

1 Balearic insular isolation and large continental spread framed the phylogeography of the
2 western Mediterranean *Cheirolophus intybaceus* s.l. (Asteraceae)

3

4 Teresa Garnatje^{1,*}, Ernesto Pérez-Collazos^{2,*}, Jaume Pellicer³ and Pilar Catalán²

5

6 ¹Institut Botànic de Barcelona (IBB-CSIC-ICUB). Passeig del Migdia s.n. 08038
7 Barcelona, Catalonia, Spain.

8 ²Departamento de Agricultura y Economía Agraria, Escuela Politécnica Superior de
9 Huesca, Universidad de Zaragoza, Carretera de Cuarte Km 1, 22071 Huesca, Spain.

10 ³Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB,
11 United Kingdom.

12

13 * Both authors contributed equally to this work

14

15 **Corresponding author:** Ernesto Pérez-Collazos. **E-mail:** ernextop@unizar.es

16 Escuela Politécnica Superior de Huesca, Universidad de Zaragoza, Carretera de Cuarte
17 Km 1, 22071 Huesca, Spain. Telephone: +34(974239301 ext 3072)

18

19 **Running title:** Phylogeography of *Cheirolophus intybaceus*

20

21 **Key words:** AFLPs, Balearic Islands, Dianic range, Ibero-Provençal region, genetic
22 diversity and structure.

23

24 8055 words including references (max. 9000)

25

1 **Abstract**

2 Recent Quaternary geological and climatic events have shaped the evolutionary
3 histories of plant species in the Mediterranean basin, one of the most important hotspots
4 of biodiversity. Genetic analyses of the western Mediterranean *Cheirolophus intybaceus*
5 *s.l.* (Asteraceae) based on AFLP were conducted aiming to establish the relationships
6 between its close species and populations, to reconstruct the phylogeography of the
7 group and to analyse potential unidirectional versus bidirectional dispersals between the
8 Ibero-Provençal belt and the Balearic islands. AFLP data revealed two main genetic
9 groups, one constituted by the Balearic populations and Garraf (NE Iberia) and the other
10 formed by the remaining mainland populations that were further substructured into two
11 geographically separated subgroups (SE+E Iberia and NE Iberia+SW France). Genetic
12 diversity and spatial structure analyses suggested a mid-Pleistocene scenario for the
13 origin of *C. intybaceus* in southern Iberia followed by dispersal to the north and a single
14 colonisation event of the Balearic archipelago from the near Dianic NE Iberian area.
15 This hypothesis was supported by paleogeographical data, which showed the existence
16 of terrestrial connections between the continent and the islands during the middle-late
17 Pleistocene marine regressions, whereas the more recent single back-colonisation of the
18 mainland from Mallorca might be explained by several hypotheses such as long-
19 distance dispersal mediated by migratory marine birds or sea currents.

20

21

22

23

24

25

1 **Introduction**

2

3 The western Mediterranean region experienced drastic geological and climatic
4 changes during the late Tertiary and Quaternary ages (Kadereit & Comes 2005). Those
5 changes influenced the present rich floral hot-spots of the area (Médail & Quézel 1997).
6 The particular and complex paleogeographical configuration of the Ibero-Provençal
7 region (eastern Iberian Peninsula, Balearic Isles and the South of France), together with
8 its history of climatic oscillations and oceanic influences since the Oligocene, modelled
9 a variety of areas that resulted in refugia for a wide range of plants (Médail & Diadema
10 2009). This has had consequences not only on the floral diversity observed today but
11 also on the high genetic diversities of populations of the taxa growing in this western
12 Mediterranean region (Jakob *et al.* 2007; Mansion *et al.* 2008). The few available
13 paleoenvironmental and evolutionary studies indicate that both (i) Tertiary vicariance of
14 some western Mediterranean endemic plants, suggesting an ancient wider range of
15 distribution of the species (Mansion *et al.* 2008), and (ii) Quaternary long distance
16 dispersals (Kropf *et al.* 2006; Kadereit *et al.* 2005; Kadereit & Yaprak 2008) have
17 played an important role in the speciation of the flora.

18 Insular isolation has also contributed to foster the speciation processes of several
19 plant lineages in the Balearic archipelago (Cardona & Contandriopoulos 1979;
20 Contandriopoulos & Cardona 1984). During the Oligocene, the southern Iberian Betic
21 mountains, the eastern Iberian Dianic range (cap de la Nau and surrounding areas of the
22 province of Alicante and the Southern border of Valencia), the Balearic isles and
23 Provence formed a continuous geological territory that split into several microplates
24 (Cohen 1980). In the late Oligocene (30-28 Ma) the Balearic microplate separated from
25 the eastern proto-Iberian peninsula (Cohen 1980; Rosenbaum *et al.* 2002). During the

1 Messinian drought and salinity crisis of the late Miocene (c. 6-5 Ma), the Balearic
2 islands formed a single land mass (Gautier *et al.* 1994) and land corridors were re-
3 established between its western side and the mainland Dianic range, presently extended
4 across the province of Alicante in the eastern Iberian Peninsula (Lalueza-Fox *et al.*
5 2005). After the opening of the Gibraltar strait and the refilling of the Mediterranean
6 basin (c. 5 Ma), the Balearic archipelago became truly isolated; however the venue of
7 the Pliocene and Pleistocene glaciations probably created new land bridges (Cuerda
8 1975; Vesica *et al.* 2000). The main connections between all the eastern Balearic
9 Gymnesian isles (Mallorca, Menorca, Cabrera) and the western Balearic Pythiusan isles
10 (Ibiza and Formentera) occurred during a maximum sea regression in the Middle-Upper
11 Pleistocene (Cuerda 1975), roughly coincident with the end of the Mindel glaciation (c.
12 0.40 Ma) that also connected the Pythiusas with the mainland Dianic range.

13 Phylogeographical reconstructions of the Ibero-Provençal biota have mostly
14 been based on narrow endemic species that show restricted distributions in some
15 western Mediterranean islands (mainly the Balearic isles) or between this archipelago
16 and some close mainland localities (Prentice *et al.* 2003; Juan *et al.* 2004; Brown *et al.*
17 2008; Falchi *et al.* 2009; Molins *et al.* 2009). Other studies focused on species with
18 wider geographical distributions have allowed to discriminate between alternative
19 hypotheses of vicariance and dispersal strategies (Médail & Diadema 2009 and
20 references therein). Vicariant floristic relationships between the Dianic territory and the
21 Balearic Islands have been reported for several groups (Guinea 1954; Rouy 1888; Vogt
22 1991; De la Torre *et al.* 1996; Crespo 2000; Moreno Saiz 2011), and long-dispersals
23 have been suggested for others (Prentice *et al.* 2003; Juan *et al.* 2004; Moreno Saiz
24 2011). However, despite the ‘cross-roading’ biogeographical nature of the area, no
25 records exist on plausible multidirectional colonisations of Ibero-Provençal plants.

1 Furthermore, the role of their potential dispersal vectors, such as migratory sea birds or
2 sea currents, has never been assessed in depth.

3 To investigate the potential Dianic-Balearic and other alternative NE Iberian-
4 Balearic floristic dispersals suggested by Cardona and Contandriopoulos (1979), we
5 have chosen a perennial, broadly distributed, coastal and subcoastal plant, *Cheirolophus*
6 *intybaceus* (Asteraceae), which is distributed across eastern Iberia, the Balearic Isles
7 (except Menorca) and southern France (Fig. 1A) and shows evidence of both insular
8 and continental speciation events (Garnatje *et al.* 2007). Recent phylogenetic studies of
9 *Cheirolophus*, based on both nuclear and cpDNA sequences and genome sizes, showed
10 that the four closely related taxa attributed to *C. intybaceus s.l.* (e. g. *C. intybaceus s.s.*,
11 *C. grandifolius*, *C. lagunae*, *C. mansanetianus*; Appendix 1) formed a monophyletic
12 supported group nested within a western Mediterranean clade (Garnatje *et al.* 2007;
13 Vitales 2010).

14 Molecular techniques have proven to be useful in discriminating alternative
15 colonisation scenarios (Martínez-Ortega *et al.* 2004; Piñeiro *et al.* 2007; Pérez-Collazos
16 *et al.* 2009; Valente *et al.* 2010). Amplified Fragment Length Polymorphism (AFLP;
17 Vos *et al.* 1995) is a powerful fingerprinting technique that has shown to be appropriate
18 in a wide range of applications including phylogeography (Tremetsberger *et al.* 2004;
19 Pérez-Collazos *et al.* 2009) and molecular differentiation of cryptic species (Martínez-
20 Ortega *et al.* 2004; Diaz-Pérez *et al.* 2008).

21 The main hypotheses of our study were: i) The widespread *C. intybaceus* is the
22 taxon which could have originated *C. grandifolius*, *C. lagunae*, *C. mansanetianus*. In
23 order to test this hypothesis we have assessed population genetic variability and
24 structure across the whole geographical distribution area of these taxa; and ii) During
25 the Oligocene the southern Iberian Betic mountains, the Dianic range, the Balearic isles

1 and Provence formed a continuous geological territory, at this time the populations of
2 the ancestor of *C. intybaceus s.l.* were distributed in several localities along the area,
3 followed by the geographical isolation of the Balearic isles, where the populations
4 diverged to *C. grandifolius*. To test this hypothesis we have clarified the genetic
5 relationships among the four taxonomic divisions of this group, testing their genetic
6 differentiation and evaluating the origin and potential colonisation routes followed by
7 the three taxa geographically restricted to islands and capes, as well as the widespread
8 *C. intybaceus*.

9

10

11 **Materials and methods**

12 *Plant sampling, DNA extraction and AFLP analysis*

13

14 *Cheirolophus intybaceus s.l.* is a morphologically polymorphic group that
15 comprises four close diploid taxa. The commonest one, *C. intybaceus s.s.*, shows a wide
16 distribution range in the mainland Ibero-Provençal belt, from its southernmost point in
17 eastern Andalusia (SE Spain) to its northernmost point at the Rhône estuary (S France)
18 (Fig. 1A, Appendix 1). The remaining three are strict endemic plants; *Cheirolophus*
19 *lagunae* is a narrow Dianic-range endemic, only known from a single locality (Cape
20 Prim, Alicante, Spain) where it constitutes two small populations (Olivares *et al.* 1995;
21 Segarra-Moragues com. pers.), *C. grandifolius* is endemic to the Balearic isles (Stübing
22 *et al.* 1997) and *C. mansanetianus* is only known from the Cabo de Gata (Almería, SE
23 Spain). They differ from each other in their habit, the shape and size of their leaves and
24 achenes, and the ornamentation of their pollen grains (Olivares *et al.* 1995; Stübing
25 *et al.* 1997). In addition, several varieties have also been described across the distribution

1 area of *C. intybaceus* s.s. due to the great morphological variability of its leaves
2 (Susanna 1989). Nonetheless, all these taxa are closely related and descend from a
3 common ancestor (Garnatje *et al.* 2007; Vitales 2010). *Cheirolophus intybaceus* s.l.
4 occurs in open habitats in the shrublands and garriga forests of sea watersheds,
5 including coastal cliff waterfronts, and in continental areas (up to 840 m. a. s. l.) of the
6 inland Mediterranean rivers' basins, avoiding low thermal inversion temperatures
7 caused by winter fogs (Garnatje *et al.* 2009).

8 A total of 270 individuals from 37 natural isolated populations of *Cheirolophus*
9 *intybaceus* s.l. were collected in the field covering the whole distribution of *C.*
10 *intybaceus* s.s. (25 populations), *C. grandifolius* (8 populations), *C. lagunae* (2
11 populations) and *C. mansanetianus* (2 populations) (Fig. 1A and Appendix 1).

12 Genomic DNA was extracted from silica gel-dried leaves following the protocol
13 of Doyle & Doyle (1987) with slight modifications. AFLP technique was carried out
14 following the protocol described in Vos *et al.* (1995) and in accordance with the
15 manufacturer's instructions (Invitrogen Ltd, UK) with some modifications. Three
16 primer pair combinations (E-AAC/M-CAG, E-ACA/M-CTC and E-AGC/M-CAA)
17 from ten combinations tested in a previous screening assay were selected. The obtained
18 products were separated on 5% polyacrylamide gels and silver stained (Bassam *et al.*
19 1991). AFLP reproducibility test was conducted by re-extracting DNA from ten random
20 individuals (each from a different population) and repeating the entire AFLP procedure
21 in an independent assay. The error rate was calculated for every primer combination as
22 the number of phenotypic differences related to the total number of phenotypic
23 comparisons, and subsequently averaged over the three combinations.

24
25

1 *Genetic diversity and structure and spatial analyses*

2 Genetic diversity analyses were performed to determine the ancestries of the
3 taxa, the levels of genetic variability and to infer the origins of the populations.
4 Bayesian genetic diversity, defined as average panmictic heterozygosity within each
5 population (h_s), and its average value across populations (H_t) were computed with
6 Hickory v. 1.0.4 (Holsinger & Lewis 2003). Additionally, heterozygosity (H_j) was
7 calculated using TFPGA v1.3 (Miller 1997). The number of rare (fr ; frequencies less
8 than 0.1) private (fu , unique bands), and diagnostic (f_d , unique bands with a frequency
9 equal to 1) fragments, and Nei's (1973) diversity index (h) were calculated from the
10 binary data matrix. Further measures of genetic diversity were estimated through: (1)
11 the frequency-down-weighted marker values (DW) index of Schönswetter & Tribsch
12 (2005) using AFLPDAT (Ehrich 2006), (2) the band richness (Br), which is the number
13 of phenotypes expected at each locus and can be interpreted as an analogue of the allelic
14 richness, ranging from 1 to 2 (Coart *et al.* 2005), and (3) the percentage of polymorphic
15 loci (PLP) with a significance of 1% ($p=0.99$). Br and PLP indices were calculated
16 according to the rarefaction method of Hurlbert (Petit *et al.* 1998), and conditioned to
17 the smallest studied population (4) using the software AFLPdiv v1.0
18 (<http://www.pierroton.inra.fr/genetics/labo/Software/Aflpdiv/>). To assess whether
19 genetic diversity differed among different geographical ranges of *C. intybaceus s.l.*,
20 diversity range values were compared in a non-parametric Kruskal-Wallis test using
21 PAST v 1.98 software (Hammer *et al.* 2001). Bonferroni corrections of Mann-Whitney
22 pairwise comparisons were applied to probabilities of the analysis.

23 Genetic structure and genetic relationship analyses allowed us to test if the *C.*
24 *intybaceus s.l.* populations were widely distributed across the whole distribution range
25 or were restricted to specific areas (e. g., the Betic mountains, the Iberian Dianic range

1 and the Balearic isles) at specific geological times. Bayesian genetic structure among
2 populations was analysed through the unbiased derived estimate θ_{β} (analogue of
3 Wright's F_{ST} coefficient) based on the random model of Weir & Cockerham (1984)
4 using Hickory. For this analysis, prior distributions were specified using the default
5 values of the program (burn-in set to 50000, sampling set to 250000, and thin set to 50)
6 and data were analysed under the free model, suiting the characteristics of dominant
7 AFLP markers (Díaz-Pérez *et al.* 2008). Additionally, analyses of molecular variance
8 (AMOVA) were conducted using Arlequin 3.11 (Excoffier *et al.* 2005). The AMOVAs
9 were conducted using: 1) the entire data set considering all populations belonging to the
10 same group and within and among groups in order to test the existence of genetic
11 structure among geographical or taxonomic divisions, and 2) using geographical
12 division data set to infer the partition of the variance within geographical groups. The
13 F_{ST} coefficient of Wright (1951) was also obtained from AMOVA.

14 A Bayesian model-based analysis was performed to infer the structure of the
15 populations and the spatial ranges of *C. intybaceus s.l.* using STRUCTURE v.2.2
16 (Falush *et al.* 2007). Due to the monophyly of the *C. intybaceus s.l.* group (Garnatje *et*
17 *al.* 2007; Vitales 2010), we imposed the admixture ancestry model with correlated allele
18 frequencies. We ran the analysis for a range of K values starting from 1 to 38, using a
19 burn-in period and a run length of the Monte Carlo Markov Chain of 75000 and 150000
20 iterations, respectively. Chain convergence was estimated through visual inspection of
21 the posterior values excluding the burnin. Eight iterations were conducted using the *ad*
22 *hoc* parameter ΔK of Evanno *et al.* (2005) to estimate the rate of change of likelihood
23 values between successive K values. The admixture proportion of an individual (α) was
24 obtained from STRUCTURE; α values close to zero imply that most individuals derive
25 from one or another ancestral Bayesian group, whereas α values >1 imply that most

1 individuals are admixed (Diaz-Pérez *et al.* 2008). Genetic substructuring within the
2 inferred spatial ranges was further assessed through independent analyses of the split
3 data matrices using the same procedures indicated above (for K groups ranging from 1
4 to the number of analysed populations plus two).

5 To evaluate the origins of populations and their colonisation routes, genetic
6 relationships among individuals and populations were assessed using a Principal
7 Coordinates Analysis (PCO) and a Minimum Spanning Tree (MST) analysis on
8 NTSYSpc v. 2.11a (Rohlf 2002). In order to test the robustness of the PCO+MST
9 analysis and the genetic link found between the NE Spain Garraf population and the
10 Balearic Island populations (see Results), we conducted PCO+MST pseudosampling
11 analyses using: i) 10 pseudo-matrices constructed with all the original individuals but
12 bootstrapping the AFLP markers over loci, ii) 20 pseudo-matrices constructed with
13 seven randomly chosen individuals in each of the three main geographical areas (SE
14 Iberia, NE Iberia + S France, Balearic Islands) plus all the seven individuals from
15 Garraf, iii) 10 pseudo-matrices constructed with seven randomly chosen individuals in
16 each of the three main geographical areas (SE Iberia, NE Iberia + S France, Balearic
17 Islands) plus all the seven individuals from Garraf but bootstrapping the AFLP markers
18 over loci. To further investigate the genetic relationships between pairs of regions,
19 autocorrelograms between the four K optimal spatial groups detected with
20 STRUCTURE and AMOVA (NE Iberia+ SW France, SE + E Iberia, Balearics, and
21 Garraf, see Results) were calculated (see Pérez-Collazos *et al.* 2009 for methodology).
22 Isolation by distance was assessed by matrix correlation analyses between a matrix of
23 pairwise linearised F_{ST} values computed with ARLEQUIN and a matrix of pairwise
24 geographical distances obtained with the Geographic Distance Matrix Generator v.1.2.3

1 software (Ersts 2006). Significance of the correlation was tested for 1000 Mantel test
2 permutations using NTSYSpC.

3

4

5 **Results**

6 *Genetic diversity and structure of the Cheirolophus intybaceus s.l.*

7 A total of 103 AFLP fragments were obtained. The AFLP data provided unique
8 multilocus profiles for each individual. In general, highly reproducible AFLP patterns
9 were obtained. An average error rate of 2.1% was estimated across the ten pilot samples
10 for all three primer pairs. This value fell below the maximum error rate percentage
11 accepted for good AFLP reproducibility (5%) (Piñeiro *et al.* 2007). Rare fragments
12 were very scarce across the studied populations; only one rare fragment was detected in
13 CG14 (Appendix 1). No diagnostic fragments were found at population, taxonomic or
14 geographical-range levels, and only three populations, CG44, CI40 and CI41, showed
15 one unique fragment (Appendix 1). The Bayesian heterozygosity and Nei diversity
16 values were similar in all populations, with a mean *C. intybaceus s.l.* value of $Ht =$
17 0.211 , and $h = 0.620$, respectively. The percentage of polymorphic loci ($PLP \%$)
18 obtained by the rarefaction method, heterozygosity values (H_j), Bayesian heterozygosity
19 at species level (h_s), Nei diversity (h), and band richness (Br) were significantly lower
20 ($p < 0.05$) in the Balearic islands populations than in the mainland populations
21 (Appendix 1).

22 The DW index showed the highest mean value in the Balearic populations of *C.*
23 *grandifolius* ($DW = 3.754$; Appendix 1). Nevertheless, the widespread *C. intybaceus*
24 also showed a high mean value of fragment rarity ($DW = 3.640$), the highest value
25 being the one detected in the E Iberian population CI40 ($DW = 9.400$; Appendix 1).

1 The Bayesian estimate of genetic divergence among populations was 0.468, with
2 95% of the credible intervals of the analysis ranging between 0.446 and 0.489. The
3 separate analysis conducted on the three main geographical groups showed that the
4 divergences were higher among the Balearic populations (0.517) than among the
5 continental ones (0.417), and were similar between the SE+E Iberian (0.392) and the
6 NE Iberian+SW France (0.397) continental populations.

7 The results of non-hierarchical AMOVA indicated a high significant level of
8 genetic structure among populations (58.44%; $df = 36$; $p < 0.0001$). Two geographic
9 groups (Balearic islands + Garraf (CI8) versus remaining continental populations)
10 showed the highest percentage of variance among groups (29.82%; $df = 35$; $p < 0.001$)
11 of the hierarchical AMOVAs, indicating the existence of two main genetic groups.
12 Nested AMOVA of four taxonomic groups accounted for less variance (25.47%; $df =$
13 33 ; $p < 0.001$). The separate AMOVA of Balearic + Garraf populations into four
14 groups, the three islands and the mainland Iberian locality (Ibiza, Formentera, Mallorca,
15 Garraf), showed moderate levels of genetic structure (28.34%; $df = 3$; $p < 0.001$). By
16 contrast, a low percentage of variation among groups (9.66%; $df = 27$; $p < 0.001$) was
17 detected within the continental populations when they were separated into two
18 geographical ranges (SE-E Iberia versus NE Iberia-SW France).

19

20 *Spatio-geographical structure and correlations of genetic and geographical data*
21 *among populations and ranges*

22 The Bayesian analysis of population genetic structure conducted with
23 STRUCTURE found the highest $\ln P(D)$ and ΔK values for $K = 2$. This grouping
24 separated the insular Balearic populations (plus Garraf) from the mainland populations,
25 showing high percentages of individual memberships (87% and 77%, respectively) to

1 these predefined groups (Fig. 2A). The independent STRUCTURE analysis of the
2 mainland group (K=2) separates the southern versus northern subgroups (Fig. 2B). The
3 analysis of the 'Balearic group' differentiated four Bayesian subgroups (K=4; Ibiza,
4 Formentera, Mallorca, and the mainland Garraf) (Fig. 2C). The individual proportion of
5 admixture (α) was low in all K=2 runs for the entire data set ($\alpha = 0.079$; Fig. 2A),
6 ranging between 0.082 and 0.077. The mean value of alpha was 0.233 for the
7 continental (K=2; Fig. 2B) population structure, which showed more admixture, and
8 0.034 for the Ibiza, Formentera, Mallorca and Garraf structure (K=4; Fig. 2C),
9 indicating that these individuals were derived from different genetic groups.

10 Similar results were found in the PCO analysis. The inland and Balearic
11 populations were separated into two genetic groups in the 3D scatterplot (Fig. 3); the
12 mainland Garraf population showed an intermediate position between them, being more
13 related to the insular rather than to the mainland populations. The mainland populations
14 subdivided into SE + E Iberia and NE Iberia + SW France groups (Fig. 3). MST
15 analysis revealed a single genetic connection between the continental and the insular
16 populations; the E Iberian Alicantean CL48 population was the likely mainland genetic
17 link to the Ibizan CG44 population. MST also showed a link among the Ibiza
18 populations to the Formentera CG15 population and to the Mallorcan CG17 population,
19 and Garraf apparently derived from the Mallorcan CG20 population (Fig. 3). The E
20 Iberian Valencian CI33 population was resolved as the potential genetic link between
21 the southern (SE + E Iberian) and the northern (NE Iberian + SW France) subclusters.
22 The low, middle and high Ebro valley populations and the Rhône estuary populations
23 linked to different Catalanian and Provençal subgroups respectively (Fig. 3). In all the
24 pseudosampling PCO analyses (40) the Garraf individuals clustered near the Balearic
25 Island individuals and both groups separated from the remaining mainland individuals

1 that clustered together in the opposite side of the first PCO axis (Results not shown).
2 Similarly, the MST pseudosampling analyses showed a single genetic link between the
3 Balearic cluster and the mainland cluster, with the Garraf cluster derived from the
4 Balearic group rather than from the mainland southern (SE + E Iberian) or northern (NE
5 Iberian + SW France) clusters in more than the 70% of the cases (Results not shown).
6 These pseudoreplicate analyses confirmed the robustness of the original PCO + MST
7 results.

8 The Mantel regional autocorrelograms revealed a positive and significant genetic
9 correlation within the northern (NE Iberia + SW France; $r = 0.323$; $p < 0.005$) and
10 southern (SE + E Iberia; $r = 0.356$; $p < 0.001$) geographical ranges (results not shown).
11 The northern and southern ranges showed a positive correlation of $r = 0.257$; $p < 0.005$.
12 The Balearic and Garraf populations showed a non-significant, slightly negative
13 correlation within their range ($r = -0.014$; $p = 0.484$). The autocorrelogram also showed
14 the genetic singularity of the Balearic and the continental Garraf populations. These
15 regions were negatively correlated to each other and to the rest of the continental
16 ranges. A positive and significant correlation between the genetic and the geographical
17 distances matrices was found in the Mantel test ($r = 0.655$; $p < 0.001$).

18
19
20
21
22
23
24
25
26
27

1 **Discussion**

2 *Pleistocene origin and dispersal of the C. intybaceus populations in the Ibero-* 3 *Provençal belt: A Dianic colonisation of the Balearic islands*

4 The endemic-rich Western Mediterranean islands have received great attention
5 in the last decades for testing alternative hypotheses of plant speciation and colonisation
6 processes within a continental tectonic scenario (Cuerda 1975; Cardona &
7 Contandriopoulos 1979; Schüle 1993; Pérez-Obiol *et al.* 2001; López de Heredia *et al.*
8 2005; Fernández-Mazuecos & Vargas, 2011). Among the recently molecularly-assayed
9 and paleogeographically studied cases, early Miocene microplate splits, late Miocene
10 Messinian land-bridge links, and recent Pleistocene long-distance dispersal events have
11 been invoked to explain the historical distributions of species and populations on islands
12 and in their neighbouring regions (Schüle 1993; Juan *et al.* 2004; López-de-Heredia *et*
13 *al.* 2005). However, in contrast to models proposed for some oceanic-island plants
14 (Carine *et al.* 2004; Rodríguez-Sánchez *et al.* 2009; Fernández-Mazuecos & Vargas
15 2011), most of the studied western Mediterranean continental-island cases shows a
16 unidirectional colonisation pattern from the mainland to the island, followed by
17 speciation by insular isolation in different geological ages (Rubio de Casas *et al.* 2006,
18 and references therein). The results of our phylogeographical study of *C. intybaceus*
19 suggest a single colonisation event from the SE Iberian mainland to the archipelago
20 followed by a single back- colonisation from the island of Mallorca to the continent (NE
21 Iberia: Garraf) (Fig. 1B). This is supported by: i) the single genetic link recovered by
22 MST between the continental and the Balearic clusters in the PCO plot (Fig. 3); ii) the
23 nested derived position of the continental Garraf population within the Balearic cluster
24 and its link to the Mallorcan population CG20 in the PCO (Fig. 3); iii) the best
25 supported spatial structure of K=2 Bayesian groups that differentiated the continental

1 populations from the islands' populations, which also included Garraf within it; and iv)
2 the best supported partition of variance in the hierarchical AMOVAs that separated the
3 Balearic+Garraf group from the remaining inland populations group.

4 Paleobotanical and paleoenvironmental data evidenced Mediterranean forest and
5 scrubland in E Iberia from the late Miocene to the Pleistocene that likely favoured the
6 establishment of *C. intybaceus*, followed by *in-situ* radiation, as has occurred in other
7 Mediterranean plants (Fauquette *et al.* 1998; González-Sampériz *et al.* 2010). This
8 agrees with the relatively basal isolated placement of the monophyletic *Cheirolophus*
9 clade within the Centaureinae clade (Hidalgo 2006; Susanna & Garcia-Jacas 2009),
10 concurring in time with the Messinian crisis, and its subsequent lineage splits spanning
11 the Pliocene and the Pleistocene.

12 In the Balearic islands, deciduous trees and thermophilous vegetation have
13 remained in different refugia across the Pleistocene (Pérez-Obiol *et al.* 2001). The
14 middle-late Pleistocene divergence of *C. intybaceus s.l.* was indirectly suggested by the
15 lack of private and diagnostic alleles at both specific and geographical levels (Appendix
16 1) and a poorly bootstrap-supported NJ phenogram (results not shown). The scarcity of
17 unique AFLP fragments and a genetic structuring based mostly on differences in allelic
18 frequencies is also a characteristic of recently evolved Macaronesian oceanic-island
19 plants (Díaz-Pérez *et al.* 2008).

20 The molecular phylogenies indicate a sister relationship of *C. intybaceus s.l.* to
21 the southern Moroccan *C. benoistii/ C. tananicus* clade (Garnatje *et al.* 2007) or to the
22 western Iberian *C. sempervirens* (Vitales 2010). These findings support the hypothesis
23 that *C. intybaceus* likely originated from taxa currently growing in the southern Iberian
24 Peninsula or in Northern Africa from one of their recent ancestors. This was also
25 supported by the AFLP data that detected higher levels of genetic diversity in the

1 southern Iberian populations than in the insular and the northern Iberian ones (Appendix
2 1). Significant differences in *PLP%*, *H_j*, *h* and *Br* values were found between the
3 continental and the Balearic populations, with the latter group representing a subset of
4 the marker variability observed on the continent (Appendix 1). Higher levels of genetic
5 variability are expected in relict populations because of the accumulation of both
6 ancient polymorphisms and new rare alleles acquired by mutations (Schönswetter &
7 Tribsch 2005; Díaz-Pérez *et al.* 2008; Pérez-Collazos *et al.* 2009). A south-to-north
8 phylogeographical model could be thus retrieved for the *C. intybaceus* lineage (Fig 1B),
9 in which likely ancient SE Iberian populations migrated to the E Iberian region at the
10 onset of the Pleistocene and from there spread both to the Balearic Isles and to the NE
11 Iberian-SW France area in the middle-late Pleistocene, according to the dated
12 phylogeny and the PCO plot (Fig. 3). This scenario is consistent with northwards
13 dispersion routes across the Iberian Peninsula proposed for other plants and animals (De
14 Jong 1998; Sanmartín 2003; Pérez-Collazos *et al.* 2009, Juan *et al.* 2011).

15 The colonisation of the Balearic Isles from the continent apparently began from
16 the Dianic territory of the E Iberian Peninsula, as depicted in the PCO plot that shows a
17 single genetic link between the Alicantean population CL48 and the Balearic subgroup
18 (Fig. 3). Dianic-Balearic floristic connections are common, especially between the
19 Dianic range and the western Pythiusan isles, separated only by 100 km of lineal
20 distance, for which up to 60 paleo- and schizoendemisms have been reported (Mansion
21 *et al.* 2008). Geologically, the mountains of Ibiza and Mallorca are an extension of the
22 Betic-Dianic ranges, and the uplifting dates back to the alpine movements produced in
23 the Oligocene-Miocene (Ríos 1975). The origin of the Dianic-Pythiusan links has been
24 mostly attributed to vicariance events caused by terrestrial connections and subsequent
25 insular isolation occurred either in the late-Miocene (c. 5 Ma) or in the Middle

1 Pleistocene (ca. 0.4 Ma), with several species, such as *Micromeria inodora* (Desf.)
2 Benth. and *Carduncellus dianius* Webb, being indicative of the Alicantean-Ibizan
3 affinity (Cardona & Contandriopoulos 1979). Other studies, however, supported a
4 recent Long Distance Dispersal (LDD) Pleistocene scenario for the colonisation of Ibiza
5 from the Dianic range, or were unable to clarify the tempo and direction of the
6 colonisation, suggesting a contemporary human-mediated introduction of populations of
7 *Silene hifacensis* Rouy either in Ibiza or in the Dianic area (Prentice *et al.* 2003). Our
8 results suggest a colonisation of the islands by continental stocks (Fig. 3), a likely event
9 considering the existence of interterritorial land-bridges and potential glacial refugia for
10 this coastal plant at that time (Médail & Diadema 2009). The posterior insular isolation
11 of the archipelago from the mainland during the Mindel-Würm interglacial (Cuerda
12 1975) probably created a strong gene flow barrier that promoted the *in-situ* speciation of
13 *C. grandifolius* in the Balearic isles. No further continental colonisation have been
14 detected as deduced from; i) the monophyly of the Balearic clade, ii) its unique genetic
15 link to the Dianic cluster (Fig. 3); and iii) the negative genetic correlations of the
16 Balearic region with any continental area (results not shown). Furthermore, strong
17 insular genetic drift was probably the main factor driving the genetic depletion detected
18 within the populations of this species (Appendix 1).

19

20 *Insular spread and re-colonisation of the mainland from the Balearic archipelago*

21 The close genetic relationships observed between the more variable and
22 ancestral Ibiza-Formentera and the more recent Mallorcan *C. intybaceus s.l.* populations
23 (Appendix 1, Figs. 2C, 3) suggest a rapid colonisation of the eastern Mallorca island
24 from the western Pithyuses. This probably occurred during the post-Mindel glaciation,
25 the last geological period when the Pithyuses and the Gymnesies formed a single mass

1 after a maximum sea regression that lowered the sea level about 130 m below its current
2 level (Cuerda 1975). By contrast, the relatively strong spatial structure detected within
3 the Balearic range, evidenced by the three AFLP island-specific supported Bayesian
4 groups (Fig. 2C), was probably a consequence of the recent separation of the islands
5 that only become completely isolated from one other after the LGM, approx. 0.018 Ma
6 ago (Cuerda 1975). Similar patterns of narrow insular Pliocene-to-Pleistocene
7 population isolations, related to marine transgressions and regressions, have been
8 documented for Balearic lizards (Brown *et al.* 2008).

9 The proposed colonisation of the coastal NE Iberian locality of Garraf from
10 Mallorca requires a Pleistocene LDD event. The inclusion of this population within the
11 Balearic group was strongly supported by the Bayesian and AMOVA analyses (Fig.
12 2A), though its derived position within the Balearic cluster supported its recent origin
13 (Fig. 3). The independent genetic structuring shown by the Garraf population in the
14 restricted Balearic isles + Garraf Bayesian analysis (Fig. 2C), similar to those retrieved
15 for the Formentera, Ibiza and Mallorca groups, suggests a common time of divergence
16 for the four groups. It could be speculated that the back-colonisation of Garraf from
17 Mallorca occurred during the Late Pleistocene Mindel times and that the plausible
18 founder effect could have been counteracted by occasional crossing with genetically
19 distinct surrounding populations. This would explain the lack of morphological *C.*
20 *grandifolius* traits in the Garraf population, though most of its genetic identity remained
21 Balearic.

22 Single and multiple unidirectional colonisation models have been proposed to
23 explain island plant radiations (Carine *et al.* 2004; Díaz-Pérez *et al.* 2008). However,
24 few cases have reported back-colonisation from islands to the mainland (Prentice *et al.*
25 2003; Carine *et al.* 2004). Seabirds have shown to have mediated in some of these

1 events, such as the back-colonisation of the narrow endemic *Medicago citrina* from
2 Ibiza to the E Iberian Columbretes Islands (Juan *et al.* 2004). Several shearwater species
3 (e.g. *Calonectris diomedea*, *Puffinus mauretanicus*) move frequently from the Balearic
4 archipelago, where they nest, to their forage grounds in front of the coastal littoral zone
5 of the E Iberian Peninsula during their post-nuptial period (Louzao *et al.* 2006; Oro *et*
6 *al.* 2009; Seo/BirdLife 2009), showing an overlapping distribution with that of *C.*
7 *intybaceus*, especially in the Mallorca-Garraf area (Fig. 1C). The presence of currently
8 extinct shearwaters (e.g. *Puffinus nestori*) in the Balearic Islands has been confirmed in the
9 fossil record since the Upper Pliocene (Alcover 1989). Furthermore, recent fossil evidences
10 of currently extant species, such as the above mentioned *Calonectris diomedea* and *Puffinus*
11 *mauretanicus*, have been documented in several Mediterranean localities (Sánchez-Marco
12 2004), including the Balearic Islands (Florit *et al.* 1989; McMinn *et al.* 1992) and more
13 strikingly, a coastal cave in the Garraf massif (Cova del Gegant, Sanz *et al.* 2008) from the
14 Middle Pleistocene onwards. Based on these evidences, it could be hypothesised that
15 during recent late Pleistocene times, similar birds' movements could have dispersed
16 mature seeds of *Cheirolophus* attached to the birds' body from the islands to the near
17 continent. Epizoochory via seed adhesion has been observed in many plant groups
18 (Sorensen 1986), and this is especially frequent in coastal and subcoastal plants. This
19 bird-mediated back-colonisation scenario could therefore explain the recent dispersal of
20 *C. intybaceus* from Mallorca to Garraf, however, other potential transport mechanisms,
21 such as the sea currents, should also be considered. The characteristics of the *C.*
22 *intybaceus* achenes (short pappus, resistant exocarpus, relatively large weight, presence
23 of elaiosomes) together with its lack of appropriate structures for wind but not for water
24 dispersal support a potential overseas transport (Vitales *et al.* unpub. res.). The long-
25 standing presence of the western Mediterranean sea-current that moves from the
26 Algerian zone, along the E Iberian-Balearic coasts, to the gulf of Lyon, and collides

1 again to close the great anticyclonic gyre (Bas 2009), might have acted as a sea-corridor
2 for dispersal of rafting *C. intybaceus s.l.* achenes from Mallorca to the Catalanian coast.
3 Similar long-distance colonisations across the sea surface have been abundantly
4 documented for oceanic-island plants (Quilichini & Debussche 2000).

5

6 *Continental spread of the C. intybaceus s.l. populations*

7 In contrast to the relatively strong genetic structure and clear dispersal patterns
8 detected within the Balearic group, the continental *C. intybaceus s.l.* core showed less
9 differentiation but more complex reticulate structuring. Both the PCO and Structure
10 analyses recovered two main geographic subgroups, the older southern core that
11 included the more ancestral SE Iberian populations (Betic and Dianic ranges), and the
12 younger northern core that included the more recently derived NE Iberian-SW France
13 (Iberian, Pyrenean and Ebro valley ranges) and S France populations (Rhône valley
14 range) (Figs. 2B, 3). The intermediate position between the two cores of the
15 geographically equidistant E Iberia Valencian populations in the PCO plot (Fig. 3)
16 further supported the suggested stepping-stone northwards colonisation route (Fig. 1B).
17 However, the unclear networking PCO patterns observed within the two cores (Fig. 3)
18 and the admixed genetic nature of some individuals detected in the K=2 Bayesian
19 analysis (Fig. 2B) points towards an overall medium-to-long-distance Pleistocene
20 reticulate scenario.

21 The divergence of the continental *C. intybaceus s.l.* NE Iberian-S France lineage
22 apparently overlapped in time with that of the Balearic split, indicating that the
23 colonisation of the northern area likely occurred in the Middle Pleistocene from older
24 SE Iberian Pliocene ancestors. The reticulate pattern observed within both the southern
25 and the northern continental population cores (Figs. 2B, 3) indicated a large spread of *C.*

1 *intybaceus* s.l. genotypes in terrestrial land during the Middle and Upper Pleistocene.
2 This was corroborated by the regional correlograms that showed overall positive
3 correlations between these two geographically continuous continental areas, and
4 detected negative correlations for insular and continental areas. The colonisation of the
5 mainland Ebro valley by this coastal and subcoastal plant apparently took place at
6 different times and from different NE Iberian sources as deduced from the PCO plot
7 (Fig. 3) and the NJ tree (results not shown). By contrast, the northeasternmost Rhône
8 valley estuary populations presumably derived from SW France Roussillon-Languedoc
9 long-distance dispersers (Fig. 1B). The large and continuous spread of colonisers at
10 medium-to-long spatial distances in these continental settings could have been
11 facilitated by the existence of water courses and suitable calcareous habitats for the
12 establishment of the new populations. Although the ways and mechanisms of transport
13 of the potential plant migrants still remain unknown, potential dispersal vectors such as
14 the seagulls (*Larus cachinnans michahellis*) that follow migratory routes along the Ebro
15 valley and up to the Cantabrian coast (Martínez-Abraín *et al.* 2002), could have acted as
16 the likely transporters of *C. intybaceus* seeds to those innermost mainland localities.

17

18

19 **Acknowledgments**

20 The authors thank J. Aparicio, O. Hidalgo, I. Sánchez-Jiménez, J. G. Segarra-
21 Moragues, M. A. Soriano and J. Vallès, for collecting material of *Cheirolophus*
22 *intybaceus*, J. Donazar for his advise on seabird dispersals, I. Sánchez-Jiménez for
23 technical support, Alastair Plant and S. Pyke for linguistic assessment, and the associate
24 editor and two anonymous referees for their critical comments to an earlier version of
25 the manuscript. This work was supported by MICINN - Spanish Government grant

1 projects (grant numbers: CGL2007-64839-C02-01, CGL2010-22234-C02-01 and
2 200630I241 Intramural), the SGR (grant number: 2009SGR439) Generalitat de
3 Catalunya project and the Aragon Government Bioflora project cofounded by the
4 European FEDER program.

5

6 **Conflict of interest.**

7 The authors declare no conflict of interest.

8

9

10 **References**

- 11 Alcover J.A. (1989) Les aus fòssils de la Cova de Ca na Reia. *Endins*, **14-15**, 95–100.
- 12 Bas C. (2009) The Mediterranean: a synoptic review. *Contributions to Science*, **5**, 25–
13 39.
- 14 Bassam B.J., Caetano-Anolles G., Gresshoff P.M. (1991) Fast and sensitive silver
15 staining of DNA in polyacrylamide gels. *Annals of Biochemistry*, **196**, 80–83.
- 16 Brown R.P., Terrasa B., Pérez-Mellado V., Castro J.A., Hoskisson P.A., Picornell A.,
17 Ramon M.M. (2008) Bayesian estimation of post-Messinian divergence times in
18 Balearic Island lizards. *Molecular Phylogenetics and Evolution*, **48**, 350–358.
- 19 Cardona M.A., Contandriopoulos J. (1979) Endemism and evolution in the islands of
20 the western Mediterranean. In: Bramwell D. (Ed.) *Plants and Islands*, Academic Press;
21 London, UK: 133–169.
- 22 Carine M.A., Russell S.J., Santos-Guerra A., Francisco-Ortega J. (2004) Relationships
23 of the Macaronesian and Mediterranean floras: molecular evidence for multiple
24 colonizations into Macaronesia and back-colonization of the continent in *Convolvulus*
25 (Convolvulaceae). *American Journal of Botany*, **91**, 1070–1085.

- 1 Coart E., Van Glabeke S., Petit R.J., Van Bockstaele E., Roldán-Ruiz I. (2005) Range
2 wide versus local patterns of genetic diversity in hornbeam (*Carpinus betulus* L.).
3 *Conservation Genetics*, **6**, 259–273.
- 4 Cohen C.R. (1980) Plate-tectonic model for the Oligo-Miocene evolution of the
5 Western Mediterranean. *Tectonophysics*, **68**, 283–311.
- 6 Contandriopoulos J., Cardona M.A. (1984) Caractère original de la flore endémique des
7 Baléares. *Botanica Helvetica*, **94**, 101–132.
- 8 Crespo M.B. (2000) Comentarios sobre los táxones del grupo de *Hippocrepis balearica*
9 Jacq. (*Leguminosae*). *Flora Montiberica*, **14**, 48-51.
- 10 Cuerda J. (1975) *Los tiempos cuaternarios en Baleares*. Instituto de Estudios
11 Baleáricos: Palma de Mallorca.
- 12 De Jong H. (1998) In search of historical biogeographic patterns in the western
13 Mediterranean terrestrial fauna. *Biological Journal of the Linnean Society*, **65**, 99–164.
- 14 De la Torre A., Alcaraz F., Crespo M.B. (1996) Aproximación a la biogeografía del
15 sector Setabense (provincia Catalano-Valenciano-Provençal). *Lazaroa*, **16**, 141–158.
- 16 Díaz-Pérez A., Sequeira M., Santos-Guerra A., Catalán P. (2008) Multiple
17 colonizations, *in situ* speciation, and volcanism associated stepping-stone dispersals
18 shaped the phylogeography of the Macaronesian red fescues (*Festuca* L., Gramineae).
19 *Systematic Biology*, **57**, 732–749.
- 20 Doyle J.L., Doyle J.J. (1987) A rapid DNA isolation procedure for small quantities of
21 fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.
- 22 Ehrlich D. (2006) AFLPdat: a collection of R functions for convenient handling of
23 AFLP data. *Molecular Ecology Notes*, **6**, 603–604.

- 1 Ersts P.J. (2006) Geographic Distance Matrix Generator version 1.2.3. American
2 Museum of Natural History, Center for Biodiversity and Conservation. Available from
3 http://biodiversityinformatics.amnh.org/open_source/gdmg.
- 4 Evanno G., Regnaut S., Goudet J. (2005) Detecting the number of clusters of
5 individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*,
6 **14**, 2611–2620.
- 7 Excoffier L., Laval G., Schneider S. (2005) Arlequin ver. 3.0: An integrated software
8 package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**,
9 47–50.
- 10 Falchi A., Paolini J., Desjobert J.M., Melis A., Costa J., Varesi L. (2009)
11 Phylogeography of *Cistus creticus* L. on Corsica and Sardinia inferred by the *trnL-F*
12 and *rpl32-trnL* sequences of cpDNA. *Molecular Phylogenetics and Evolution*, **52**, 538-
13 543.
- 14 Falush D., Stephens M., Pritchard J.K. (2007) Inference of population structure using
15 multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*,
16 **7**, 574–578.
- 17 Fauquette S., Guiot J., Suc J.P. (1998) A method for climatic reconstruction of the
18 Mediterranean Pliocene using pollen data. *Palaeogeography, Palaeoclimatology,*
19 *Palaeoecology*, **144**, 183–201.
- 20 Fernández-Mazuecos M., Vargas P. (2011) Genetically depauperate in the continent but
21 rich in oceanic islands: *Cistus monspeliensis* (Cistaceae) in the Canary Islands. *PLoS*
22 *ONE*, **6**, e17172. doi:10.1371/journal.pone.0017172
- 23 Florit X., Mourer-Chauvire C., Alcover J.A. (1989) Els ocells pleistocènics d'es Pouàs,
24 Eivissa. *Butlletí de l'Institut Català d'Història Natural*, **56**, 35–46.

- 1 Garnatje T., Garcia S., Canela M.A. (2007) Genome size variation from a phylogenetic
2 perspective in the genus *Cheirolophus* Cass. (Asteraceae): biogeographic implications.
3 *Plant Systematics and Evolution*, **264**, 117–134.
- 4 Garnatje T., Garcia S., Hidalgo O., Pellicer J., Sánchez-Jiménez I., Vallès J. (2009)
5 *Cheirolophus intybaceus* (Asteraceae, Centaureinae) o la constància del valor 2C.
6 *Collectanea Botanica (Barcelona)*, **28**, 7–17.
- 7 Gautier F., Caluzon G., Suk J.P., Violanti D. (1994) Age et durée de la crise de salinité
8 Messinienne. *Comptes Rendus de l'Académie des Sciences de Paris*, **318**, 1103–1109.
- 9 González-Sampériz P., Leroy S.A.G., Carrión J., Fernández S., García-Antón M., Gil-
10 García Uzquiano P., Valero-Garcés B., Figueiral I. (2010) Steppes, savannahs, forests
11 and phytodiversity reservoirs during the Pleistocene in the Iberian Peninsula. *Review of*
12 *Palaeobotany and Palynology*, **162**, 427–457.
- 13 Guinea E. (1954) Cistáceas españolas. *Boletín del Instituto Forestal de Investigación y*
14 *Experiencias*, **71**, 63–160.
- 15 Hammer O., Harper D.A.T., Ryan P.D. (2001) PAST: Palaeontological Statistics
16 software package for education and data analysis. *Paleontología Electronica*, **4**, 1–9.
- 17 Hidalgo O. (2006) *El grupo Rhaponticum* (Asteraceae, Cardueae, Centaureinae):
18 *delimitación y filogenia*. PhD dissertation. Universitat de Barcelona.
- 19 Holsinger K.E., Lewis P.O. (2003) HICKORY v.1.0.4. <http://www.eeb.uconn.edu/>
- 20 Jakob S.L., Ihlow A., Blattner F.R. (2007) Combined ecological niche modelling and
21 molecular phylogeography revealed the evolutionary history of *Hordeum marinum*
22 (Poaceae)-niche differentiation, loss of genetic diversity, and speciation in
23 Mediterranean Quaternary refugia. *Molecular Ecology*, **16**, 1713–1727.
- 24 Juan A., Crespo M.B., Cowan R.S., Lexer C., Fay M.F. (2004) Patterns of variability
25 and gene flow in *Medicago citrina*, an endangered endemic of islands in the western

- 1 Mediterranean, as revealed by amplified fragment length polymorphism (AFLP).
2 *Molecular Ecology*, **13**, 2679–2690.
- 3 Juan A., Fay M.F., Pastor J., Juan R., Fernández I., Crespo M.B. (2011) Genetic
4 structure and phylogeography in *Juniperus oxycedrus* subsp. *macrocarpa* around the
5 Mediterranean and Atlantic coasts of the Iberian Peninsula, based on AFLP and plastid
6 markers. *European Journal of Forest Research*. DOI 10.1007/s10342-011-0558-5.
- 7 Kadereit J.W., Comes H.P. (2005) The temporal course of alpine plant diversification in
8 the Quaternary. *Plant Species Level Systematics New Perspectives on Pattern Process*,
9 **143**, 0–117.
- 10 Kadereit G., Yaprak A.E. (2008) *Microcnemum coralloides* (Salicornioideae,
11 Chenopodiaceae): an example of intraspecific East-West disjunction in the
12 Mediterranean region. *Anales del Jardín Botánico de Madrid*, **65**, 415–426.
- 13 Kadereit J.W., Arafteh R., Somogyi G., Westberg E. (2005) Terrestrial growth and
14 marine dispersal? Comparative phylogeography of five coastal plant species on a
15 European scale. *Taxon*, **54**, 861–876.
- 16 Kropf M., Comes H.P., Kadereit J.W. (2006) Long-distance dispersal vs. vicariance.
17 The origin and genetic diversity of alpine plants in the Spanish Sierra Nevada. *New*
18 *Phytologist*, **172**, 169–184.
- 19 Lalueza-Fox C., Castresana J., Sampietro L., Marquès-Bonet T., Alcover J.A.,
20 Bertranpetit J. (2005) Molecular dating of caprines using ancient DNA sequences of
21 *Myotragus balearicus*, an extinct endemic Balearic mammal. *BMC Evolutionary*
22 *Biology*, **5**, 70.
- 23 López de Heredia U., Jiménez P., Díaz Fernández P., Gil L. (2005) The Balearic
24 Islands: a reservoir of cpDNA genetic variation for evergreen oaks. *Journal of*
25 *Biogeography*, **32**, 939–949.

- 1 Louzao M., Hyrenbach K.D., Arcos J.M., Abelló P., Gil de Sola L., Oro D. (2006)
2 Oceanographic Habitat of an Endangered Mediterranean Procellariiform: Implications
3 for Marine Protected Areas. *Ecological Applications*, **16**, 1683–1695.
- 4 Mansion G., Rosenbaum G., Schoenenberger N., Bacchetta G., Rosselló J.A., Conti E.
5 (2008) Phylogenetic analysis informed by geological history supports multiple,
6 sequential invasions of the Mediterranean basin by the angiosperm family Araceae.
7 *Systematic Biology*, **57**, 269–285.
- 8 Martínez-Abraín A., Oro D., Cardá J., Del Señor X. (2002) Movements of Yellow-
9 legged Gull *Larus cachinnans michahellis* from two small Western Mediterranean
10 colonies. *Atlantic Seabird*, **4**, 101-108.
- 11 Martínez-Ortega M.M., Delgado L., Albach D.C., Elena-Rossello J.A., Rico E. (2004)
12 Species boundaries and phylogeographic patterns in cryptic taxa inferred from AFLP
13 markers: *Veronica* subgen. *Pentasepalae* (Scrophulariaceae) in the Western
14 Mediterranean. *Systematic Botany*, **29**, 965–986.
- 15 McMinn M., Altaba C.R., Alcover J.A. (1992) Els ocells del Pleistocè superior de la Cova
16 Nova (Capdepera, Mallorca). III. Noves aportacions al registre. *Bolletí de la Societat*
17 *d'Història Natural de les Illes Balears*, **35**, 17–31.
- 18 Médail F., Quézel P. (1997) Hot-spots analysis for conservation of plant biodiversity in
19 the Mediterranean basin. *Annals of the Missouri Botanical Garden*, **84**, 112–127.
- 20 Médail F., Diadema K. (2009) Glacial refugia influence plant diversity patterns in the
21 Mediterranean Basin. *Journal of Biogeography*, **36**, 1333–1345.
- 22 Miller M.P. (1997) *Tools for Population Genetic Analyses (TFPGA) version 1.3*,
23 Northern Arizona University: USA.

- 1 Molins A., Mayol M., Rosselló J.A. (2009) Phylogeographical structure in the coastal
2 species *Senecio rodriguezii* (Asteraceae), a narrowly distributed endemic Mediterranean
3 plant. *Journal of Biogeography*, **36**, 1372–1383.
- 4 Moreno Sainz J.C. (2011) La diversidad florística vascular española. In: Viejo J.L.
5 (Ed.). *Biodiversidad. Aproximación a la biodiversidad botánica y zoológica en España*.
6 Memorias de la Real Sociedad Española de Historia Natural IX: 75-108.
- 7 Nei M. (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the*
8 *National Academy of Sciences of the USA*, **70**, 3321-3323.
- 9 Olivares A., Peris J.B., Stübing G., Martín J. (1995) *Cheirolophus lagunae*, sp. nov.
10 (Asteraceae), endemismo iberolevantino. *Anales del Real Jardín Botánico de Madrid*,
11 **53**, 262-265.
- 12 Oro D., Louzao M., Genovart M. (2009) Pardela balear – *Puffinus mauretanicus*. In:
13 Salvador A., Bautista L.M. (Eds.). *Enciclopedia Virtual de los Vertebrados Españoles*.
14 Museo Nacional de Ciencias Naturales, Madrid.
- 15 Pérez-Collazos E., Sánchez-Gómez P., Jiménez J.F., Catalán P. (2009) The
16 phylogeographical history of the western Mediterranean *Ferula loscosii* (Apiaceae).
17 *Molecular Ecology*, **18**, 848–861.
- 18 Pérez-Obiol R., Yll E.I., Pantaleón-Cano J., Roure J.M. (2001) Canvis en el paisatge
19 vegetal de les Illes Balears durant el final del Quaternari. In: Pons G.X., Guijarro J.A.
20 (Eds.) *El canvi climatic: passat present i futur. Monografia de la Societat d'Història*
21 *Natural de les Balears*, **9**, 69-89.
- 22 Petit R.J., El Mousadik A., Pons O. (1998) Identifying populations for conservation on
23 the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

- 1 Piñeiro R., Fuertes J., Draper D., Nieto G. (2007) Ecology matters: Atlantic-
2 Mediterranean disjunction in the sand-dune shrub *Armeria pungens* (Plumbaginaceae).
3 *Molecular Ecology*, **16**, 2155–2171.
- 4 Prentice H.C., Malm J.U., Mateu-Andrés I., Segarra-Moragues J.G. (2003) Allozyme
5 and chloroplast DNA variation in island and mainland populations of the rare Spanish
6 endemic, *Silene hifacensis* (Caryophyllaceae). *Conservation Genetics*, **4**, 543–555.
- 7 Quilichini A., Debussche M. (2000) Seed dispersal and germination patterns in a rare
8 Mediterranean island endemic (*Anchusa crispa* Viv., Boraginaceae). *Acta Oecologica*,
9 **21**, 303–313.
- 10 Ríos J.M. (1975) El mar Mediterráneo occidental y sus costas ibéricas. *Boletín de la*
11 *Real Sociedad Española de Historia Natural*, **70**, 1–473.
- 12 Rodríguez-Sánchez F., Guzmán B., Valido A., Vargas P., Arroyo J. (2009) Late
13 Neogene history of the laurel tree (*Laurus* L., Lauraceae) based on phylogeographical
14 analyses of Mediterranean and Macaronesian populations. *Journal of Biogeography*, **36**,
15 1270–1281.
- 16 Rohlf F.J. (2002) NtSYSpc, *Numerical Taxonomy and Multivariate analysis System*.
17 *Versison 2.11a, User guide*. Exeter software: New York.
- 18 Rosenbaum G., Lister G.S., Duboz C. (2002) Reconstruction of the tectonic evolution in
19 the western Mediterranean since the Oligocene. In: Rosenbaum G., Lister G.S. (Eds.)
20 *Reconstruction of the evolution of the Alpine-Himalayan Orogeny*. *Journal of the*
21 *Virtual Explorer*, **8**, 107–130.
- 22 Rouy G. (1888) Excursions botaniques en Espagne (Mai-Juin 1883). Denia-Madrid. III.
23 Observations, remarques et diagnoses. *Bulletin de la Societe Botanique de France*, **35**,
24 115–124.

- 1 Rubio de Casas R., Besnard G., Schönswetter P., Balaguer L., Vargas P. (2006)
2 Extensive gene flow blurs phylogeographic but not phylogenetic signal in *Olea*
3 *europaea* L. *Theoretical and Applied Genetics*, **113**, 575–583.
- 4 Sánchez-Marco A. (2004) Avian zoogeographical patterns during the Quaternary in the
5 Mediterranean region and paleoclimatic interpretation. *Ardeola*, **51**, 91–132.
- 6 Sanmartín I. (2003) Dispersal vs. vicariance in the Mediterranean. Historical
7 biogeography of the Palaearctic Pachydeminae (Coleoptera, Sacarabaeoidea). *Journal of*
8 *Biogeography*, **30**, 1883–1897.
- 9 Sanz M., Daura J., Subirà E. (2008) La mandíbula de neandertal de la cova del Gegant
10 (Sitges, el Garraf). V *Trobada d'Estudiosos del Garraf*, 131-136.
- 11 Schönswetter P., Tribsch A. (2005) Vicariance and dispersal in the alpine perennial
12 *Bupleurum stellatus* L. (Apiaceae). *Taxon*, **3**, 725–732.
- 13 Schüle W. (1993) Mammals, vegetation and the initial human settlement of the
14 Mediterranean islands: a palaeoecological approach. *Journal of Biogeography*, **20**, 399–
15 412.
- 16 Seo/BirdLife. (2009) IBA Marinas; Áreas Importantes para las Aves (IBA) marinas en
17 España. LIFE04NAT/ES/000049. (<http://www.seo.org>)
- 18 Sorensen A.E. (1986) Seed dispersal by adhesion. *Annual Review of Ecology, Evolution*
19 *and Systematics*, **17**, 443–463.
- 20 Stübing G., Peris J.B., Olivares A., Martín J. (1997) *Cheirolophus mansanetianus*
21 Stübing, Peris, Olivares & Martín, sp. nov. and *Ch. grandifolius* (Font Quer) comb. &
22 stat. nov. (Asteraceae), two endemics from Spain. *Anales del Real Jardín Botánico de*
23 *Madrid*, **55**, 170–173.
- 24 Susanna A. (1989) Mapa 114. *Cheirolophus intybaceus* (Lam.) Dostál. In: Fernández

- 1 Casas J. (Ed.) *Asientos para un Atlas corológico de la flora occidental. Fontqueria*, **22**,
2 16–18.
- 3 Susanna A., Garcia-Jacas N. (2009) Cardueae (Carduoideae). In: Funk V., Susanna A.,
4 Stuessy T.F., Bayer R.J. (Eds.) *Systematics, evolution, and biogeography of*
5 *Compositae*. International Association for Plant Taxonomy: Vienna.
- 6 Tremetsberger K., Talavera S., Stuessy T.F., Angeles-Ortiz M., Weiss-Schneeweiss H.,
7 Kadlec G. (2004) Relationship of *Hypochaeris salzmanniana* (Asteraceae, Lactuceae),
8 an endangered species in the Iberian peninsula, to *H. radicata* and *H. glabra* and
9 biogeographical implications. *Botanical Journal of the Linnean Society*, **146**, 79–95.
- 10 Valente L.M., Savolainen V., Vargas P. (2010). Unparalleled rates of species
11 diversification in Europe. *Proceedings of the Royal Society B*, **277**, 1489–1496.
- 12 Vesica P.L., Tuccimei P., Turi B., Fornós J.J., Ginés A., Ginés J. (2000) Late
13 Pleistocene paleoclimates and sea-level change in the Mediterranean as inferred from
14 stable isotope and U-series studies of overgrowths on speleothems, Mallorca, Spain.
15 *Quaternary Science Reviews*, **19**, 865–879.
- 16 Vitales D. (2010) Implications of molecular systematic analyses on the conservation of
17 rare and threatened taxa: the genus *Cheirolophus* (Asteraceae). Master thesis.
18 Universitat de Barcelona.
- 19 Vogt R. (1991) Die Gattung *Leucanthemum* Mill. (Compositae-Anthemidae) auf der
20 Iberischen Halbinsel. *Monografias del Real Jardín Botánico de Madrid* 10.
- 21 Vos P., Hogers R., Bleeker M., Reijans M., Van de Lee T., Hornes M., Frijters A., Pot
22 J., Peleman J., Kuiper M., Zabeau M. (1995) AFLP: a new technique for DNA finger-
23 printing. *Nucleic Acids Research*, **23**, 4407–4414.
- 24 Weir B.S., Cockerham C.C. (1984) Estimating *F*-statistics for the analysis of
25 populations structure. *Evolution*, **38**, 1358–1370.

- 1 Wright S. (1951) The genetical structure of populations. *Annals of Eugenetics*, **15**, 323–
- 2 324.
- 3

1 Figure captions

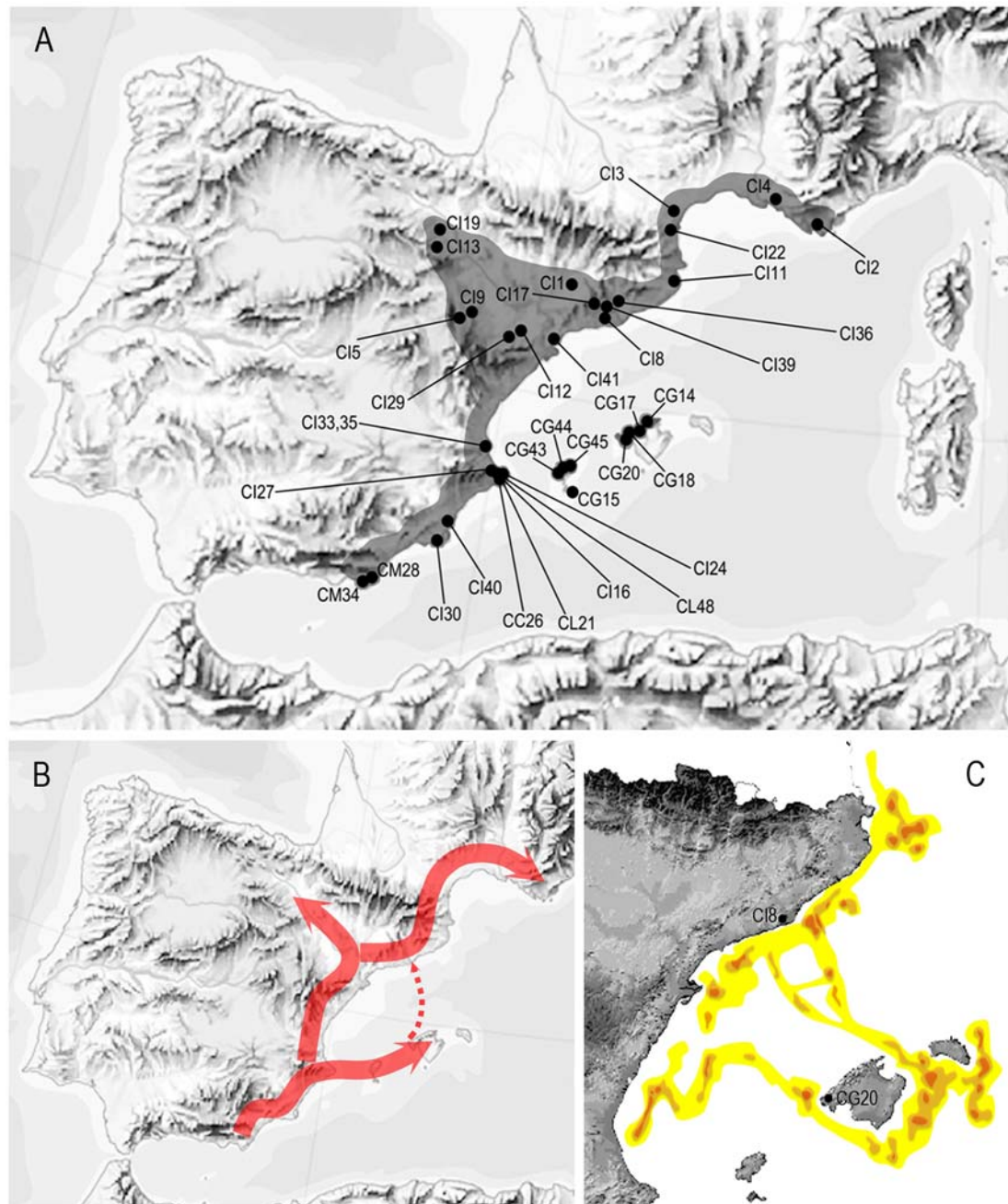
2 Figure 1. A) Distribution range of *Cheirolophus intybaceus s.l.* (shaded area) and
3 geographical location of the studied populations. Population codes correspond to those
4 indicated in Appendix 1. B) Potential colonisation routes of *C. intybaceus* populations
5 from SE Iberia towards (i) the Balearic archipelago in Middle-Upper Pleistocene times
6 (c. 0.4 Ma) through the Dianic-Phytiuses land-bridge connection, and (ii) northwards E
7 Iberia (including the Ebro valley incursion) and SW France. The dotted arrow indicates
8 the potential Late-Pleistocene LDD of *C. intybaceus* seeds from Mallorca to Garraf. C)
9 Frequent movements of the shearwater (*Calonectris diomedea*) from the Balearic islands
10 to the E Iberian coastal area, extracted from SeoBirdLife (2009).

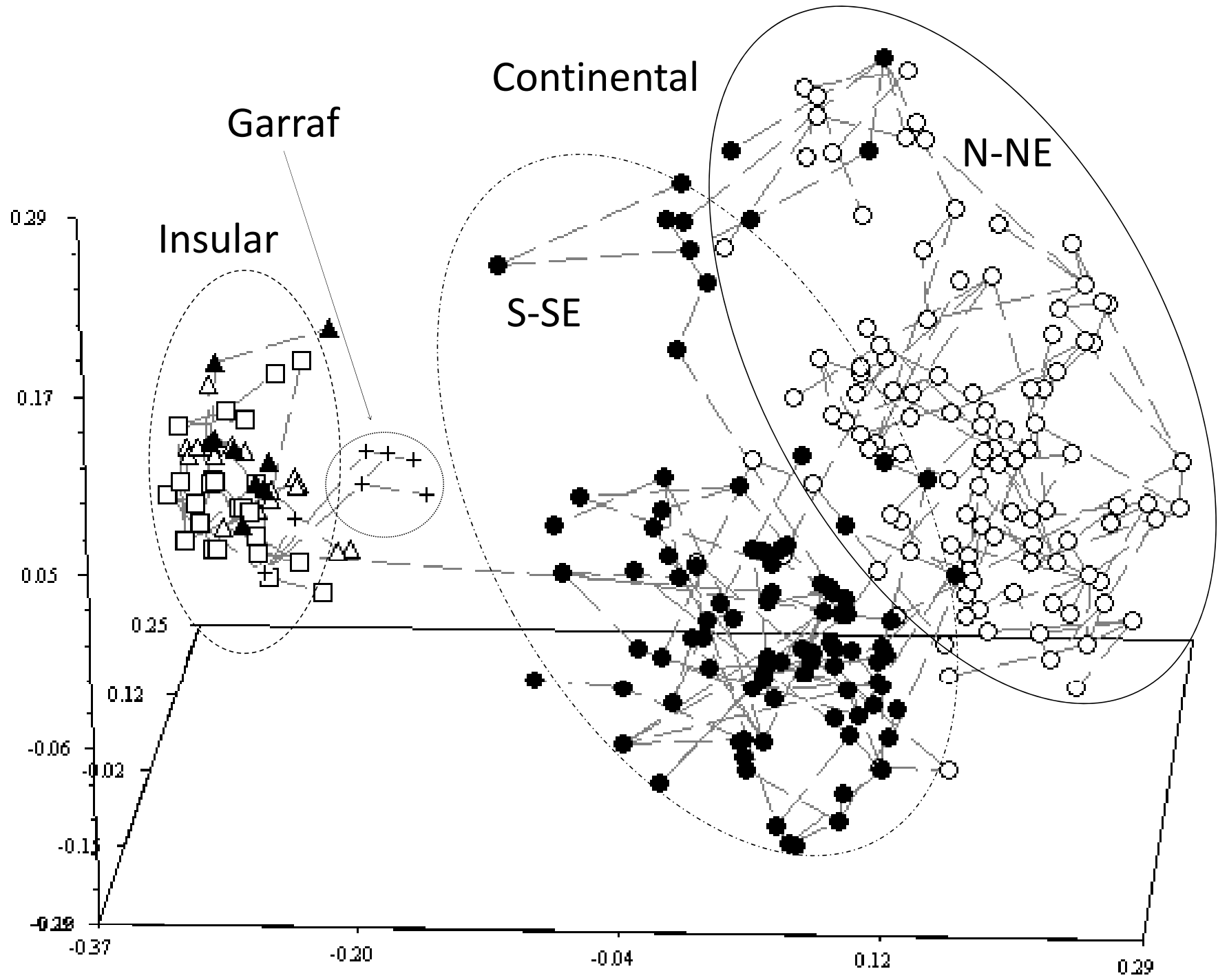
11

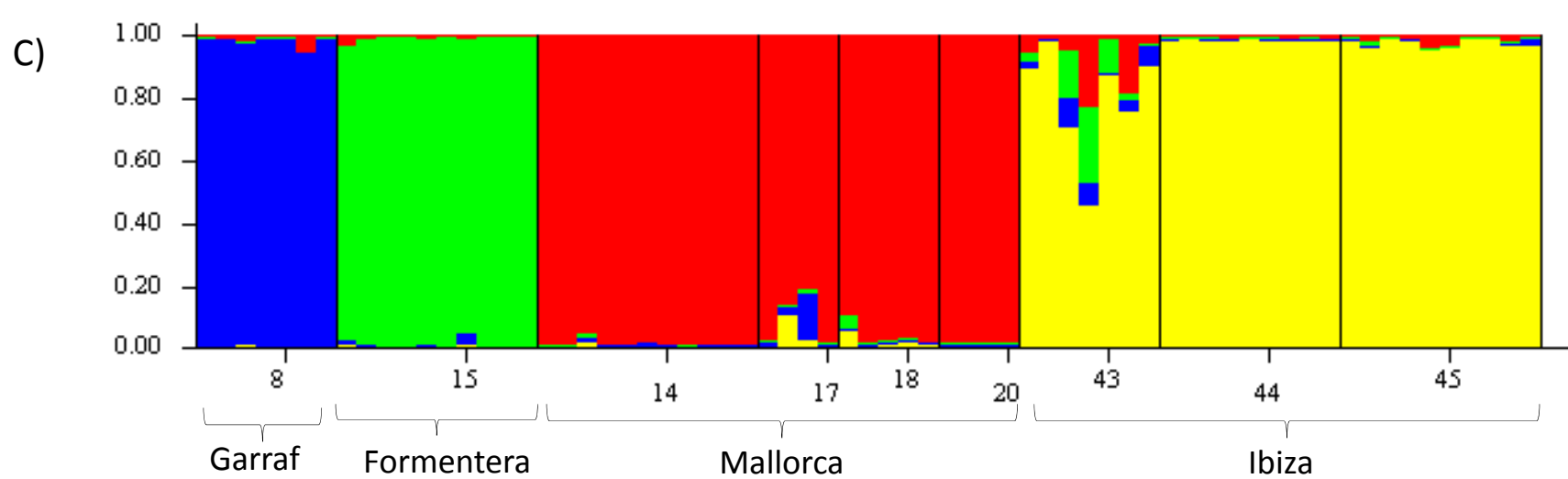
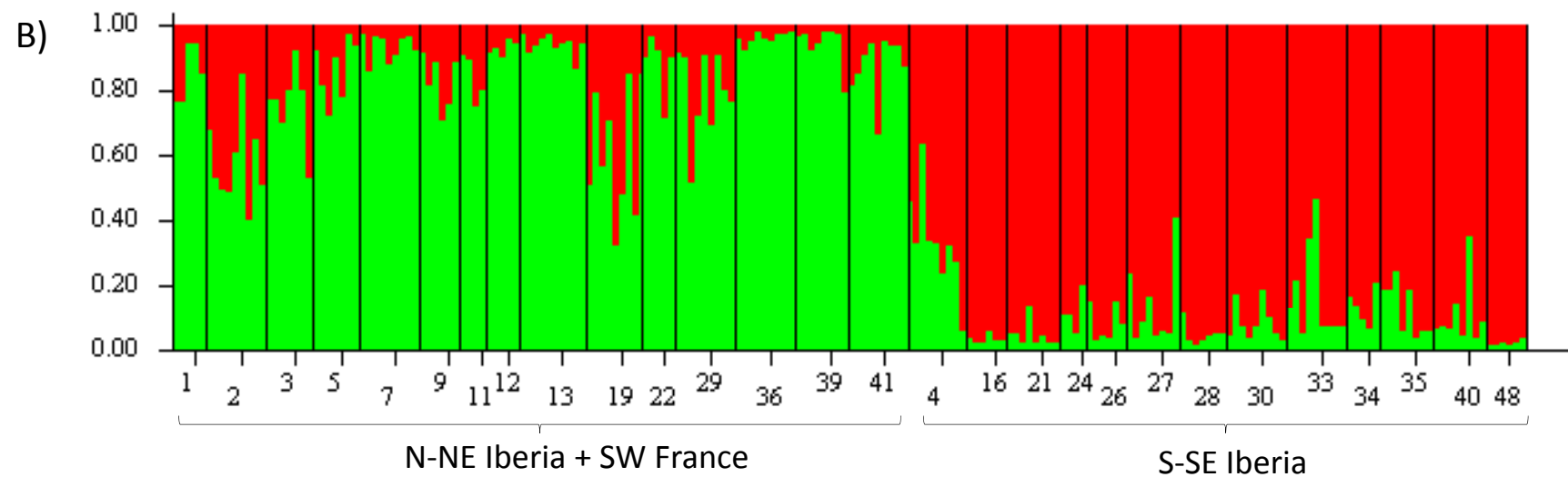
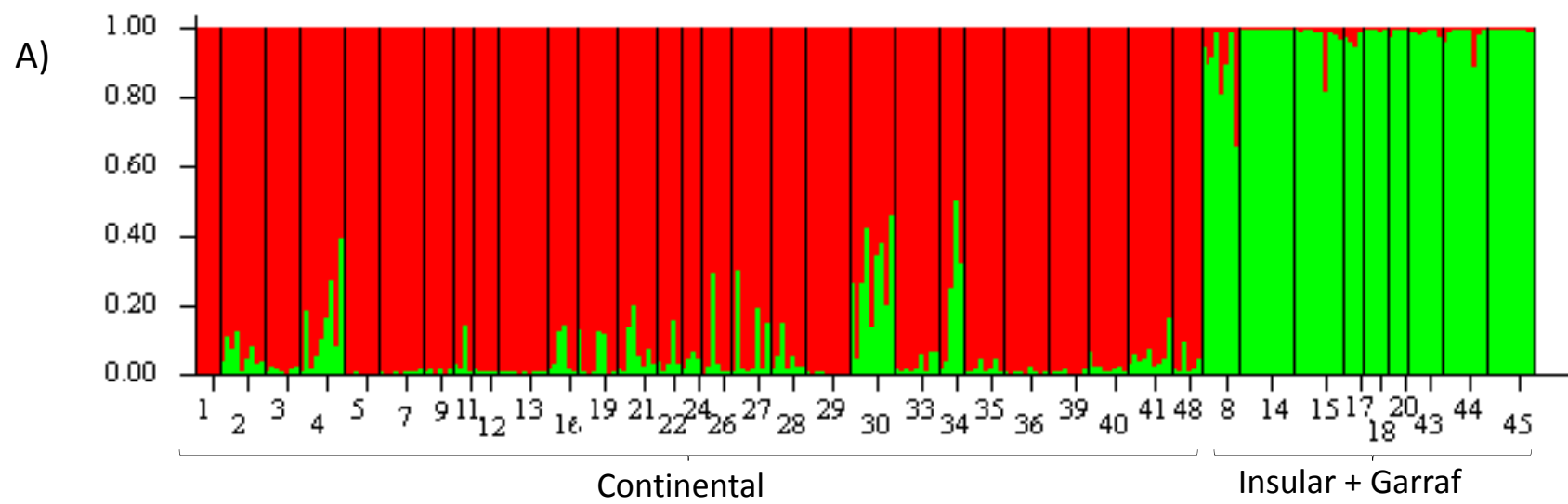
12 Figure 2. Bayesian estimation of genetic structure within *C. intybaceus s.l.* and main
13 subgroups: A. Continental and insular populations ($K = 2$); B. Continental subgroup:
14 Southern (SE + E Iberia) and Northern (NE Iberia + SW France) continental ranges
15 populations ($K = 2$). C. Insular subgroup: Balearic islands (Ibiza, Formentera, Mallorca)
16 plus mainland Garraf populations ($K = 4$). Percentages of individual memberships to
17 predefined groups are shown in colours.

18

19 Figure 3. MST superimposed into a three dimension PCO plot showing the genetic
20 relationships among the *C. intybaceus s.l.* groups. Axis one (17.76%), axis two (6.18%),
21 and axis three (4.44%) accounted for 28.38% of the total variance. *Cheirolophus*
22 *intybaceus* + *Cheirolophus lagunae* + *Cheirolophus mansanetianus* (circles),
23 *Cheirolophus intybaceus* (Garraf population: crosses), *Cheirolophus grandifolius* (Ibiza
24 populations; squares), *Cheirolophus grandifolius* (Formentera + Mallorca population:
25 triangles).







APPENDIX 1

Taxa, locality, code, number of sampled individuals (n), and genetic diversity indexes assessed by AFLP in 41 populations of the *Cheirolophus intybaceus* complex. Genetic indices: Number of rare fragments (f_r); Number of private fragments (f_u); Percentage of polymorphic loci for a standardised sample size of four (PLP(4)1%); Heterozygosity (Hj); Nei genetic diversity index (h); Band richness for a standardised sample size of four (Br (4)); and frequency-down-weighted marker values index (DW).

	Locality	Code	n	f_r	f_u	PLP (4) %	Hj	Nei diversity index (h)	Br (4)	DW	
11	<i>Cheirolophus grandifolius</i>	Spain: Ibiza, Cala Albarca	CG43	7	–	–	0.223	0.129	0.501	1.163	5.041
12	<i>Cheirolophus grandifolius</i>	Spain: Ibiza, Cap Negret	CG44	9	–	1	0.184	0.092	0.517	1.112	6.249
13	<i>Cheirolophus grandifolius</i>	Spain: Ibiza, Cala Xemena	CG45	10	–	–	0.097	0.061	0.479	1.064	6.046
14	<i>Cheirolophus grandifolius</i>	Spain: Formentera, La Mola	CG15	10	–	–	0.136	0.082	0.522	1.085	4.335
15	<i>Cheirolophus grandifolius</i>	Spain: Mallorca, Estellencs 1	CG18	5	–	–	0.019	0.038	0.515	1.035	1.498
16	<i>Cheirolophus grandifolius</i>	Spain: Mallorca, Estellencs 2	CG20	4	–	–	0.000	0.010	0.509	1.010	1.090
17	<i>Cheirolophus grandifolius</i>	Spain: Mallorca, Sa Talaia	CG17	4	–	–	0.000	0.062	0.536	1.068	1.138
18	<i>Cheirolophus grandifolius</i>	Spain: Mallorca, Son Morro	CG14	11	1	–	0.078	0.043	0.496	1.044	4.633
19	<i>Cheirolophus grandifolius</i>	Total		60	13	2	0.092	0.064	0.539	1.073	3.754
20	<i>Cheirolophus intybaceus</i>	Spain: Alicante, Benirrama	CI27	8	–	–	0.194	0.113	0.600	1.141	2.315
21	<i>Cheirolophus intybaceus</i>	Spain: Alicante, Cala de la Barraca	CI24	4	–	–	0.000	0.132	0.593	1.165	1.162
22	<i>Cheirolophus intybaceus</i>	Spain: Alicante, Cap Negre	CI16	6	–	–	0.194	0.172	0.586	1.216	4.298
23	<i>Cheirolophus intybaceus</i>	Spain: Alicante, Dehesa de Campoamor	CI40	8	–	1	0.184	0.106	0.499	1.132	9.400
24	<i>Cheirolophus intybaceus</i>	Spain: Barcelona, Garraf	CI8	7	–	–	0.117	0.182	0.498	1.089	2.492
25	<i>Cheirolophus intybaceus</i>	Spain: Barcelona, La Puda de Montserrat	CI39	8	–	–	0.233	0.130	0.541	1.159	3.172
26	<i>Cheirolophus intybaceus</i>	Spain: Barcelona, Sant Miquel del Fai	CI36	9	–	–	0.165	0.100	0.545	1.123	5.901
27	<i>Cheirolophus intybaceus</i>	Spain: Girona, L'Escala	CI11	4	–	–	0.000	0.181	0.566	1.223	1.333
28	<i>Cheirolophus intybaceus</i>	Spain: Lleida, Arbeca.	CI7	9	–	–	0.291	0.168	0.531	1.196	3.988
29	<i>Cheirolophus intybaceus</i>	Spain: Lleida, Monteró	CI1	5	–	–	0.136	0.181	0.562	1.208	1.976
30	<i>Cheirolophus intybaceus</i>	Spain: Tarragona, Montroig	CI41	9	–	1	0.214	0.085	0.537	1.117	5.165
31	<i>Cheirolophus intybaceus</i>	Spain: Logroño, Santa Eulalia Somera	CI13	10	–	–	0.184	0.115	0.549	1.135	4.377
32	<i>Cheirolophus intybaceus</i>	Spain: Murcia, From la Unión to Portman	CI30	9	–	–	0.359	0.186	0.586	1.245	5.003
33	<i>Cheirolophus intybaceus</i>	Spain: Navarra, Mendavia	CI19	8	–	–	0.320	0.191	0.594	1.252	2.792
34	<i>Cheirolophus intybaceus</i>	Spain: Teruel, Calanda	CI29	9	–	–	0.233	0.129	0.579	1.152	3.385
35	<i>Cheirolophus intybaceus</i>	Spain: Valencia, Cullera 1	CI33	9	–	–	0.233	0.132	0.557	1.175	3.844
36	<i>Cheirolophus intybaceus</i>	Spain: Valencia, Cullera 2	CI35	8	–	–	0.291	0.170	0.575	1.211	9.354
37	<i>Cheirolophus intybaceus</i>	Spain: Zaragoza, Rueda de Jalón	CI9	6	–	–	0.262	0.218	0.598	1.269	2.786
38	<i>Cheirolophus intybaceus</i>	Spain: Zaragoza, La Almunia	CI5	7	–	–	0.359	0.228	0.538	1.274	3.931
39	<i>Cheirolophus intybaceus</i>	Spain: Zaragoza, Mequinenza	CI12	5	–	–	0.078	0.114	0.576	1.130	1.624
40	<i>Cheirolophus intybaceus</i>	France: Aude, La Clape	CI3	7	–	–	0.175	0.124	0.560	1.149	2.200
41	<i>Cheirolophus intybaceus</i>	France : Bouches-du-Rhône, Sormiou	CI4	9	–	–	0.291	0.147	0.554	1.183	3.258
42	<i>Cheirolophus intybaceus</i>	France: Pyrénées Orientales, Tautavel	CI22	5	–	–	0.117	0.135	0.546	1.161	1.664
43	<i>Cheirolophus intybaceus</i>	France: Var, Toulon	CI2	9	–	–	0.320	0.181	0.562	1.215	3.835
44	<i>Cheirolophus intybaceus</i>	Spain: Alicante, Cala Moraig	CC26	6	–	–	0.175	0.154	0.586	1.183	1.948
45	<i>Cheirolophus intybaceus</i>	var. <i>capillifolia</i>									
46	<i>Cheirolophus intybaceus</i>	Total		217	23	12	0.205	0.151	0.618	1.180	3.648
47	<i>Cheirolophus lagunae</i>	Spain: Alicante, Cap Prim	CL21	8	–	–	0.282	0.157	0.601	1.201	2.687
48	<i>Cheirolophus lagunae</i>	Spain: Alicante, La Caleta	CL48	6	–	–	0.097	0.171	0.601	1.214	2.174
49	<i>Cheirolophus lagunae</i>	Total		14	4	–	0.189	0.164	0.611	1.208	2.430
50	<i>Cheirolophus mansanetianus</i>	Spain: Almería, El Sabinar	CM34	5	–	–	0.146	0.183	0.632	1.243	1.327
51	<i>Cheirolophus mansanetianus</i>	Spain: Almería, Rodalquilar	CM28	7	–	–	0.243	0.144	0.600	1.171	1.888
52	<i>Cheirolophus mansanetianus</i>	Total		12	8	–	0.194	0.163	0.631	1.207	1.607
53	All taxa	Total		303				0.134	0.620	1.160	3.495