

Bayesian admixture clustering (Corander *et al.* 2003; Corander and Marttinen 2006) was used to estimate the number of genetically coherent clusters in the perennial species of *Helminthotheca* with BAPS ver. 5.2 (Corander *et al.* 2004, 2008). The maximum number of populations for mixture clustering was set to 20 (5x), 25 (5x), and 30 (5x). For admixture clustering based on results of mixture clustering, the minimum size of a population was set to 1, the number of iterations to 200, and the number of reference individuals to 200 with 20 iterations each.

In order to determine which taxon might be ancestral, we generated a phylogenetic tree, which also included three individuals of the outgroup, *H. echioides*, by neighbour joining analysis based on Nei-Li distances in PAUP* ver. 4.0b10 (Swofford 2003). The support of nodes was estimated by 500 bootstrap replicates. Clusters containing only individuals of the same

population or taxon were collapsed using FigTree ver. 1.3.1 (Rambaut 2006-2009).

In order to distinguish between incongruent species trees, we adopted the approach of Bryant *et al.* (2012) to infer species trees from AFLP data in the framework of the multispecies coalescent, implemented in SNAPP ('SNP and AFLP Phylogenies'), a Bayesian MCMC sampler, which interfaces with the BEAST package (Drummond and Rambaut 2007). Nine individuals per taxon were included in the analysis: *H. aculeata* subsp. *aculeata*, A1-A5; *H. aculeata* subsp. *maroccana*, M1-M3; *H. comosa* subsp. *comosa*, C2-C6; and *H. comosa* subsp. *lusitanica*, S1-S3, S5 and S6. Individuals of putative hybrid origin (see below) were not included. We ran two separate Markov Chain Monte Carlo (MCMC) chains, each with 7 million states, without correction for dominance, in BEAST ver. 2.1.3 (Bouckaert *et al.* 2002-2014). Trees and associated parameter values were logged every 1,000 states. Results of the MCMC runs were analysed with Tracer ver. 1.5.0 (Rambaut and Drummond 2003-2005), discarding the first 10% as burn-in. Trees created in the two MCMC runs were combined with LogCombiner ver. 2.1.3, again discarding the first 10% of each run as burn-in. A species tree was drawn from the combined trees with DensiTree ver. 2.1.11 (Bouckaert 2010).

Within-population genetic diversity was calculated as 'average gene diversity over loci' in Arlequin ver. 3.5.1.2 (Excoffier *et al.* 2005; only for populations with a minimum of three individuals). As indicators of divergence due to long-lasting isolation of populations, we estimated the number of private AFLP bands in a population (i.e., those bands occurring exclusively in the respective population) and the rarity index. Only populations with a

minimum of two individuals were used. Furthermore, some populations with more than two individuals were also omitted, so that a similar number of populations was included for each taxon. The number of private bands was counted with FAMD ver. 1.25. The rarity index equivalent to DW of Schönswetter and Tribsch (2005) was calculated as 'rarity 1' using the R-script AFLPdat (Ehrich 2006) in R ver. 2.11.1 (© The R Foundation for Statistical Computing). For each individual, each AFLP band was divided by the total number of occurrences of this band in the data set. These relative values were then summed to the rarity index for this particular individual. Population values were calculated as averages of individual values.

DNA sequences

The internal transcribed spacer (ITS) region of nuclear ribosomal DNA was amplified by using primers ITS4 and ITS5 of White *et al.* (1990). The PCR mix for amplification (total volume 20 μ l) contained: 18 μ l 1.1 \times Reddy Mix PCR Master Mix (2.5 mM MgCl₂; ABgene), 0.4 μ l forward and reverse primer [20 μ M] each, 0.8 μ l dimethyl sulfoxide (DMSO), and 0.4 μ l DNA extract. Amplification was performed in a Mastercycler (Eppendorf) with the following reaction conditions: 95 °C/4.5 min, 45 °C/1 min, 72 °C/1 min (1 cycle); 95 °C/1 min, 48 °C/1 min, 72 °C/1 min (36 cycles); and 72 °C/7 min, 4 °C/thereafter. Amplification products were purified by enzyme treatment with 0.5 μ l Exonuclease I and 1 μ l Calf Intestine Alkaline Phosphatase (Fermentas; incubation at 37 °C for 45 min followed by enzyme inactivation at 85 °C for 15 min) and cycle sequenced on a GeneAmp PCR System 9700 (Applied Biosystems) by following the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit Protocol (Applied Biosystems). As a

modification, the total volume of the cycle sequencing reaction was 10 μl (1 μl Terminator Ready Reaction Mix, 1.5 μl 5 \times sequencing buffer [10 mM], 1 μl forward or reverse primer [5 μM], and 6.5 μl purified amplification product). Unincorporated dye terminators were removed by centrifugation through MultiScreen-HV filter plates (Millipore) filled with Sephadex G-50 Fine (GE Healthcare) prior to electrophoresis on a 3130xl Genetic Analyzer (Applied Biosystems).

We tested several plastid intergenic spacer regions, namely *atpI-atpH*, *ndhF-rpl32*, *petL-psbE*, *psbD-trnT*, *psbJ-petA*, *rpl32-trnL*, *3'rps16-5'trnK*, *trnQ-5'rps16*, and *3'trnV-ndhC* (Shaw *et al.* 2007) and *psbA-trnH* (Sang *et al.* 1997), for their phylogenetic information content in *Helminthotheca*. The two most variable regions, *ndhF-rpl32* and *rpl32-trnL*, were sequenced in a larger number of individuals. Addition of the remaining regions did not substantially improve phylogenetic resolution. The PCR mix for amplification (total volume 18 μl) contained: 14.4 μl 1.1 \times Reddy Mix PCR Master Mix (2.5 mM MgCl_2 ; ABgene), 0.9 μl forward and reverse primer [10 μM] each, 0.9 μl 0.4% bovine serum albumin (BSA), and 0.9 μl DNA extract. Amplification was performed with the following reaction conditions: 80 $^\circ\text{C}$ /5 min; 95 $^\circ\text{C}$ /30 sec, 50 $^\circ\text{C}$ /30 sec, 65 $^\circ\text{C}$ /4 min (36 cycles); and 65 $^\circ\text{C}$ /5 min, 4 $^\circ\text{C}$ /thereafter. Amplification products were purified by enzyme treatment with 0.5 μl Exonuclease I and 1 μl FastAP Thermosensitive Alkaline Phosphatase (Fermentas; same conditions as for ITS) and cycle sequenced by following the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit Protocol (Applied Biosystems) in a 10 μl total reaction volume (0.6 μl Terminator Ready Reaction Mix, 1.7 μl 5 \times sequencing buffer [10 mM], 1 μl forward or reverse

primer [3.2 μ M], and 6.7 μ l purified amplification product). Products of the cycle sequencing reaction were purified with Sephadex G-50 Fine and run on a capillary sequencer (as for ITS).

Forward and reverse sequences were assembled with SeqMan II ver. 5.05 (DNASTAR, Inc., Madison, WI, USA). The sequences were aligned manually in BioEdit ver. 7.0.5.3 (Hall 1999). SeqScape ver. 2.5.0 (Applied Biosystems) was used to check all mutations in the alignments as well as multiple states (additivities) found in many ITS sequences. Because the peaks of the two bases contained in the multiple states had similar heights, they were interpreted as real polymorphisms due to incomplete concerted evolution and not as sequencing uncertainties. BioEdit was also used to calculate G+C contents of ITS sequences. We also reconstructed the secondary structure of ITS2 of the perennial species of *Helminthotheca*, using the RNA folding form of mfold ver. 3.5 (Zuker 2003), based on the common structure described for Asteraceae by Goertzen *et al.* (2003). Sequences were deposited in the EMBL Nucleotide Sequence Database (Stoesser *et al.* 2003; Table 1).

For phylogenetic analysis of ITS sequences, we adopted the maximum likelihood (ML) approach using the RAxML software on the Vital IT unit of the Swiss Institute of Bioinformatics web server (<http://embnet.vital-it.ch/raxml-bb/index.php>; Stamatakis *et al.* 2008), with individual general time reversible (GTR) models of nucleotide substitution with gamma-distributed site-to-site rate variation and a proportion of invariable sites being estimated for the three partitions ITS1, 5.8S rDNA, and ITS2. Support values were obtained by bootstrapping (100 replicates).

For analysis of concatenated *ndhF-rpl32* and *rpl32-trnL* sequences, the maximum parsimony (MP) approach was adopted due to its ease of incorporating information contained in indels, which were coded using the simple indel coding method implemented in SeqState ver. 1.3.2 (Müller 2005). A heuristic search under the MP criterion was carried out with 1000 random-addition-sequence replicates and TBR swapping with no more than 10 trees of score ≥ 10 retained in each replicate with PAUP* ver. 4.0b10 (Swofford 2003). Bootstrap support was estimated with 500 bootstrap replicates (Felsenstein 1985; heuristic search with 10 random-addition-sequence replicates and TBR swapping with no more than 10 trees of score ≥ 10 retained in each replicate).

A species tree was constructed from ITS and plastid DNA sequences (without indel coding) in BEAST ver. 2.1.3 using the *BEAST template (Heled and Drummond 2010), with a Jukes-Cantor (JC69) site model, a relaxed log normal clock model, a birth-death tree model with constant population size, and 10 million states of the MCMC chain. Individuals of putative hybrid origin (see below) were not included. Populations were assigned to the following taxa: *H. aculeata* subsp. *aculeata*, A1-A5; *H. aculeata* subsp. *maroccana*, M1-M3 (the plastid DNA sequences of two individuals that have the 'comosa/lusitanica'-plastid type [populations M1 and M3; see below] have not been included); *H. comosa* subsp. *comosa*, C1-C6; *H. comosa* subsp. *lusitanica*, S1-S8; and *H. echioides*, E1-E5. Use of other models (e.g., the GTR site model) did not affect the result. Trees were displayed in DensiTree ver. 2.1.11 (10% burn-in).

Biogeographic analysis

Ancestral areas were reconstructed by ML-based biogeographic analysis implemented in the Lagrange version released on 26 May 2013 (Ree and Smith 2008). The species tree generated from DNA sequences was used as input. Taxa were assigned to the geographic areas western North Africa (Rif and Middle Atlas; all taxa except *H. comosa* subsp. *lusitanica*), central North Africa (Tell Atlas; both subspecies of *H. aculeata*), Sicily (*H. aculeata* subsp. *aculeata*), the W Betic region (García-Barros *et al.* 2002; *H. comosa* subsp. *comosa*), and the W Iberian region (García-Barros *et al.* 2002; both subspecies of *H. comosa*). *Helminthotheca echioides* occurs in all areas. The maximum number of areas in ancestral ranges was set to 3. The area adjacency matrix is detailed in Fig. 5A. The dispersal rate was set to 1.0 between adjacent areas and 0.1 between non-adjacent areas.

Molecular clock

We adopted the approach of Tremetsberger *et al.* (2013) for dating the age of the perennial species of *Helminthotheca*. The ITS1 and ITS2 alignments of Cichorieae (Tremetsberger *et al.* 2013; available from TreeBASE) and of *Helminthotheca* were merged. To simplify the calculation, some sequences were removed; totals of retained sequences were 25 for *Helminthotheca* and 46 for Cichorieae and outgroup taxa. BEAST ver. 1.6.2 was used as outlined in Tremetsberger *et al.* (2013) with stem group node calibration. Results of four independent MCMC runs were analysed with Tracer ver. 1.5.0 and combined with LogCombiner ver. 1.6.2 (10% burn-in). The maximum clade credibility tree was obtained using TreeAnnotator ver. 1.6.2 and graphed with FigTree ver. 1.3.1.

RESULTS

AFLP

In total, 231 AFLP bands in the range of 90–430 bp were scored for the 96 individuals of the perennial species of *Helminthotheca*. Of these, 207 (i.e., 90%) are polymorphic. All individuals have unique AFLP phenotypes. The overall error rate was 4% (ranging from 0–41% for individual loci). Removal of seven loci with individual error rates > 25% did not significantly alter the phylogenetic network (Fig. 2A). We thus decided to rely on the original matrix for all further calculations.

In the Bayesian mixture clustering analysis, the first nine best visited partitions revealed four clusters, whereas the tenth best visited partition revealed five clusters. The values for the log of the marginal likelihood of the nine optimal partitions range from -5916 to -5960 (the tenth best visited partition has a value of -5971). The posterior probability for $K = 4$ is 1. The four clusters revealed (Figs. 1, 2) are (1) populations A1-A5 of *H. aculeata* subsp. *aculeata* (Algeria, Sicily), (2) populations M1-M3 of *H. aculeata* subsp. *maroccana* together with populations AM1 and AM2, which morphologically resemble *H. aculeata* subsp. *aculeata*, and population CM, which morphologically resembles *H. comosa* subsp. *comosa* (all Morocco), (3) populations C2-C6 of *H. comosa* subsp. *comosa* (Spain), and (4) populations S1-S8 of *H. comosa* subsp. *lusitanica* (Portugal, Spain) together with population CS (Spain). Some evidence for admixture was also found, namely in some individuals of populations CM and CS, which show admixture from the *H. comosa*-cluster (Fig. 2B).

Network analysis (Fig. 2A) shows relationships among individuals of the four clusters. The major split is across the Strait of Gibraltar separating the *H. comosa* subsp. *comosa*/*H. comosa* subsp. *lusitanica*-cluster from the *H. aculeata* subsp. *maroccana*/*H. aculeata* subsp. *aculeata*-cluster. Within the Afro-Sicilian group, the *H. aculeata* subsp. *maroccana*-populations M1 and M2 from the Middle Atlas branch off basally, whereas the *H. aculeata* subsp. *aculeata*-populations (A1-A5) are derived (Fig. 2C). Within the Iberian group, the *H. comosa* subsp. *comosa*-populations (C2-C6) branch off basally, whereas the *H. comosa* subsp. *lusitanica*-populations (S1-S8) are derived.

Evaluation of sampled trees from the coalescent approach to species tree reconstruction revealed an effective sample size (ESS) of > 10,000 for the two combined runs. The traces did not show signs of non-convergence. All clades have posterior probabilities of 100%, yielding a single consensus tree (Fig. 3A). In a tree lacking short branches, results of analyses with and without correction for dominance do not differ (Bryant *et al.* 2012), so that we relied only on the analysis without correction for dominance. The sequence of divergence suggested by the obtained species tree is: first, divergence of an Afro-Sicilian and an Afro-Iberian lineage; second, divergence of *H. aculeata* subsp. *maroccana* and *H. aculeata* subsp. *aculeata* in the Afro-Sicilian lineage; and third, divergence of *H. comosa* subsp. *comosa* and *H. comosa* subsp. *lusitanica* in the Afro-Iberian lineage.

Accumulation of private and rare AFLP bands in populations due to mutations indicates long-term in-situ persistence of populations. This is in contrast to newly established populations, which should have fewer such bands. The mean number of private bands and the rarity index are slightly

higher in populations of the Afro-Sicilian lineage than in populations of the Afro-Iberian lineage (Table 2). This trend is accentuated, when the estimates are calculated for taxa rather than for populations. *Helminthotheca aculeata* subsp. *aculeata* (populations A1-A5 [$N_{\text{individuals}} = 11$]) has 20 private bands and a rarity index of 4.1. Similarly, *H. aculeata* subsp. *maroccana* (populations M1-M3 [$N_{\text{individuals}} = 13$]) has 19 private bands and a rarity index of 4.1. *Helminthotheca comosa* subsp. *comosa* (populations C3-C6 [$N_{\text{individuals}} = 18$]) has 16 private bands and a rarity index of 2.6, and *H. comosa* subsp. *lusitanica* (populations S2, S5, S7, and S8 [$N_{\text{individuals}} = 20$]) has 12 private bands and rarity index of 2.2. In parallel to estimates of divergence, within-population genetic diversity is also slightly higher in populations of the Afro-Sicilian lineage than in populations of the Afro-Iberian lineage (Table 2).

DNA sequences

The G+C content of ITS1 and ITS2 sequences is very similar among accessions of the perennial species of *Helminthotheca* (mean = 53.3%; SD = 0.2%) and *H. echioides* (mean = 54.0%; SD = 0.3%), suggesting absence of pseudogenes. The ITS2 secondary structure is common to all individuals of *Helminthotheca* analysed [see **SUPPORTING INFORMATION**]. It is also well in accordance with the ITS2 secondary structure of Asteraceae presented by Goertzen *et al.* (2003). No compensatory base changes (Coleman 2007) exist among individuals of *Helminthotheca*, neither when comparisons are made among perennial species nor when these are compared with *H. echioides*.

The ITS tree (Fig. 4A) reveals two groups. Sequences of populations M1-M3 are in basal position together with one sequence of population AM1

and one sequence of population CM. All remaining sequences are in a derived group with 72% bootstrap support. The plastid DNA tree (Fig. 4B) also reveals two main groups. One group with 79% bootstrap support contains sequences of the Afro-Sicilian AFLP group (all but two sequences of populations M1-M3 and all sequences of populations AM1, AM2, and A1-A5). The other group with 81% bootstrap support contains sequences of the Iberian AFLP group (populations C1-C6, CS and S1-S8) together with one sequence of population M1, one sequence of population M3 and sequences of population CM.

Hybrids

Differences in the individual trees derived from the different data sets allow identifying individuals of hybrid origin. Individuals of population CM largely correspond to the 'maroccana'-AFLP group (with two individuals showing admixture from the 'comosa'-AFLP group), one individual has the more basal 'maroccana'- and the other the more derived 'aculeata/comosa/lusitanica'-ribotype, but all individuals have the 'comosa/lusitanica'-plastid type. Similarly, one individual of population M1 and one individual of population M3 also have the 'comosa/lusitanica'-plastid type. One individual of population AM1 and both individuals of population AM2 morphologically resemble *H. aculeata* and have the 'aculeata/comosa/lusitanica'-ribotype, but correspond to the 'maroccana'-AFLP group. Finally, some individuals of population CS show evidence of admixture by Bayesian admixture clustering of AFLP data. When individuals of putative hybrid origin are excluded, Bayesian inference results in a best-

scoring species tree derived from ITS and plastid DNA sequences (Fig. 3B) that is congruent with the species tree derived from AFLP data.

Ancestral area and age estimation

Model-based biogeographic reconstruction of ancestral areas (Fig. 5) gives an idea of the biogeographic history of the perennial species of *Helminthotheca*, albeit with little confidence. Western North Africa with its mountain ranges Rif and Middle Atlas is revealed as the ancestral area of the entire clade as well as of the common ancestor of *H. aculeata* subsp. *aculeata* and *H. aculeata* subsp. *maroccana*. The common ancestor of *H. comosa* subsp. *comosa* and *H. comosa* subsp. *lusitanica* is inferred to have lived in an area spanning W North Africa, the W Betic region, and the W Iberian region. Molecular clock analysis reveals that the minimum age of the MRCA of the perennial clade dates well to the middle of the Pleistocene (median = 1.09 mya; 95% HPD interval = 0.48–1.93 mya) [see **SUPPORTING INFORMATION**].

DISCUSSION

Specific entities

Genomic variation as evidenced by Bayesian mixture clustering of AFLP data (Fig. 2B) shows a basic concordance with morphological variation (S. Talavera et al., unpubl. data). Diagnostic characters for distinguishing *H. comosa* subsp. *lusitanica* from the nominal subspecies are provided by the involucre bracts of the heads. In the former, the subapical appendix of the inner involucre bracts is shorter (1–3.5 mm long in *H. comosa* subsp. *lusitanica* vs. 5–8 mm long in *H. comosa* subsp. *comosa*) and projects less above the bracts (1–2 mm in *H. comosa* subsp. *lusitanica* vs. 2–6 mm in *H.*

comosa subsp. *comosa*). The outer involucral bracts usually have fewer pairs of marginal spines (2–4 in *H. comosa* subsp. *lusitanica* vs. 10–21 in *H. comosa* subsp. *comosa*). Similarly, diagnostic characters for distinguishing *H. aculeata* subsp. *maroccana* from the nominal subspecies are also provided by the involucral bracts of the heads. The outer involucral bracts of *H. aculeata* subsp. *aculeata* usually lack marginal spines. Only rarely do they have one or two pairs of marginal spines in addition to the terminal spine, whereas the outer involucral bracts of *H. aculeata* subsp. *maroccana* have considerably more pairs of marginal spines (10–21).

Following De Queiroz (2007) in treating the existence as separately evolving segments of lineages as the only necessary property of species, we therefore propose to distinguish *H. comosa* subsp. *lusitanica* as a separate species from *H. comosa* and *H. aculeata* subsp. *maroccana* as a separate species from *H. aculeata*. As outlined above, lines of evidence for lineage separation come from AFLP data and morphology. The lineages also have distinct geographic distributions (Fig. 1) and ecological preferences (see below). In this sense, we will refer to *H. comosa* subsp. *lusitanica* as *H. spinosa*, comb. nov., and to *H. aculeata* subsp. *maroccana* as *H. maroccana*, comb. nov., in the following (the new combinations are presented in the section Taxonomic Implications).

North-western African origin

Although ancestral area reconstruction is not quite conclusive, it suggests that the perennial clade of *Helminthotheca* originated in western North Africa, possibly in the Middle Atlas or Rif. A caveat of the biogeographic analysis, however, is that few populations from central North Africa (Tell Atlas;

especially *H. glomerata*) and no populations from the High Atlas and Anti-Atlas have been included, so that we cannot more precisely determine the ancestral area.

Our results for the perennial clade of *Helminthotheca* are well in line with observations in the closely related *Hypochaeris* sect. *Hypochaeris* (Ortiz *et al.* 2009). Both groups originated in western North Africa in the Pleistocene or Pliocene; their origins were thus associated with the ascent of the Mediterranean-type climate and/or Pleistocene climatic oscillations. The importance of the High and Middle Atlas region in addition to the Betic-Rif region for the evolution of Mediterranean plant taxa has been suggested by Médail and Quézel (1997). The High and Middle Atlas Mountains, just as the Rif, are known to have a high rate of endemism, which is explained by the long isolation and high altitude of these massifs, which made them serve as refugium for the Tertiary residual flora. In addition, there is also a large number of schizoendemics that demonstrate the role of the Atlas Mountains in more recent speciation processes (Médail and Quézel 1997). This is in line with climatic reconstructions, which suggest that a Mediterranean-type climate with summer drought was present in northern Africa earlier than in southern Europe (Quézel 1978; Suc 1984). Fauquette *et al.* (1999) describe Pliocene vegetation assemblages resembling the modern thermo-Mediterranean formation for the regions Andalusia (southern Spain), North Africa (Morocco, Algeria, Tunisia), and Sicily, corresponding to those regions inhabited by perennial species of *Helminthotheca*. In accordance with our age estimates, we might hypothesize that after the common ancestor of the perennial clade of *Helminthotheca* had originated in response to the ascent of the

Mediterranean-type climate in northern Africa, it held on during the Pliocene and diversified in the Pleistocene.

Migratory routes out of western North Africa

The evolutionary scenario derived from our data suggests two migratory routes out of western North Africa and across the Mediterranean Sea. The first expansion was to the north, across the Strait of Gibraltar, to the Iberian Peninsula. The second expansion was to the east, along central North African territories, to Sicily and the Apennine Peninsula.

Numerous studies have investigated genetic differentiation between populations on both sides of the Strait of Gibraltar (e.g., Arroyo *et al.* 2008; Escudero *et al.* 2008), with differing outcomes. Whereas the Strait of Gibraltar has been an effective barrier to gene flow in *Carex helodes* after its dispersal from the Iberian Peninsula to North Africa (Escudero *et al.* 2008), extensive gene flow across the Strait of Gibraltar via seed dispersal has been documented in *Calicotome villosa* (Arroyo *et al.* 2008). In the central Mediterranean region, the Strait of Sicily separating North Africa and Sicily has also functioned as an important route of dispersal for plants and animals (e.g., Habel *et al.* 2008; Fernández-Mazuecos and Vargas 2011). Dispersal was probably facilitated by lowered sea water levels during the Pleistocene glacial periods (Thiede 1978). In *Helminthotheca*, seed dispersal could have been effectuated by wind or migratory birds (Finlayson 1992).

Ecogeographic isolation

As in other Mediterranean plant groups of Pliocene/Pleistocene age such as *Antirrhinum* (Vargas *et al.* 2009) and *Reseda* sect. *Glaucoreseda* (Martín-Bravo *et al.* 2010), there is also a clear geographic component to

speciation in the perennial clade of *Helminthotheca*. The biogeographic analysis suggests two lineage divergence events within western North Africa, the first yielding an Afro-Sicilian and an Afro-Iberian lineage and the second yielding *H. maroccana* (as progenitor) and *H. aculeata* (as derivative) within the Afro-Sicilian lineage. The placement of *H. glomerata* in this scenario still has to be determined. The third lineage divergence event is situated in the western Iberian region, yielding *H. comosa* (as progenitor) and *H. spinosa* (as derivative).

Accurate identification of species pairs is crucial for investigating isolating barriers. The limit between the distributions of *H. comosa* and *H. spinosa* closely follows the Guadalquivir river basin, which separates the Mesozoic and Cenozoic Betic Cordillera ('Western Betic Region'; García-Barros *et al.* 2002) inhabited by *H. comosa* and the Paleozoic Iberian Massif ('Western Iberia') inhabited by *H. spinosa*. Characteristics of the two regions such as those associated with the substrate (calcareous in the Betic Cordillera, acidic in the Iberian Massif) are expected to exert divergent selective pressure on populations, resulting in ecogeographic isolation of the diverging lineages (Sobel *et al.* 2010). Ecogeographic isolation might also be hypothesized for the species pair *H. aculeata*/*H. maroccana*, in response to differences in substrate or precipitation. *Helminthotheca aculeata* grows on calcareous substrate of the Cenozoic, whereas *H. maroccana* grows on crystalline rocks and schist. Hybrids exist in contact zones between the regions. For example, some populations of *H. comosa* grow in calcareous enclaves, which exist in the western Iberian region. Hybrids between *H.*

spinosa and *H. comosa* are found precisely in the ecotones between acidic and calcareous soils.

As already supposed by Sauvage (1961), assessment of ploidy levels could shed more light on the issue of hybridization coupled with polyploidization in *Helminthotheca*. To the present day, there are few chromosome counts available, whereby a diploid chromosome number ($2n = 10$) has been reported for *H. comosa* and *H. spinosa* (Fernandes and Queirós 1971; Gallego 1981; Luque 1981). A *H. comosa*-like tetraploid population ($2n = 20$) has been found in northern Morocco (*Talavera et al.* 327/03, SEV237787; S. Talavera, unpubl. data), close to the here analysed hybrid population CM.

Taxonomic implications

As a result of the genetic analyses and the morphological survey of the populations included in this work (S. Talavera et al., unpubl. data), two new combinations and descriptions of three hybrids that are new to science are now necessary.

New combinations:

***Helminthotheca maroccana* (Sauvage) Talavera & Tremetsberger, comb.**

nov.: *Picris aculeata* subsp. *maroccana* Sauvage in Trav. Inst. Sci. Chérifien, Sér. Bot. 22: 202 (1961) \equiv *Helminthotheca aculeata* subsp. *maroccana* (Sauvage) Greuter in Willdenowia 33: 233 (2003). Ind. loc.: “Forêt des Bni-Âbid.” Holotype: herbarium C.S. [Charles Sauvage] no. 14689.

***Helminthotheca spinosa* (DC.) Talavera & Tremetsberger, comb. nov.:**

Helminthia spinosa DC. in Candolle & Lamarck, Fl. Franç., ed. 3, 4: 58 (1805) \equiv *Picris spinosa* (DC.) Poir. in Lamarck, Encyc., Suppl. 4: 408 (1816). Ind.

loc.: “Je décrit cette plante d’après des échantillons originaires des Pyrénées et qui proviennent de l’herbier de Lemonnier.” Lectotype, here designated: G00317181 [G-DC]. The type material that is conserved consists of the upper part of three plants, mounted on two “échantillons” or vouchers. One of the vouchers comprises two plants and the other one plant. In the voucher with two plants, there are also two handwritten labels. On one label, which possibly stems from the herbarium of Lemonnier, it is indicated “*Picris aculeata* Desf.?/*P. sprengeriana* Gaert.?/Pyrénées”, and on the other label, A. P. de Candolle indicated “*Helminthia spinosa* Fl. Fr.”. The other voucher, with a single plant, does not contain any label, but the plant is similar to the other two. We select as lectotype the plant mounted on the right side of the voucher that contains the two labels, the more complete of the two plants. The other plant of this voucher and the plant from the other voucher are isolectotypes.

= *Helminthotheca comosa* subsp. *lusitanica* (Welw. ex Schldl.) P. Silva & Escud. in Bol. Soc. Brot., sér. 2, 60: 156 (1987) [for synonyms see Greuter, Med-Checkl. 2: 240 (2008)].

New hybrids:

***Helminthotheca* × *hispanica* Tremetsberger & Talavera, *hybr. nov.*:**

Helminthotheca comosa (Boiss.) Holub × *H. spinosa* (DC.) Talavera & Tremetsberger

Holotype: SEV237782. Spain, Huelva, Hinojos, “Las Porqueras”, in *Quercus suber* forest, 80 m, 37.29°N-6.42°W, 30 June 2007, S. Talavera, n° 228/07. Isotypes: SEV237783, SEV237784.

Plants with leaves that surround the involucre (outer involucral bracts) similar to those of *H. comosa*, but the subapical appendix of the inner involucral bracts is similar in length to that of *H. spinosa* (5–8 mm long in *H. comosa*; 1–3.5 mm long in *H. spinosa*; 2–3 mm long in *H. x hispanica*).

Helminthotheca x tingitana Tremetsberger & Talavera, hybr. nov.:

Helminthotheca comosa (Boiss.) Holub x *H. maroccana* (Sauvage) Talavera & Tremetsberger

Holotype: SEV256853. Morocco, Tangier-Tétouan, Chefchaouen, pr. Aguelman, carretera de Chefchaouen a Tetuán, 536 m, 35°22'08"N-5°22'38.2"W, 17 June 2008, E. Rico et al., n° MS-999. Isotypes: MA779452, SALA132080.

Plants with the subapical appendix of the inner involucral bracts (6.5–7 mm long) similar to that of *H. comosa*, but the leaves that surround the involucre (outer involucral bracts) have more than 50 spines in the margin, similar to those of *H. maroccana*.

Helminthotheca x riphaea Tremetsberger & Talavera, hybr. nov.:

Helminthotheca aculeata (Vahl) Lack x *H. maroccana* (Sauvage) Talavera & Tremetsberger

Holotype: SEV256845. Morocco, Taza-Al Hoceima-Taounate, pr. Ait Isa, en la carretera de Targuist a Al Hoceima, 1012 m, 35°1'27.2"N-4°10'4"W, 18 June 2008, E. Rico et al., n° MS-1099. Isotypes: MA782913, SALA132081.

Plants morphologically more similar to *H. aculeata* than to *H. maroccana*, but the leaves that surround the involucre (outer involucral bracts)

have 2–7 pairs of marginal spines and some of the peduncles are doliform in the apex, as in *H. aculeata*, and some are cylindrical, as in *H. maroccana*.

CONCLUSIONS

This study demonstrates that phylogeography above the species level is a powerful tool for investigating patterns and processes at the boundary between divergent and reticulate relationships. The western Mediterranean region, and specifically western North Africa, is highlighted as a region of intensive recent speciation. The inferred evolutionary history is compatible with the concept of ecogeographic isolation, which refers to the fact that geographic ranges of diverging lineages are largely non-overlapping due to adaptive differentiation (Sobel et al. 2010).

SUPPORTING INFORMATION

The following [**SUPPORTING INFORMATION**] is available in the online version of this article:

File 1. Figure. Common secondary structure of ITS2 rRNA of *Helminthotheca*.

File 2. Figure. Chronogram of *Helminthotheca* and the two closest outgroup genera, *Leontodon* and *Picris*, based on ITS1 and ITS2 sequences.

ACCESSION NUMBERS

All sequences obtained in this study have been deposited in the EMBL Nucleotide Sequence Database under the accession numbers LN830755–LN830895.

SOURCES OF FUNDING

This work was supported by a Juan de la Cierva fellowship [to K.T.] and projects of the Ministerio de Educación y Ciencia resp. Ministerio de

Ciencia e Innovación resp. Ministerio de Economía y Competitividad [Spain; grant numbers CGL2009-08178, CGL2012-32914, CGL2008-02486-E and CGL2008-02531-E].

CONTRIBUTIONS BY AUTHORS

S.T. conceived the idea of the paper, F.B., R.C.-S., M.Á.O, M.T., S.T., A.T., and K.T. collected plant material, M.T. and S.T. studied herbarium material, A.T. and K.T. performed the analyses, and K.T. led the writing of the manuscript.

CONFLICTS OF INTEREST

No conflicts of interest

ACKNOWLEDGEMENTS

We thank Christiane König (University of Vienna, Austria), Enrique Rico Hernández (Universidad de Salamanca, Spain), Iñigo Sánchez (Zoobotánico Jerez, Spain) and Errol Véla (Université de Montpellier, France) for collecting plant material, Ali Loukkas and Ali Mahmoudi, former directors of National Parks of Theniet-el-Had and Gouraya (Algeria), for their help in the field and courtesy for collecting fresh material, Eric Chambers (Vienna) for proofreading and two anonymous referees for constructive criticism.

LITERATURE CITED

- Arroyo J, Aparicio B, Albaladejo RG, Muñoz J, Braza R. 2008. Genetic structure and population differentiation of the Mediterranean pioneer spiny broom *Calicotome villosa* across the Strait of Gibraltar. *Biological Journal of the Linnean Society* **93**:39–51.
- Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**:489–522.
- Avise JC. 2000. *Phylogeography: The history and formation of species*. Cambridge, MA: Harvard University Press.
- Bandelt H-J, Dress AWM. 1992. A canonical decomposition theory for metrics on a finite set. *Advances in Mathematics* **92**:47–105.
- Benham JJ. 2001. Genographer, ver. 1.6.0. Bozeman: Montana State University.
- Bocquet G, Widler B, Kiefer H. 1978. The Messinian Model - A new outlook for the floristics and systematics of the Mediterranean area. *Candollea* **33**:269–287.
- 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.
- Bouckaert RR, Drummond AJ, Rambaut A, Suchard MA. 2002-2014. BEAST, ver. 2.1.3, available from <http://beast2.cs.auckland.ac.nz/>.
- Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* **26**:1372–1373.

- Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution* **29**:1917–1932.
- Coleman AW. 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Research* **35**:3322–3329.
- Corander J, Waldmann P, Sillanpää MJ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* **163**:367–374.
- Corander J, Waldmann P, Marttinen P, Sillanpää MJ. 2004. BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* **20**:2363–2369.
- Corander J, Marttinen P. 2006. Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology* **15**:2833–2843.
- Corander J, Marttinen P, Sirén J, Tang J. 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9**:539.
- De Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* **56**:879–886.
- Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**:214.
- Ehrich D. 2006. AFLPdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes* **6**:603–604.
- Escudero M, Vargas P, Valcárcel V, Luceño M. 2008. Strait of Gibraltar: an effective gene-flow barrier for wind-pollinated *Carex helodes*

- (Cyperaceae) as revealed by DNA sequences, AFLP, and cytogenetic variation. *American Journal of Botany* **95**:745–755.
- Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**:47–50.
- Fauquette S, Suc J-P, Guiot J, Diniz F, Feddi N, Zheng Z, Bessais E, Drivaliari A. 1999. Climate and biomes in the West Mediterranean area during the Pliocene. *Palaeogeography, Palaeoclimatology, Palaeoecology* **152**:15–36.
- Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791.
- Fernandes A, Queirós M. 1971. Contribution à la connaissance cytotaxinomique des *Spermatophyta* du Portugal. II. Compositae. *Boletim da Sociedade Broteriana, 2.^a Série*, **45**:5–121.
- Fernández-Mazuecos M, Vargas P. 2011. Historical isolation versus recent long-distance connections between Europe and Africa in bifid toadflaxes (*Linaria* sect. *Versicolores*). *PLoS One* **6**:e22234.
- Finlayson C. 1992. *Birds of the Strait of Gibraltar*. London: T. & A. D. Poyser.
- Fiz-Palacios O, Valcárcel V. 2013. From Messinian crisis to Mediterranean climate: A temporal gap of diversification recovered from multiple plant phylogenies. *Perspectives in Plant Ecology, Evolution and Systematics* **15**:130–137.
- Gallego MJ. 1981. Números cromosómicos para la flora Española. Números 197-200. *Lagasalia* **10**:231–232.

- García-Barros E, Gurrea P, Luciáñez MJ, Cano JM, Munguira ML, Moreno JC, Sainz H, Sanz MJ, Simón JC. 2002. Parsimony analysis of endemism and its application to animal and plant geographical distributions in the Ibero-Balearic region (western Mediterranean). *Journal of Biogeography* **29**:109–124.
- Goertzen LR, Cannone JJ, Gutell RR, Jansen RK. 2003. ITS secondary structure derived from comparative analysis: implications for sequence alignment and phylogeny of the Asteraceae. *Molecular Phylogenetics and Evolution* **29**:216–234.
- Habel JC, Meyer M, El Mousadik A, Schmitt T. 2008. Africa goes Europe: The complete phylogeography of the marbled white butterfly species complex *Melanargia galathea*/*M. lachesis* (Lepidoptera: Satyridae). *Organisms, Diversity & Evolution* **8**:121–129.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**:95–98.
- Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27**:570–580.
- Hijmans RJ, Guarino L, Cruz M, Rojas E. 2001. Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter* **127**:15–19.
- Hsü KJ, Ryan WBF, Cita MB. 1973. Late Miocene desiccation of the Mediterranean. *Nature* **242**:240–244.
- Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**:254–267.

- Lack HW. 1974. *Die Gattung Picris L., sensu lato, im ostmediterranean westasiatischen Raum*. Dissertationen der Universität Wien, Band 116. Wien: Verband der wissenschaftlichen Gesellschaften Österreichs (VWGÖ).
- Lack HW. 1975. A note on *Helminthotheca* Zinn (Compositae). *Taxon* **24**:111–112.
- Lonergan L, White N. 1997. Origin of the Betic-Rif mountain belt. *Tectonics* **16**:504–522.
- Luque T. 1981. Números cromosómicos para la flora Española. Números 210-213. *Lagascalia* **10**:236–237.
- Martín-Bravo S, Valcárcel V, Vargas P, Luceño M. 2010. Geographical speciation related to Pleistocene range shifts in the western Mediterranean mountains (*Reseda* sect. *Glaucoseda*, Resedaceae). *Taxon* **59**:466–482.
- Médail F, Quézel P. 1997. Hot-spots analysis for conservation of plant biodiversity in the Mediterranean basin. *Annals of the Missouri Botanical Garden* **84**:112–127.
- Médail F, Diadema K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography* **36**:1333–1345.
- Müller K. 2005. SeqState: primer design and sequence statistics for phylogenetic DNA datasets. *Applied Bioinformatics* **4**:65–69.
- Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA* **76**:5269–5273.

- Ortiz MÁ, Tremetsberger K, Stuessy TF, Terrab A, García-Castaño JL, Talavera S. 2009. Phylogeographic patterns in *Hypochaeris* section *Hypochaeris* (Asteraceae, Lactuceae) of the western Mediterranean. *Journal of Biogeography* **36**:1384–1397.
- Pignatti S. 1982. *Flora d'Italia. Volume 3*. Bologna: Edagricole.
- Quézel P, Santa S. 1963. *Nouvelle Flore de l'Algérie et des régions désertiques méridionales*. Paris: Centre National de la Recherche Scientifique.
- Quézel P. 1978. Analysis of the flora of Mediterranean and Saharan Africa. *Annals of the Missouri Botanical Garden* **65**:479–534.
- Rambaut A, Drummond AJ. 2003-2005. Tracer, ver. 1.3, available from <http://evolve.zoo.ox.ac.uk/beast/>.
- Rambaut A. 2006-2009. FigTree, ver. 1.3.1, available from <http://tree.bio.ed.ac.uk/software/figtree/>.
- Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* **57**:4–14.
- Sang T, Crawford DJ, Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**:1120–1136.
- Sauvage C. 1961. Flore des subéeraies marocaines. Catalogues des cryptogames vasculaires et des phanerogames. *Travaux de l'Institut Scientifique Chérifien, Série Botanique* **22**:1–252.
- Schlüter PM, Harris SA. 2006. Analysis of multilocus fingerprinting data sets containing missing data. *Molecular Ecology Notes* **6**:569–572.

- Schönswetter P, Tribsch A. 2005. Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L. (Apiaceae). *Taxon* **54**:725–732.
- Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94**:275–288.
- Sobel JM, Chen GF, Watt LR, Schemske DW. 2010. The biology of speciation. *Evolution* **64**:295–315.
- Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web-Servers. *Systematic Biology* **75**:758–771.
- Stoesser G, Baker W, Van den Broek A, Garcia-Pastor M, Kanz C, Kulikova T, Leinonen R, Lin Q, Lombard V, Lopez R, Mancuso R, Nardone F, Stoehr P, Tuli M, Tzouvara K, Vaughan R. 2003. The EMBL Nucleotide Sequence Database: major new developments. *Nucleic Acids Research* **31**:17–22.
- Suc JP. 1984. Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature* **307**:429–432.
- Swofford DL. 2003. PAUP*: Phylogenetic analyses using parsimony (*and other methods), ver. 4.0b10. Sunderland: Sinauer.
- Thiede J. 1978. A Glacial Mediterranean. *Nature* **276**:680–683.
- Tremetsberger K, Urtubey E, Terrab A, Baeza CM, Ortiz MÁ, Talavera M, König C, Tensch EM, Kohl G, Talavera S, Stuessy TF. 2009. Pleistocene refugia and polytopic replacement of diploids by tetraploids in the Patagonian and Subantarctic plant *Hypochaeris incana* (Asteraceae, Cichorieae). *Molecular Ecology* **18**:3668–3682.

- Tremetsberger K, Gemeinholzer B, Zetsche H, Blackmore S, Kilian N, Talavera S. 2013. Divergence time estimation in Cichorieae (Asteraceae) using a fossil-calibrated relaxed molecular clock. *Organisms, Diversity & Evolution* **13**:1–13.
- Vargas P, Carrió E, Guzmán B, Amat E, Güemes J. 2009. A geographical pattern of *Antirrhinum* (Scrophulariaceae) speciation since the Pliocene based on plastid and nuclear DNA polymorphisms. *Journal of Biogeography* **36**:1297–1312.
- White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. *PCR protocols: a guide to methods and applications*. San Diego: Academic Press, 315–322.
- Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**:3406–3415.

FIGURE LEGENDS

Figure 1. Distribution ranges of *H. aculeata* subsp. *aculeata*, *H. aculeata* subsp. *maroccana* (*H. maroccana*, comb. nov.), *H. comosa* subsp. *comosa*, and *H. comosa* subsp. *lusitanica* (*H. spinosa*, comb. nov.) and populations analysed in this study (colour of triangles refers to belonging of accessions to one of four clusters inferred by Bayesian mixture clustering of AFLP data).

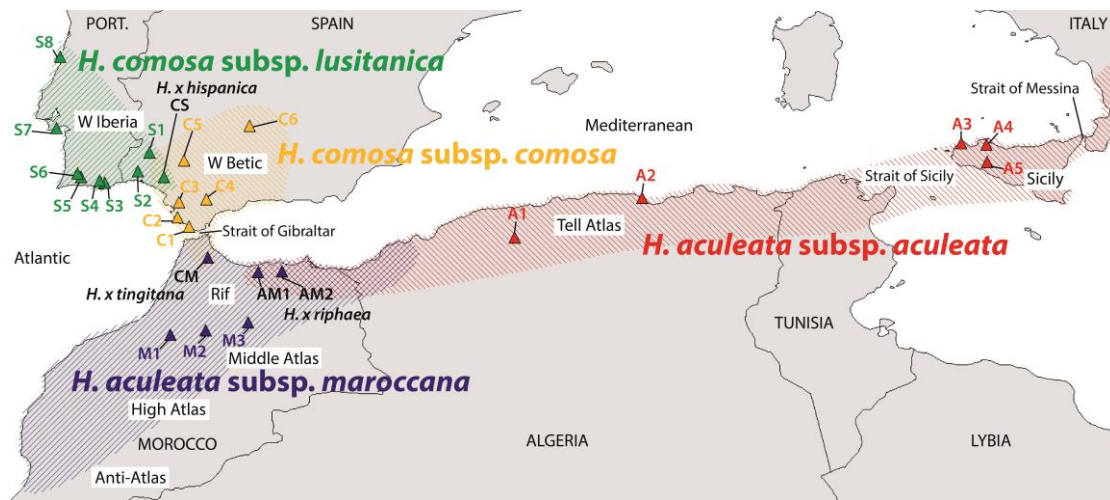
Figure 2. Results of AFLP analysis showing division of population samples of perennial species of *Helminthotheca* into four clusters. A, Phylogenetic network (colour coding according to Bayesian mixture clustering). B, Bayesian admixture clustering with evidence for admixture in populations CM and CS. C, Neighbour-joining tree including the outgroup, *H. echioides* (colour coding according to Bayesian mixture clustering). Numbers above branches indicate bootstrap support values >50%.

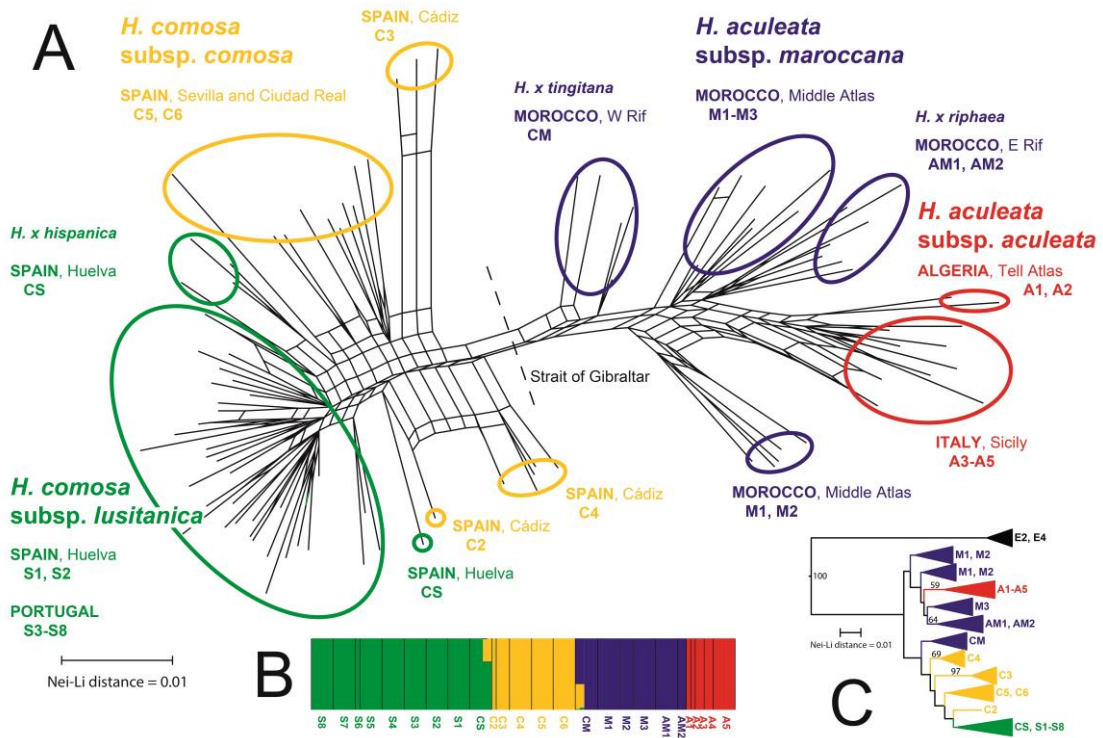
Figure 3. DensiTree of the perennial clade of *Helminthotheca* showing the complete tree set (thin lines) and consensus trees (thick lines; 10% burn-in) from (A) SNAPP analysis of AFLP data (single consensus tree) and (B) *BEAST analysis of ITS and plastid DNA sequences. The posterior probability of the *H. aculeata* subsp. *maroccana*/*H. aculeata* subsp. *aculeata*-clade is 58%. With coded indels of plastid DNA sequences included, it increases to 92% (not shown).

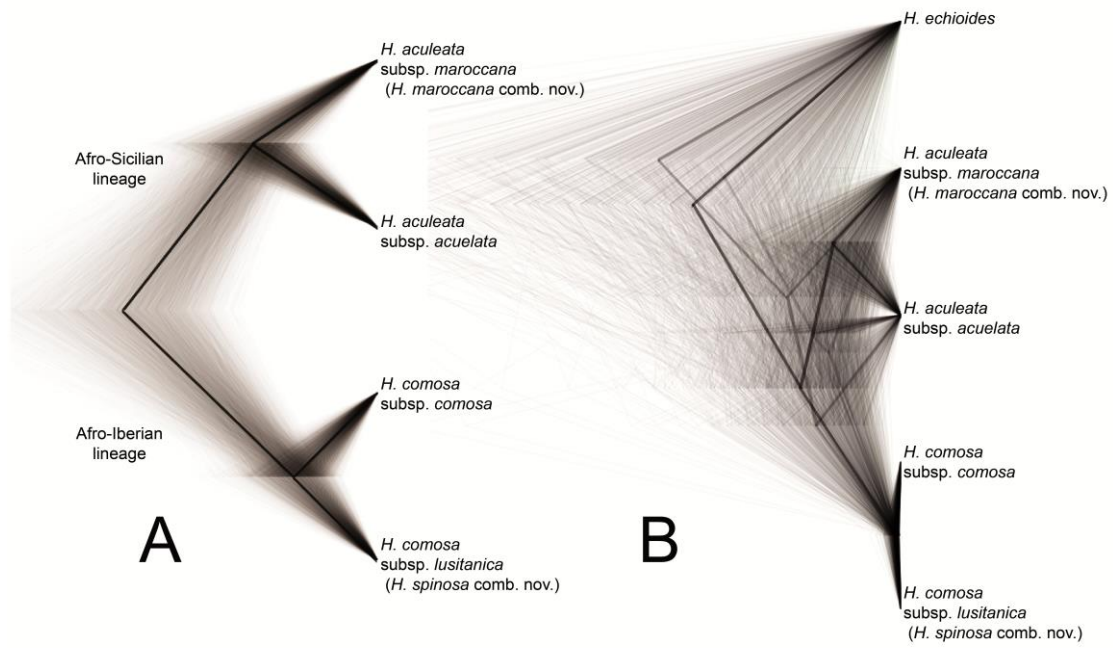
Figure 4. Phylogenetic trees of perennial species of *Helminthotheca*, with *H. echioides* designated as outgroup. A, Best-scoring ML tree of ITS1, 5.8S rDNA, and ITS2 sequences. B, Bootstrap 50% majority rule consensus tree (plus other groups compatible with it) based on MP analysis of concatenated

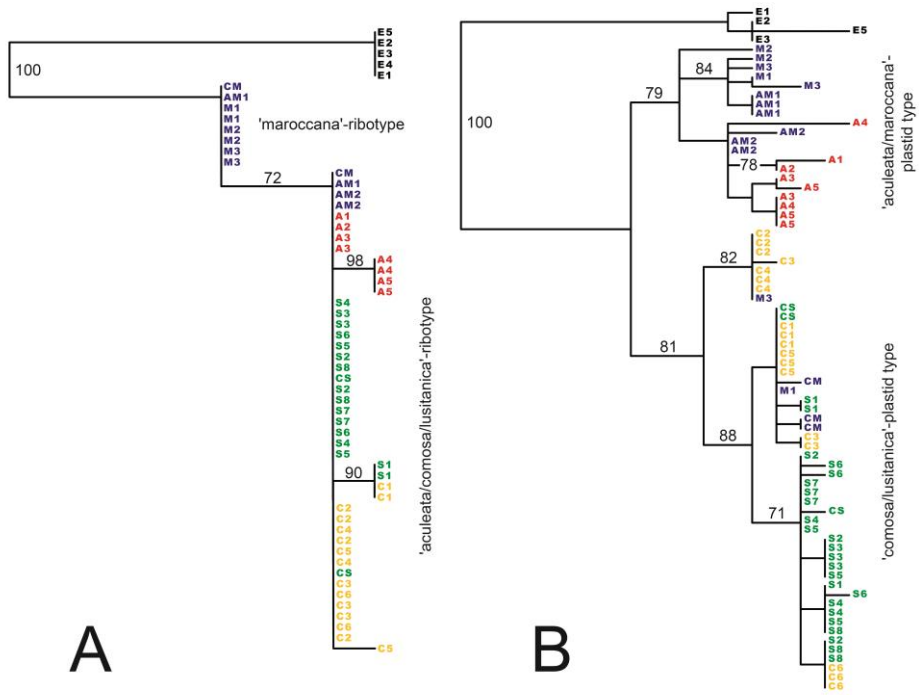
ndhF-rpl32 and *rpl32-trnL* sequences. Numbers above branches indicate bootstrap support values >70%. Colour coding refers to belonging of individuals to one of four groups obtained by Bayesian mixture clustering of AFLP data.

Figure 5. Model-based ancestral area reconstruction in the perennial clade of *Helminthotheca* using the software Lagrange (Ree and Smith 2008). Global -lnL at root node is 9.405, the estimated dispersal rate is 1.124, and the estimated local extinction rate is 4.285e-09. A, Area adjacency matrix used for input. B, Species tree (consensus tree of *BEAST analysis) displaying most probable ancestral areas. C, Splits with associated log likelihoods (lnL) and relative probabilities (Rel. Prob.) for each node.







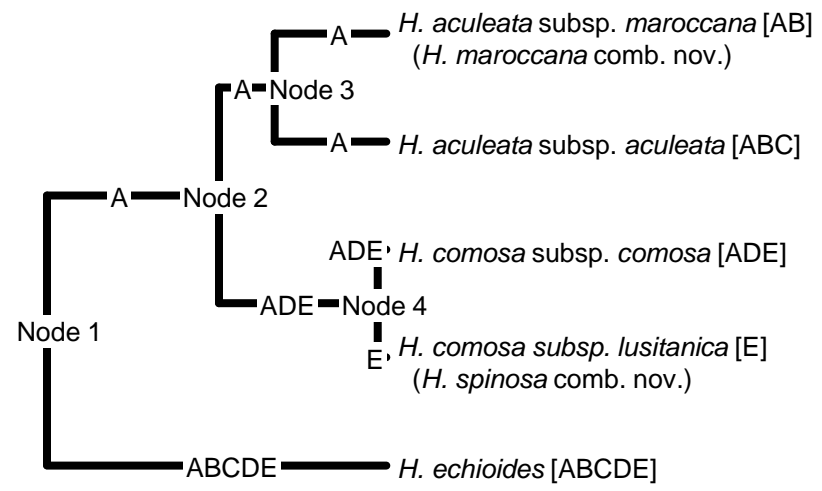


A

Areas	
A	Western North Africa (Rif and Middle Atlas)
B	Central North Africa (Tell Atlas)
C	Sicily
D	Western Betic Region (Betic Cordillera)
E	Western Iberian Region (Iberian Massif)

	A	B	C	D	E
A	—	adjacent	non-adjacent	adjacent	non-adjacent
B		—	adjacent	non-adjacent	non-adjacent
C			—	non-adjacent	non-adjacent
D				—	adjacent
E					—

B



C

Node	Split	InL	Rel. Prob.	
1	[A BCDE]	-10.59	0.3057	
	[D ABCE]	-10.66	0.2843	
	[E ABCD]	-11.17	0.1714	
	[B ABCDE]	-11.19	0.1671	
	[C ABCDE]	-12.04	0.0714	
2	[A ADE]	-10.83	0.2406	
	[A DE]	-11.03	0.1963	
	[B ADE]	-11.23	0.1611	
	[AB DE]	-11.44	0.1313	
	[ABC A]	-11.66	0.1051	
	[BC A]	-11.97	0.0769	
	[A D]	-13.25	0.0214	
	[A AD]	-13.34	0.0196	
	3	[A A]	-10.94	0.2153
		[ABC B]	-11.61	0.1104
[ABC A]		-11.61	0.1104	
[B B]		-11.82	0.0892	
[AB B]		-12.00	0.0748	
[AB A]		-12.00	0.0748	
[BC A]		-12.02	0.0733	
4	[B A]	-12.19	0.0615	
	[B AB]	-12.29	0.0560	
	[BC B]	-12.62	0.0400	
	[C AB]	-12.98	0.0281	
	[A B]	-13.06	0.0259	
	[ADE E]	-9.49	0.9151	
	[AD E]	-12.04	0.0713	

Table 1. Population samples of *Helminthotheca* used in this study with EMBL accession numbers of DNA sequences

Taxon with		EMBL accession numbers		
population	Locality, collector(s) and number	ITS	<i>ndhF-rpl32</i>	<i>rpl32-trnL</i>
code				
<i>Helminthotheca aculeata</i> (Vahl) Lack subsp. <i>aculeata</i>				
Pop. A1	Algeria, Tissemsilt Province (Tell Atlas): Theniet-el-Had National Park (35.85° N, 1.98° E), <i>Véla s/n</i> (9 June 2006)	LN830805	LN830890	LN830846
Pop. A2	Algeria, Béjaïa Province (Tell Atlas): Gouraya National Park, <i>Véla s/n</i> (2 July 2008)	LN830806	LN830891	LN830847
Pop. A3	Italy, Sicilian Region: Monte Cofano NW of Custonaci (38.10° N, 12.67° E), <i>Ortiz &</i> <i>Tremetsberger 24/08</i>	LN830807 (two identical individuals)	LN830892 (two identical individuals)	LN830848, LN830849
Pop. A4	Italy, Sicilian Region: Castellaccio W of Monreale (38.08° N, 13.27° E), <i>Ortiz &</i> <i>Tremetsberger 27/08</i>	LN830808, LN830809	LN830893, LN830894	LN830850, LN830851

Pop. A5	Italy, Sicilian Region: Chiusa Sclafani (37.66° N, 13.28° E), <i>Ortiz & Tremetsberger 30/08</i>	LN830810, LN830811	LN830895 (three identical individuals)	LN830852, LN830853 (two identical individuals)
<i>Helminthotheca aculeata</i> subsp. <i>maroccana</i> (Sauvage) Greuter [<i>H. maroccana</i> (Sauvage) Talavera & Tremetsberger, comb. nov.]				
Pop. M1	Morocco, Rabat-Salé-Zemmour-Zaer Region (Middle Atlas): Montes de Zaïan between Tiddas and Oulmes (33.52° N, 6.26° W), <i>Talavera et al. 171/06</i>	LN830799, LN830800	LN830885, LN830886	LN830842, LN830843
Pop. M2	Morocco, Meknès-Tafilalet Region (Middle Atlas): S of Bou Fekrane on road to Jemaa de Mrirt (33.62° N, 5.43° W), <i>Talavera et al. 796/05</i>	LN830797, LN830798	LN830885 (two identical individuals)	LN830840, LN830841
Pop. M3	Morocco, Fès-Boulemane Region (Middle Atlas): Ribate El Kheir (33.82° N, 4.41° W), <i>Talavera et al. 37/08</i>	LN830795, LN830796	LN830882, LN830883, LN830884	LN830837, LN830838, LN830839
<i>H. comosa</i> (Boiss.) Holub subsp. <i>comosa</i>				
Pop. C1	Spain, Cádiz Province (W Betic): Zahara de los	LN830782,	LN830875 (three	LN830831 (three

	Atunes (36.11° N, 5.82° W), <i>König s/n</i> (22 May 2008)	LN830783	identical individuals)	identical individuals)
Pop. C2	Spain, Cádiz Province (W Betic): Conil de la Frontera (36.33° N, 6.09° W), <i>Talavera et al. 11/08</i>	LN830789, LN830790, LN830791, LN830792	LN830879 (three identical individuals)	LN830835 (three identical individuals)
Pop. C3	Spain, Cádiz Province (W Betic): Jerez de la Frontera (36.69° N, 6.06° W), <i>Sánchez s/n</i> (30 July 2007)	LN830786, LN830787, LN830788	LN830877 (two identical individuals), LN830878	LN830833 (two identical individuals), LN830834
Pop. C4	Spain, Cádiz Province (W Betic): Grazalema (36.76° N, 5.40° W), <i>Talavera & Talavera 112/08</i>	LN830784, LN830785	LN830876 (three identical individuals)	LN830832 (three identical individuals)
Pop. C5	Spain, Seville Province (W Iberia): Castilblanco de los Arroyos (37.70° N, 5.94° W), <i>Talavera et al. 7/06</i>	LN830778, LN830779	LN830873 (three identical individuals)	LN830829 (three identical individuals)
Pop. C6	Spain, Ciudad Real Province (W Iberia)::	LN830780,	LN830874 (three	LN830830 (three

Puertollano (38.52° N, 4.37° W), *Talavera et al.* LN830781 identical individuals) identical individuals)
205/06

H. comosa subsp. *lusitanica* (Welw. ex Schltld.) P. Silva & Escud. [*H. spinosa* (DC.) Talavera & Tremetsberger, comb. nov.]

Pop. S1	Spain, Huelva Province (W Iberia): Almonaster la Real (37.88° N, 6.77° W), <i>Talavera et al.</i> 324/06	LN830774, LN830775	LN830870 (two identical individuals), LN830871	LN830825 (three identical individuals)
Pop. S2	Spain, Huelva Province (W Iberia): between Gibraleón and San Bartolomé de la Torre (37.43° N, 7.05° W), <i>Tremetsberger & Ortiz</i> 1/06	LN830776, LN830777	LN830872 (three identical individuals)	LN830826, LN830827, LN830828
Pop. S3	Portugal, Faro District (W Iberia): between Sao Bras de Alportel and Santa Catarina (37.15° N, 7.83° W), <i>Talavera et al.</i> 214/06	LN830760, LN830761	LN830858 (three identical individuals)	LN830816 (three identical individuals)
Pop. S4	Portugal, Faro District (W Iberia): between Querenca and Barranco do Velho (37.21° N, 7.96° W), <i>Talavera et al.</i> 219/06	LN830770, LN830771	LN830866, LN830867 (two identical individuals)	LN830823 (three identical individuals)
Pop. S5	Portugal, Faro District (W Iberia): between	LN830762,	LN830859 (two	LN830817 (two

	Silves and Sao Marcos da Serra (37.30° N, 8.40° W), <i>Talavera et al. 224/06</i>	LN830763	identical individuals), LN830860	identical individuals), LN830818
Pop. S6	Portugal, Beja District (W Iberia): Serra de Monchique between Monchique and Odemira (37.39° N, 8.49° W), <i>Talavera et al. 244/06</i>	LN830764, LN830765	LN830861 (two identical individuals), LN830862	LN830819 (three identical individuals)
Pop. S7	Portugal, Setubal District (W Iberia): Serra da Arrábida between Setubal and Portinho da Arrábida (38.47° N, 9.00° W), <i>Talavera et al. 262/06</i>	LN830766, LN830767	LN830863 (three identical individuals)	LN830820 (three identical individuals)
Pop. S8	Portugal, Coimbra District (W Iberia): Figueira da Foz, Cabo Mondego (40.18° N, 8.90° W), <i>Talavera et al. 263/06</i>	LN830768, LN830769	LN830864 (two identical individuals), LN830865	LN830821 (two identical individuals), LN830822
<i>Helminthotheca echioides</i> (L.) Holub				
Pop. E1	Portugal, Coimbra District: Figueira da Foz, Cabo Mondego (40.18° N, 8.90° W), <i>Talavera et al. 266/06</i>	LN830755	LN830854	LN830812

Pop. E2	Spain, Huelva Province: between Gibrleón and Cartaya (37.33° N, 7.03° W), <i>Tremetsberger & Ortiz 2/06</i>	LN830756	LN830855	LN830813
Pop. E3	Spain, Cádiz Province: El Palmar (36.32° N, 6.08° W), <i>Talavera & Ortiz s/n</i> (26 April 2006)	LN830757	LN830856	LN830814
Pop. E4	Morocco, Tangier-Tetouan Region: S of Ksar es-Seghir (35.77° N, 5.53° W), <i>Talavera et al. 23/07</i>	LN830758 (four identical individuals)	–	–
Pop. E5	Algeria, Annaba Province: El Kerma on road Annaba-Guelma (36.70° N, 7.62° E), <i>Véla s/n</i> (5 June 2006)	LN830759	LN830857	LN830815
<i>Helminthotheca x hispanica</i> Tremetsberger & Talavera, hybr. nov.				
Pop. CS	Spain, Huelva Province (W Iberia): Hinojos (37.29° N, 6.42° W), <i>Talavera 228/07</i>	LN830772, LN830773	LN830868, LN830869 (two identical individuals)	LN830824 (three identical individuals)

Helminthotheca x riphaea Tremetsberger & Talavera, hybr. nov.

Pop. AM1	Morocco, Taza-Al Hoceima-Taounate Region (Rif): between Targuist and Al Hoceima (35.02° N, 4.17° W), <i>Rico et al. MS-1099</i>	LN830801, LN830802	LN830883 (three identical individuals)	LN830844 (three identical individuals)
Pop. AM2	Morocco, Taza-Al Hoceima-Taounate Region (Rif): between Tafersit and Temsaman (35.04° N, 3.59° W), <i>Rico et al. MS-1145</i>	LN830803, LN830804	LN830884 (three identical individuals)	LN830845 (three identical individuals)
<i>Helminthotheca x tingitana</i> Tremetsberger & Talavera, hybr. nov.				
Pop. CM	Morocco, Tangier-Tetouan Region (Rif): between Chefchaouen and Tetouan (35.37° N, 5.38° W), <i>Rico et al. MS-999</i>	LN830793, LN830794	LN830885 (two identical individuals), LN830884	LN830836 (three identical individuals)

Table 2. Estimates of divergence and diversity derived from AFLP data in populations of *Helminthotheca*. *N*, sample size; ND, not determined

Species and region	<i>N</i>	Estimates of divergence		Within-population genetic diversity
		<i>N</i> _{private bands}	Rarity index	
<i>H. aculeata</i> (Vahl) Lack subsp. <i>aculeata</i> , Algeria and Sicily				
A3	2	2	3.5	ND
A4	2	2	3.0	ND
A5	5	6	3.3	0.099
Mean ± SD		3.3 ± 2.3	3.3 ± 0.3	0.099
<i>H. × riphaea</i> Tremetsberger & Talavera, hybr. nov., Morocco				
AM1	5	3	2.9	0.098
AM2	2	9	7.4	ND
Mean ± SD		6.0 ± 4.2	5.2 ± 3.2	0.098
<i>H. aculeata</i> subsp. <i>maroccana</i> (Sauvage) Greuter				
[<i>H. maroccana</i> (Sauvage) Talavera & Tremetsberger, comb. nov.], Morocco				
M1	5	1	2.3	0.113
M2	3	4	3.3	0.115
M3	5	4	4.3	0.121
Mean ± SD		3.0 ± 1.7	3.3 ± 1.0	0.116 ± 0.004
<i>H. × tingitana</i> Tremetsberger & Talavera, hybr. nov., Morocco				
CM	5	1	2.5	0.109
<i>H. comosa</i> (Boiss.) Holub subsp. <i>comosa</i> , Spain				
C3	3	7	4.5	0.069
C4	5	3	2.3	0.068

C5	5	0	1.4	0.104
C6	5	3	2.1	0.097
Mean ± SD		3.3 ± 2.9	2.6 ± 1.3	0.084 ± 0.019

H. x hispanica Tremetsberger & Talavera, hybr. nov., Spain

CS	5	ND	ND	0.100
----	---	----	----	-------

H. comosa subsp. *lusitanica* (Welw. ex Schldl.) P. Silva & Escud.

[*H. spinosa* (DC.) Talavera & Tremetsberger, comb. nov.], Spain and Portugal

S1	5	ND	ND	0.103
S2	5	1	1.9	0.096
S3	5	ND	ND	0.080
S4	5	ND	ND	0.086
S5	5	5	2.7	0.097
S7	5	0	1.7	0.055
S8	5	3	2.2	0.083
Mean ± SD		2.3 ± 2.2	2.1 ± 0.5	0.087 ± 0.015
