



**HAL**  
open science

## Phylogeography sheds light on the central–marginal hypothesis in a Mediterranean narrow endemic plant

Marine Pouget, Sami Youssef, Jérémy Migliore, Marianick Juin, Frédéric Médail, Alex Baumel

### ► To cite this version:

Marine Pouget, Sami Youssef, Jérémy Migliore, Marianick Juin, Frédéric Médail, et al.. Phylogeography sheds light on the central–marginal hypothesis in a Mediterranean narrow endemic plant. *Annals of Botany*, 2013, 112 (7), pp.1409-1420. 10.1093/aob/mct183 . hal-03616925

**HAL Id: hal-03616925**

**<https://hal.science/hal-03616925>**

Submitted on 23 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 *Revised manuscript for Annals of Botany Ms12740*

2 **Phylogeography sheds light on the central-marginal hypothesis in a Mediterranean**  
3 **narrow endemic plant**

4 **Authors**

5 Marine Pouget<sup>1</sup>, Sami Youssef<sup>2</sup>, Jérémy Migliore<sup>1</sup>, Marianick Juin<sup>1</sup>, Frédéric Médail<sup>1</sup>, Alex  
6 Baumel<sup>1</sup>.

7

8 <sup>1</sup> *Aix-Marseille Université. Institut Méditerranéen de Biodiversité et d'Ecologie (IMBE,*  
9 *UMR CNRS 7263). Technopôle de l'Environnement Arbois-Méditerranée, BP 80, 13545*  
10 *Aix-en-Provence cedex 04. France.*

11 <sup>2</sup> *Faculty of Agriculture and Forestry. University of Duhok, Kurdistan Region. Iraq.*

12

13 **Abstract**

14 • *Background and Aims* Understanding the factors that shape variation in genetic diversity  
15 across the geographical ranges of species is an important challenge in the effort for  
16 conservation of evolutionary processes sustaining biodiversity. The historical influences  
17 leading to a central-marginal organization of genetic diversity have been explored for species  
18 whose range is known to have expanded from refugia after glacial events. However, this  
19 question has rarely been addressed for Mediterranean endemic plants of azonal habitats such  
20 as rocky slopes or screes. In this context, our comprehensive study examined molecular and  
21 field data from *Arenaria provincialis* (Caryophyllaceae), a narrow endemic plant of south-  
22 eastern France.

23 • *Methods* We investigated, across the whole geographical range, whether high levels of  
24 abundance and genetic diversity (estimated from AFLP) are centrally distributed, to evaluate  
25 the relevance of the central-marginal hypothesis. We used phylogeographic patterns inferred  
26 from chloroplastic DNA (cpDNA), applying Bayesian methods to test the influence of past  
27 biogeographical events. Finally, we used multivariate analysis combining phylogeographic  
28 and ecological data to reveal the historical and ecological distinctiveness of populations.

29 • *Key Results* Despite the narrow distribution of *A. provincialis*, a high level of nucleotidic  
30 variation is found within cpDNA loci, supporting its persistence throughout the Pleistocene  
31 period. The area characterized by the highest genetic diversity is centrally located. Structured  
32 phylogeography and Bayesian factor analysis support the hypothesis that the central area of  
33 the distribution was the source of both westward and eastward migrations, probably during  
34 arid periods of the Pleistocene, and more recently was a crossroads of backward migrations.  
35 By contrast, the two areas located today at the range limits are younger, have reduced genetic  
36 diversity and are marginal in the ecological gradients.

37 • *Conclusions* This study highlights a case of strong population distinctiveness within a  
38 narrow range. Phylogeography sheds light on the historical role of the areas centrally situated  
39 in the distribution. The current range size and abundance patterns are not sufficient to predict  
40 the organization of genetic diversity.

41

42 **Keywords**

43 Genetic diversity, geographical range, ecological niche, pleistocene, Provence,  
44 conservation, migration, population distinctiveness.

45 Running title: Phylogeography of *Arenaria provincialis*

46

48 A fundamental issue in ecology and conservation biology is how evolutionary processes shape  
49 variation across species' whole geographical range (Lesica and Allendorf, 1995; Brown, 1984;  
50 Gaston, 2003; Guo, 2012). According to the “central-marginal” hypothesis (Eckert *et al.*,  
51 2008), genetic diversity decreases and genetic differentiation increases toward the periphery  
52 of the range. The ecological and evolutionary features of populations living at range limits of  
53 species (Sexton *et al.*, 2009) are of particular importance in the context of global change  
54 (Sexton *et al.*, 2011). Low habitat quality at the range margins may lead to higher rates of  
55 extinction, while increased isolation reduces the rate of recolonization as well as of gene flow.  
56 However, the decline in genetic diversity at range limits is not a ubiquitous trend, neither in  
57 multispecific comparisons (Eckert *et al.*, 2008; Eckstein *et al.*, 2006; Hardie and Hutchings,  
58 2010; Guo, 2012), nor within intra-specific case studies (Hoban *et al.*, 2010). As emphasized  
59 by several authors (Garner *et al.*, 2004; Eckert *et al.*, 2008; Moeller *et al.*, 2011), looking for  
60 the causes of a central-marginal structure of genetic diversity is particularly challenging  
61 because of the diverse alternative models and the effects of current demographic and  
62 ecological factors combined with persistent historical influences. Exploring historical  
63 influences, *e.g.* whether marginal populations are younger than core populations (Eckert *et al.*,  
64 2008), has been seen as relevant in the Northern Hemisphere, where Pleistocene climatic  
65 oscillations are still perceptible in the layout of the current patterns of widespread species  
66 distribution, through persistence in refugial areas and post-glacial migration roads (Taberlet *et al.*,  
67 1998; Hewitt, 2000, Petit *et al.*, 2003; Hewitt, 2004; Hampe and Petit, 2005; Soltis *et al.*,  
68 2006). Thus, range shifts and distance from areas ensuring the persistence of populations –  
69 notably refugia (Stewart *et al.*, 2010) – are often invoked to explain the coarse geographical  
70 changes in genetic variation (Garner *et al.*, 2004; Hoban *et al.*, 2010; and references in Hardie  
71 and Hutchings, 2010).

72 Yet neither the historical influences shaping genetic variation, nor the central-marginal  
73 hypothesis, have been invoked in connection with Mediterranean endemic plants. Our aim  
74 here was to explore these questions, for three reasons. Firstly, as underlined recently by Nieto  
75 Feliner (2011), the post-glacial colonization paradigm pattern is insufficient as a basis for  
76 exploring Mediterranean evolutionary patterns and processes during the Pleistocene.  
77 Secondly, in the highly structured landscapes of the Mediterranean, narrow endemic plants are  
78 so frequent (Médail and Verlaque, 1997; Casazza *et al.*, 2005) that they are one of the main  
79 features of Mediterranean biodiversity (Thompson *et al.*, 2005). Thirdly, because their

80 distribution is structured in small areas, narrow endemic plants could make it easier to test the  
81 central-marginal hypothesis at the scale of the whole range; moreover, they constitute key  
82 candidates to study persistence and speciation (Stebbins and Major, 1965; Chown, 1997).  
83 Recent studies of some narrow endemic Mediterranean plants (e.g. Medrano and Herrera,  
84 2008; Molins *et al.*, 2011; Mayol *et al.*, 2012; Nicoletti *et al.*, 2012) support the hypothesis of  
85 their long-term persistence during the Pleistocene and their fluctuating genetic diversity.  
86 Nevertheless, the link between historical influences, geographical range and genetic diversity  
87 in Mediterranean narrow endemic plants remains relatively unelucidated in the context of the  
88 central-marginal hypothesis. To explore this link, we focus on the Caryophyllaceae *Arenaria*  
89 *provincialis* Chater and Hallid. (Chater and Halliday, 1964), which inhabits the calcareous  
90 outcrops of Southern Provence, in south-eastern France.

91 Recent studies on geographical distribution and ecological niches (Véla *et al.*, 2008; Youssef  
92 *et al.*, 2011) provide the basis for examining central and marginal populations of *A.*  
93 *provincialis* throughout its entire range, taking into account both abundance and ecological  
94 aspects. According to the central-marginal hypothesis, populations of *A. provincialis* living  
95 near the centre of the distribution are expected to exhibit more genetic diversity than  
96 populations living further afield (Fig. 1). Moreover, at range limits, populations of *A.*  
97 *provincialis* experience contrasting ecological conditions (Youssef *et al.*, 2011) that could  
98 raise conservation issues if, concurrently, their genetic diversity decreased and their isolation  
99 increased with respect to more central populations.

100 Here, occurrence data are used to show that abundance is not uniformly distributed across the  
101 range of *A. provincialis*, which confirms the importance of exploring the central-marginal  
102 hypothesis. First, we test the consistency of our hypotheses on a central-marginal structure of  
103 genetic diversity using multilocus markers, through analysis of amplified fragment length  
104 polymorphism (AFLP) markers. Second, we infer the historical influences throughout the  
105 distribution from the phylogeography, based on nucleotidic variation of four chloroplastic loci  
106 analysed in a Bayesian framework. Finally, we use ecological data to link marginality in the  
107 ecological gradient with range position and phylogeography, thus revealing the conservation  
108 significance of our results.

109

110

112 *Study system and occurrence data*

113 In south-eastern France, the winter annual plant *Arenaria provincialis* (Chater and Halliday,  
114 1964) is distributed within an area of less than 145 km<sup>2</sup> along an altitudinal gradient and in  
115 different rocky habitats (Véla *et al.*, 2008; Baumel *et al.*, 2009). Its southern limits occur in  
116 the driest context of steep scree slopes at low altitudes and short distances from the coastline,  
117 whereas its northern limits lie in colder conditions near the summits of the outcrops.

118 Variation in abundance of *Arenaria provincialis* throughout its entire range was estimated  
119 from the variation in density of occurrence on samples collected during several field  
120 campaigns. Data were filtered to ensure that occurrences were at least 100 metres apart, to  
121 avoid redundancy and overestimation of abundance due to oversampling. 591 occurrences  
122 were used to model abundance variation by applying an isotropic Gaussian kernel function  
123 (density.ppp function; “spatstat” package; R Development Core Team, 2011).

124 Occurrence data were also used to set discrete, non-overlapping geographical units to  
125 organize measures of genetic diversity and for phylogeography analysis (see below).  
126 Preliminary analyses distinguished 5 geographical units (A-E; Fig. 1a; Table 1). These  
127 geographical units were delineated via non-hierarchical clustering analysis (kmeans function,  
128 “stats” package; R Development Core Team, 2011) from coordinates of *A. provincialis*  
129 occurrences. The barycentre of the distribution is situated in the northern part of area C.

130

131 *DNA extraction and molecular methods*

132 Forty populations of *Arenaria provincialis* were sampled throughout its entire geographical  
133 range (Table 1) between April 2009 and May 2010. Leaves were dried with silica gel, and  
134 then stored at -20°C. Total DNA extraction followed the CTAB protocol of Doyle and Doyle  
135 (1987).

136 The polymorphism of the chloroplast genome was examined on two or three randomly  
137 sampled individuals per population (n = 94) for four plastid regions: internal part of *matK*  
138 *gene*, *trnK* intron plus 5' part of *matK*, *trnL-trnF* and *trnT-trnL* intergenic spacer  
139 (supplementary material Table S1). Sequence characterization of the four chloroplastic  
140 sequences and the combined matrix are reported in Table 2. According to the most recent

141 phylogeny published on this genus (Youssef *et al.*, 2011), *A. provincialis* and *A. cinerea* are  
142 sister species. The same cpDNA regions were sequenced for three samples of *A. cinerea*  
143 sampled at three different locations, as well as one sample of *A. ciliata* and one sample of *A.*  
144 *hispidia* (Table 1; Genbank accession number in supplementary material Table S2).

145 Polymerase Chain Reactions (PCR) were performed in a 50 µl volume reaction that contained  
146 2 µl of each primer (5 mM), 1.25 units of Taq polymerase (Q-Biogen), 5 µl of incubation  
147 buffer, 6 µl of MgCl<sub>2</sub> (10 mM), 4 µl of dNTP (25 mM) and 5 µl of DNA (50 ng.µl<sup>-1</sup>).  
148 Amplification was performed using PTC-200 thermocycler (MJ Research, Inc.; cycling  
149 conditions in supplementary material Table S3). Direct sequencing (ABI 3730 xl, Applied  
150 Biosystems) was carried out by LGC genomics services (Germany). After quality checking  
151 and alignment, the cpDNA data set contained 94 sequences.

152 Variation in genetic diversity throughout the range was investigated by multilocus genotyping  
153 based on the Amplified Fragment Length Polymorphism technique (AFLP) developed by Vos  
154 *et al.* (1995). Two primer combinations for the selective PCR (fluorescent dyes in brackets,  
155 Eurofins MWG Operon, Ebersberg, Germany) were used: EcoRI (6-FAM) - AAGG / MseI  
156 CCAG, EcoRI (6-FAM)-AACG / MseI-CAAG. Selective amplification was performed in  
157 20µL volumes with 5 pmol of each primer, 0.65 mM of MgCl<sub>2</sub>, 0.5 mM of dNTPs, 1 unit of  
158 Taq DNA polymerase (Q-Biogen) and 5 µL of 100x diluted preamplification. The selective  
159 amplification thermocycle profile was 94°C for 2 min, 10 cycles of 94°C for 30 s, 65°C for 30  
160 s (step -0.7°C per cycle), 72°C for 1 min, followed by 20 cycles at 94°C for 30 s, 56°C for 30  
161 s, 72°C for 1 min and 72°C for 5 min. PCR products were separated and quantified on an ABI  
162 3730xl DNA analyzer (Applied Biosystems, Foster City, CA). The reliability of AFLP  
163 markers was checked by repeating the complete analysis from DNA amplification to AFLP  
164 screening on 10 samples for each pair of primers.

165 More than forty polymorphic markers were generated, the minimum threshold for the  
166 resolution of a genetic structure with AFLPs (Gaudeul *et al.*, 2004). Amplified labelled  
167 fragments were revealed on an ABI 3730 DNA Analyzer (PE Applied Biosystems) and  
168 fingerprint profiles were scored with Genemapper 3.7 (PE Applied Biosystems). The results  
169 of the scoring were exported as a presence/absence matrix and used for further analysis. Two  
170 individuals per population were studied for all populations. However, AFLP failed for some  
171 samples, and the final AFLP data set comprised 73 genotypes.

172



173 *Variation in AFLP across the geographical range*

174 Genetic diversity was estimated by the percentage of polymorphic loci, the number of private  
175 markers and the Shannon entropy index, which were computed from AFLP data for each of  
176 the 5 geographical units (A-E) described above (Table 3), using GenAlEx 6.41 (Genetic  
177 Analysis in Excel; Peakall and Smouse, 2006). Genetic differentiation was estimated by the  
178 mean of genotype divergence within each geographical unit. Genotype divergences were  
179 based on jaccard index (dist.binary function; “ade4” package; R Development Core Team,  
180 2011).

181

182 *Phylogeography of A. provincialis*

183 A network of relationships among cpDNA haplotypes was established from a NeighborNet  
184 analysis (Bryant and Moulton, 2004) using SplitsTree4 (Huson and Bryant, 2006). *Arenaria*  
185 *cinerea*, *A. hispida*, and *A. ciliata* were included in the analysis so as to identify the branch  
186 representing the oldest divergence in *A. provincialis*. To compute pairwise distances between  
187 haplotypes, a distance using a GTR + gamma substitution model was chosen according to  
188 jModeltest (Posada, 2008) with empirical estimation of nucleotide frequencies (Fig. 2).

189 The phylogeography of *A. provincialis* was based on the Bayesian MCMC methods of  
190 BEAST 1.7 (Drummond *et al.*, 2012), using all sequences obtained for *A. provincialis* (four  
191 cpDNA genes and n=94 samples). All five geographical units (A-E) were used to set a  
192 geographic state for each sample. The ancestral geographic states were inferred at each node  
193 using the Bayesian Stochastic Search Variable Selection (BSSVS; Lemey *et al.*, 2009) (Fig.  
194 3). Without *a priori* knowledge about the history of *A. provincialis*, we decided not to specify  
195 prior rate for the ancestral geographic states; thus, geographic state transitions (dispersal) are  
196 estimated solely from the genealogical processes. This method was recently used to build  
197 discrete phylogeographies in intra-specific studies (*e.g.* Winkler *et al.*, 2012). We used Bayes  
198 factor (BF) analysis to identify well-supported geographic state transitions with strong  
199 posterior support (*i.e.* BF values >3, Lemey *et al.*, 2009).

200 The maximum clade credibility (MCC) tree produced by BEAST 1.7 was used to map the  
201 ancestral distributions reconstructed for each node via the BSSVS method. We ran the MCMC  
202 analysis for 10 million steps, sampling states every 1000 generations. The coalescent model  
203 for molecular data was set with a GTR + gamma substitution model, an empirical estimation

204 of nucleotide frequencies, a coalescent-constant size tree prior, a random starting tree, a strict  
205 molecular clock and default program prior distribution on other model parameters. The  
206 coalescent model for the discrete geographical data was set with a symmetric substitution  
207 model and the BSSVS was then enabled. Convergence and acceptable mixing of the sampled  
208 parameters were checked using TRACER (part of BEAST) (supplementary material Fig. S1);  
209 all ESS values were higher than 200. TREEANNOTATOR (part of BEAST) was used to  
210 choose the MCC tree after a 10% burn-in.

211 The most probable location of each node was labelled on the MCC tree with the  
212 corresponding geographic state. For the main phylogroups and the root nodes, the distribution  
213 of posterior probabilities of geographic states was also shown to account for uncertainties in  
214 ancestral reconstructions (Fig. 3). The history of migration events within the distribution of *A.*  
215 *provincialis* was synthesized on Figure 4 according to ancestral geographic state  
216 reconstruction.

217

#### 218 *Correlation between phylogeography and ecological gradients*

219 Youssef *et al.* (2011) found that the flora of open rocky habitats in Southern Provence is  
220 organized along two ecological gradients, according to an Outlying Mean Index analysis  
221 (OMI; Dolédec *et al.*, 2000) based upon 623 floristic relevés and 14 variables : altitude, slope,  
222 exposition, mean annual temperature and precipitation (BIOCLIM variables; Hijmans *et al.*,  
223 2005), cover percentage of high and low shrubs, high and low trees, herbaceous plants and  
224 total plant cover, as well as the proportion of rocks, pebbles and gravel. The first gradient is  
225 mainly climatic, related to altitude and distance from the coastline, from thermophilous  
226 coastline habitats to colder and wetter habitats located at high altitude in the study area. The  
227 second gradient is in vegetation cover, from scrublands to scree slopes, measured within a 100  
228 m<sup>2</sup> circular area. This vegetation structure is mainly determined by substrate granulometry and  
229 slope steepness.

230 In our analysis, each location, corresponding to a leaf of the MCC tree of cpDNA sequences,  
231 was reported with respect to its contribution to the two ecological gradients described above  
232 using the dotchart.phylog function (“adephylo” and “ade4” packages; R Development Core  
233 Team, 2011). The contributions were centred; therefore the mean value on each gradient  
234 corresponded to the barycentre of the ecological distribution of *A. provincialis*. Finally, the

235 phylogenetic auto-correlation on the two gradients (N=94) was tested using the Abouheif test  
236 with the abouheif.moran function (“adephylo” package, R Development Core Team, 2011),  
237 based on the Moran statistic and a randomization procedure (1000 permutations) (Pavoine *et*  
238 *al.*, 2008).

239

## 240 RESULTS

### 241 *Abundance variation across the geographical range*

242 The density graph obtained via the Kernel method shows, in different intensities of grey, the  
243 smoothing of occurrence density inside a 500x500 metre plot (Fig. 1b), with a maximum of  
244 17 occurrences. Density is higher and more continuous in the south-western part of the  
245 distribution and lower, as well as more fragmented, in the north-eastern part. The core of  
246 abundance is not in line with the geographic barycentre, but located in the south-western part  
247 (mainly B area). Moreover, abundance decreases on the northern and eastern margins but not  
248 on the southern and western margins.

249

### 250 *Variation in AFLP across the geographical range*

251 When AFLP data were collated according to the 5 geographical units (Table 3), the C area was  
252 distinguished as having the highest number of private markers (13) and the highest Shannon  
253 entropy index (0.33). The B area was also characterized by high diversity, while the lowest  
254 diversity was observed in the A area for all indexes (Table 3). Medium diversity was observed  
255 in both areas D and E. Genetic differentiation, estimated by mean pairwise AFLP genotype  
256 divergence within each geographical unit, was lowest in the C area.

257

### 258 *Phylogeography of A. provincialis*

259 When the four cpDNA sequences for *A. provincialis* (Table 2) were combined, the alignment  
260 reached 2,151 bp in length. Twenty eight haplotypes (H1 to H28; Table 1) were identified for  
261 the 94 samples of *A. provincialis* from 59 variable sites (51 parsimony informative positions),  
262 and 5 indels. The *trnL-trnF* DNA region presents the highest nucleotide diversity (0.011), in  
263 spite of having the fewest number of nucleotides (267 bp) and haplotypes (only 9).

264 Conversely, the longest region, *trnK-matK* (830 bp), is characterized by the highest number of  
265 haplotypes (16) and the lowest nucleotide diversity (0.003).

266 NeighborNet analysis revealed two main phylogroups within *A. provincialis*. The outgroup  
267 taxa (*A. cinerea*, *A. cilliata* and *A. hispida*) are connected to the branch separating these two  
268 main phylogroups (Fig. 2). The geographical distribution of cpDNA haplotypes revealed a  
269 strong phylogeographical signal: the first phylogroup (H1 to H15) was mainly situated in the  
270 northern part of the distribution range (groups D and E), while the second phylogroup (H15 to  
271 H28) was restricted to the south (areas A, B and C). The two phylogroups are separated by the  
272 Huveaune river valley, but a few haplotypes of the first phylogroup were also found in the  
273 southern part of the distribution range (group C, haplotypes H5, H11 and H15).

274 The MCC tree provided by BEAST and based on the 94 samples (Fig. 3) is congruent with  
275 the result of the NeighborNet analysis. The root of the MCC tree is located on the branch  
276 separating the two main phylogroups. The BSSVS method used to reconstruct the ancestral  
277 areas of *Arenaria provincialis* identified area C as ancestral (highest posterior probability for  
278 the root, Fig. 3). After the initial divergence of the two phylogroups, the history of *A.*  
279 *provincialis* was characterized by migrations to E and B areas. The first phylogroup originated  
280 from area E and migrated eastwards to area D. The second phylogroup originated from area  
281 B, migrated westwards to area A and then migrated back to area C. There were at least three  
282 dispersal events back towards former areas (e.g. from D to E (H4, H5), from E to C (H5), or  
283 from B to C; Fig. 3 and Fig. 4), but no migration event from the southern areas to the north  
284 occurred after the divergence of the phylogroups. BF analysis revealed that the geographic  
285 state transitions with strongest posterior support were between C and B (BF=24) and between  
286 C and E (BF= 22).

287

### 288 *Correlation between ecological and phylogenetic structures*

289 The Abouheif tests exhibit significant phylogenetic autocorrelation on the two ecological  
290 gradients provided by OMI analysis for all haplotypes (p value = 0.001) as well as for the two  
291 phylogroups separately (p value = 0.001). Descriptive results of the Abouheif tests are given  
292 as supplementary material (Table S4). Fig. 5 shows that samples closely related in phylogeny  
293 tend to share the same position in the ecological gradients.

294 Across the first gradient, samples from sites with above average climatic conditions (first  
295 OMI axis) are mainly in the first phylogroup, while those below average are mainly in the  
296 second phylogroup. In the former areas of both phylogroups, *i.e.* C and E (H16, H17, H18,  
297 H1, H2, and H3), climatic conditions differed but were not extreme. The most marginal  
298 values, corresponding to either the coldest and wettest or the warmest and driest climatic  
299 conditions, were for locations at the periphery, where the most recent haplotypes were found  
300 (D versus A areas) and for which migrations from former areas (B, C and E) were inferred  
301 (see above results).

302 Across the second gradient, that of vegetation cover, the distribution of samples was also  
303 phylogenetically structured. Samples belonging to the second phylogroup are mainly  
304 distributed in sites characterized by low vegetation cover, high slope values and high gravel  
305 cover values (corresponding to screes), whereas samples belonging to the first phylogroup are  
306 mainly distributed in sites with low vegetation and low gravel cover values and high rock  
307 cover values (representing lapiaz). Again, the most marginal values were observed in areas A,  
308 B and D, while values for the C and E areas were at the centre of the gradient.

309

310

## DISCUSSION

311 A hotspot of biodiversity, the Mediterranean Basin has been described as a mosaic of refugial  
312 areas harboring numerous endemic taxa (Médail and Diadema, 2009). But few studies have  
313 investigated precisely the phylogeography of narrow endemic plants, especially outside the  
314 major regional Mediterranean biodiversity hotspots (Médail and Quézel, 1997). Furthermore,  
315 annual plants are generally not well represented in phylogeographical studies. Therefore our  
316 study focusing on a winter annual plant of southern Provence, a region that remains relatively  
317 unexplored in molecular studies, represents a contribution towards more comprehensive  
318 knowledge of plant evolution in the Mediterranean region.

319

### 320 *Unexpected chloroplast DNA variation*

321 In a recent study of its ecological niche and phylogenetic position (Youssef *et al.*, 2011),  
322 *Arenaria provincialis* was shown to be the only one of the six *Arenaria* species of Provence  
323 able to colonize scree slopes. This suggests a probable relationship between its history and the

324 emergence of lapiaz and scree slopes derived from the intense frost erosion processes  
325 occurring during the Pleistocene glacial periods.

326 Here, NeighborNet analysis based on the combined sequences of four cpDNA loci detected a  
327 well-resolved genealogy with two phylogroups in *A. Provincialis*, themselves connected to  
328 the outgroups *A. cinerea*, *A. hispida* and *A. ciliata* (Fig. 2). After the divergence of the two  
329 main phylogroups of *A. provincialis*, the cpDNA haplotypic diversity accumulated, leading to  
330 the current high numbers of haplotypes (28) observed within a small area (less than 145 km<sup>2</sup>).  
331 In comparison, a study based on *Arenaria humifusa*, endemic to the Arctic Arc, showed only  
332 two substitutions and one indel for 1,578 bp in two cpDNA loci, revealing three haplotypes  
333 distributed over more than 5,000 km<sup>2</sup> (Westergaard *et al.*, 2010).

334 Our study, like other Mediterranean case studies [*e.g.* in insular endemics, Molins *et al.*, 2011;  
335 Mayol *et al.*, 2012], confirms that strong genetic structuring is plausible in any part of the  
336 distribution range of Mediterranean endemics, and therefore that the size of a distribution  
337 range is not correlated with levels and structure of genetic diversity. The well-known climatic  
338 impact of the last glacial maximum (21,000 years bp) on species distributions (Hewitt, 2000)  
339 has also been shown for vegetation in the western Mediterranean area (Cheddadi *et al.*, 2009;  
340 Médail and Diadema, 2009; Fady and Connors, 2010). However, the high cpDNA diversity  
341 and polymorphism of *A. provincialis* we observed in such a restricted area support the idea  
342 that the preservation and accumulation of genetic diversity during the Ice Ages could be a  
343 common pattern in the Mediterranean, as underlined recently by Nieto Feliner (2011). Our  
344 study yields new evidence that extreme or azonal habitats, *i.e.* rocky terrain, cliffs and steep  
345 slopes, may have promoted the persistence of endemic plants and enhanced specific traits  
346 (Thompson 2005). *Arenaria provincialis*, with its phylogeographic structure, therefore affords  
347 an excellent opportunity to relate long-term persistence and patterns of genetic diversity to the  
348 central-marginal hypothesis (Eckert *et al.*, 2008; Sexton *et al.*, 2009).

349

#### 350 *Spatial distribution and the central-marginal hypothesis*

351 The kernel model of density (Fig. 1b) clearly shows that abundance of *Arenaria provincialis*  
352 is not homogeneous across the geographical range. Area B has the highest abundance. In the  
353 northern part of its distribution, *A. provincialis* is markedly less abundant and more  
354 fragmented. A bias in field inventories is unlikely to explain this pattern, since inventories of

355 *A. provincialis* have been made for several years and its northern and eastern range limits  
356 have been intensively investigated. The north-eastern parts of the distribution correspond to  
357 isolated patches of open rocky habitats most frequently embedded in matorral or forest  
358 matrixes. A decrease in abundance is observed towards the northern and eastern range limits.

359 Genetic diversity showed a different pattern from that of contemporary abundance: measured  
360 by the polymorphism and Shannon entropy of AFLP markers (Table 3), it increased towards  
361 the centre (areas B and C), whereas AFLP genotype divergence, used here to describe genetic  
362 differentiation, was significantly lower in area C (Table 3). Molecular markers thus show that  
363 the populations of *A. provincialis* with the highest diversity and that are the least differentiated  
364 are near the central part of the distribution (mostly C and secondary B areas), whereas  
365 marginal populations with the lowest genetic diversity are in areas A, D, and E, near the  
366 western, eastern and northern parts of the distribution respectively. The structure of genetic  
367 diversity seems roughly consistent with the hypotheses formulated here (see introduction),  
368 and we cannot rule out the contemporary effects of isolation and small effective population  
369 size, notably at the northern and eastern limits. However, we make the point below that this  
370 pattern of contemporary genetic diversity is also well correlated with Pleistocene range  
371 dynamics and expansion.

372

### 373 *Past migrations from the centre to the periphery*

374 The structured phylogeography of *A. provincialis* and the reconstruction of ancestral areas,  
375 based on the BSSVS method of Lemey *et al.* (2009), enabled us to reconstruct a step-by-step  
376 history of the migrations responsible for the current range of *A. provincialis*. The ancestral  
377 area of *A. provincialis* was thus found to be situated near the centre of the distribution (C  
378 area), with evidence that after migrations to the west (B area) and north (E area), migrations  
379 occurred toward the eastern and western periphery (A and D areas, Fig. 3 and Fig. 4).  
380 Moreover, phylogeography and the reconstruction of ancestral states revealed that backward  
381 migrations also occurred to C and E areas (Fig. 3 and Fig. 4). BF analysis revealed strong  
382 posterior support for dispersal between C and B and between E and C areas, thus confirming  
383 central position C as both an ancestral and a crossroads area. In this phylogeographic context,  
384 it should be noted that the frequency of AFLP private markers (Table 3) supported a longer  
385 persistence of *A. provincialis* in area C, and secondarily in area B. Consequently, area C,

386 which is central in the distribution, appears to have been crucial in the history of *A.*  
387 *provincialis*.

388 The history of *A. provincialis* therefore reveals that the central-marginal pattern is strongly  
389 associated with a history of expansion from former areas (Fig. 4). Such correlations between  
390 genetic diversity and distance from refugia have been found elsewhere (e.g. Garner *et al.*,  
391 2004; Hoban *et al.*, 2010), but over a coarser scale and in the context of glaciations. To the  
392 best of our knowledge, this correlation has never before been observed for a Mediterranean  
393 narrow endemic plant. In the light of the extensive cpDNA nucleotidic variation and marked  
394 phylogeographic structure revealed here (28 haplotypes in two phylogroups), it is likely that  
395 the rocky habitats situated in the central parts (C, and thereafter B and E areas) of the  
396 distribution range have sheltered *A. provincialis* for a long time (*i.e.* probably since the early  
397 Pleistocene). Originating from these central-ancestral areas, the populations situated today at  
398 the range limits (A and D areas) are younger and have more limited AFLP diversity (Table 3).

399 An extensive survey of the literature yielded no conclusive evidence from recent geological  
400 history of the focus area that might have helped to calibrate precisely in time the haplotype  
401 genealogy of *A. provincialis*. Regarding past environmental change, paleoecological data are  
402 also very scarce for rocky outcrops of the focus area, and ancestral distribution modelling  
403 would be difficult to realize for a narrow endemic plant living in azonal habitats. Of the plants  
404 common to the rocky habitats of *A. provincialis*, the pioneer tree *Juniperus phoenicea* has  
405 been hypothesised to have undergone expansion during arid and cold periods of the  
406 Pleistocene, such as the Last Glacial Maximum (Boratynski *et al.*, 2009). Similarly, it has  
407 been proposed that *Cheirolophus intybaceus* reached southern France from the Iberian  
408 Peninsula during the mid-late Pleistocene (Garnatje *et al.*, 2013). The sole relevant  
409 hypothesis, therefore, is that a decrease in vegetation cover due to climate cooling and drying  
410 during the major glaciation phases beginning around 0.9 Mya (De Beaulieu *et al.*, 2005) may  
411 have favored the expansion of pioneer plants like *A. provincialis*.

412

#### 413 *Conservation significance of the phylogeographic pattern*

414 The habitat of *A. provincialis* is structured around two main ecological gradients mainly  
415 determined by altitude and vegetation cover (Baumel *et al.*, 2009; Youssef *et al.*, 2011; Fig.  
416 5A). As supported by the significant phylogenetic autocorrelation on these two ecological



417 gradients (Abouheif test), the populations having the most divergent haplotypes are now  
418 living at the extremes of its two ecological gradients (Fig. 5C). During the history of its  
419 expansion, *A. provincialis* reached the top of its elevation gradient in southern Provence (*i.e.* a  
420 range of 1,000 metres), where its populations are currently experiencing contrasting climatic  
421 conditions (Axis 1, Fig. 5A and Fig. 5C). For example, *A. provincialis* is the sole *Arenaria*  
422 species living both near the littoral and near the summit of the Sainte Baume mountain range  
423 (almost 1,000m above sea level), in contrast to its congeners *A. grandiflora* and *A. aggregata*  
424 which are present only in the second location. Yet populations located near the littoral, and  
425 belonging mainly to the second phylogroup, are living in the most open and steep habitats of  
426 the species, such as huge south-facing seaside screes (Axis 2, Fig. 5 A, C). This means that  
427 today's populations of *A. provincialis* are organized across an ecological range explained both  
428 by the history of the species and by the geographical structure of its genetic diversity.

429 In conclusion, population distinctiveness based on molecular variation and ecological data is  
430 well supported here: populations from the two phylogroups are not genetically, nor  
431 ecologically, exchangeable (Crandall *et al.*, 2000). Further study will be needed in order to  
432 evaluate the probable adaptive polymorphism of *A. provincialis* suggested by the strong  
433 correlation between phylogeography and ecological gradients (Moritz, 2002). Our results  
434 point to key issues in terms of conservation. Firstly, the two phylogroups of *A. provincialis*  
435 have different ecologies, and exhibit different patterns of fragmentation and abundance (Fig.  
436 1, Fig. 5). The populations situated near the south-western and north-eastern range limits  
437 contribute significantly to this fragmented structure. Secondly, the central populations have  
438 persisted the longest and been the source of expansions, but they are not currently  
439 characterized by the highest abundance.

440 Exploring the abundance, genetic diversity and phylogeography of a Mediterranean narrow  
441 endemic plant, this study highlights the fact that current range size and abundance patterns are  
442 not sufficient to predict the organization of genetic diversity. However, our study reveals the  
443 crucial role played in *A. provincialis*' history by the areas centrally situated in its distribution.

444

445 Acknowledgements: We thank the Conservatoire Botanique National Méditerranéen de  
446 Porquerolles for providing unpublished data about *Arenaria* taxa locations and allowing us to  
447 collect *Arenaria* material. We are grateful to Errol Véla for his help on the field and botanical  
448 assessment. Our thanks to Marjorie Sweetko for English language editing. This research was

449 funded by the CNRS, the French National Office of Forests (ONF) and the General council of  
450 Bouches du Rhône district (CG13). Marine Pouget was supported by the Provence–Alpes–  
451 Côte d’Azur region (PACA).

452

- 454 **Baumel A, Affre L, Véla E, Auda P, Torre F, Youssef S, Tatoni T. 2009.** Ecological  
455 magnitude and fine scale dynamics of the Mediterranean narrow endemic therophyte,  
456 *Arenaria provincialis* (Caryophyllaceae). *Acta Botanica Gallica* **156**: 259–272.
- 457 **Boratynski A, Lewandowski A, Boratynska K, Montserrat JM, Romo A. 2009.** High level  
458 of genetic differentiation of *Juniperus phoenicea* (Cupressaceae) in the Mediterranean region:  
459 geographic implications. *Plant systematic and Evolution* **277**:163–172.
- 460 **Brown JH. 1984.** On the relationship between abundance and distribution of species. *The*  
461 *American Naturalist* **124**: 255.
- 462 **Bryant D, Moulton V. 2004.** Neighbor-Net: an agglomerative method for the construction of  
463 phylogenetic networks. *Molecular biology and evolution* **21**: 255–65.
- 464 **Casazza G, Barberis G, Minuto L. 2005.** Ecological characteristics and rarity of endemic  
465 plants of the Italian Maritime Alps. *Biological Conservation* **123**: 361–371.
- 466 **Chater AO, Halliday G. 1964.** *Arenaria* L. In: Tutin TG, Heywood VH, Burges NA,  
467 Valentine DH, Walters SM, Webb DA, eds. *Flora Europaea 1*. Cambridge University Press,  
468 Cambridge, 140–148.
- 469 **Cheddadi R, Fady B, François L, Hajar L, Suc J-P, Huang K, Demarteau M, Vendramin**  
470 **G, Ortu E. 2009.** Putative glacial refugia of *Cedrus atlantica* deduced from Quaternary pollen  
471 records and modern genetic diversity. *Journal of Biogeography* **36**: 1361–1371.
- 472 **Chown SL. 1997.** Speciation and rarity: separating cause from consequence. In: Kunin WE,  
473 Gaston KJ, eds. *The biology of rarity*. Chapman & Hall, London, 91–109.
- 474 **Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK. 2000.** Considering  
475 evolutionary processes in conservation biology. *TREE* **15**: 290–295.
- 476 **De Beaulieu J-L, Miras Y, Andrieu-Ponel V, Guiter F. 2005.** Vegetation dynamics in north-  
477 western Mediterranean regions: Instability of the Mediterranean bioclimate. *Plant Biosystems*  
478 **139**: 114–126.
- 479 **Dolédec S, Chessel D, Gimaret-Carpentier C. 2000.** Niche separation in community  
480 analysis: a new method. *Ecology* **81**: 2914–2927.

481 **Doyle JJ, Doyle JL. 1987.** A rapid DNA isolation method for small quantities of fresh tissues.  
482 *Phytochemical bulletin* **19**: 11–15.

483 **Drummond AJ, Rambaut A. 2007.** BEAST: Bayesian evolutionary analysis by sampling  
484 trees. *BMC evolutionary biology* **7**: 214.

485 **Drummond AJ, Suchard M a, Xie D, Rambaut A. 2012.** Bayesian phylogenetics with  
486 BEAUti and the BEAST 1.7. *Molecular biology and evolution* **29**: 1969–73.

487 **Eckert CG, Samis KE, Loughheed SC. 2008.** Genetic variation across species' geographical  
488 ranges: the central-marginal hypothesis and beyond. *Molecular ecology* **17**: 1170–88.

489 **Eckstein RL, O'neill RA, Danihelka J, Otte A, Köhler W. 2006.** Genetic structure among  
490 and within peripheral and central populations of three endangered floodplain violets.  
491 *Molecular ecology* **15**: 2367–79.

492 **Fady B, Conord C. 2010.** Macroecological patterns of species and genetic diversity in  
493 vascular plants of the Mediterranean basin. *Diversity and Distributions* **16**: 53–64.

494 **Garnatje T, Perez-Collazos E, Pellicer J, Catalan P. 2013.** Balearic insular isolation and  
495 large continental spread framed the phylogeography of the western Mediterranean  
496 *Cheirolophus intybaceus s.l.* (Asteraceae). *Plant Biology* **15**: 166–175.

497 **Garner TWJ, Pearman PB, Angelone S. 2004.** Genetic diversity across a vertebrate species'  
498 range: a test of the central-peripheral hypothesis. *Molecular ecology* **13**: 1047–53.

499 **Gaston KJ. 2003.** *The structure and dynamics of geographic range*. Oxford, UK: Oxford  
500 University Press

501 **Gaudeul M, Till-Bottraud I, Barjon F, Manel S. 2004.** Genetic diversity and differentiation  
502 in *Eryngium alpinum* L. (Apiaceae): comparison of AFLP and microsatellite markers.  
503 *Heredity* **92**: 508–18.

504 **Guo Q. 2012.** Incorporating latitudinal and central-marginal trends in assessing genetic  
505 variation across species ranges. *Molecular ecology* **21**: 5396–403.

506 **Hampe A, Petit RJ. 2005.** Conserving biodiversity under climate change: the rear edge  
507 matters. *Ecology letters* **8**: 461–7.

508 **Hardie DC, Hutchings JA. 2010.** Evolutionary ecology at the extremes of species' ranges.  
509 *Environmental Reviews* **18**: 1–20.

510 **Hewitt GM. 2000.** The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–13.

511 **Hewitt GM. 2004.** Genetic consequences of climatic oscillations in the Quaternary.  
512 *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **359**:  
513 183–95.

514 **Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005.** Very high resolution  
515 interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**:  
516 1965–1978.

517 **Hoban SM, Borkowski DS, Brosi SL, McCleary TS, Thompson LM, McLachlan JS,  
518 Pereira MA, Schlarbaum SE, Romero-Severson J. 2010.** Range-wide distribution of  
519 genetic diversity in the North American tree *Juglans cinerea*: a product of range shifts, not  
520 ecological marginality or recent population decline. *Molecular ecology* **19**: 4876–91.

521 **Huson DH, Bryant D. 2006.** Application of phylogenetic networks in evolutionary studies.  
522 *Molecular biology and evolution* **23**: 254–67.

523 **Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009.** Bayesian phylogeography finds  
524 its roots. *PLoS computational biology* **5**: e1000520.

525 **Lesica P, Allendorf FW. 1995.** When are peripheral populations valuable for conservation□?  
526 *Conservation Biology* **9**: 753–760.

527 **Mayol M, Palau C, Rosselló JA, González-Martínez SC, Molins A, Riba M. 2012.**  
528 Patterns of genetic variability and habitat occupancy in *Crepis triasii* (Asteraceae) at different  
529 spatial scales: insights on evolutionary processes leading to diversification in continental  
530 islands. *Annals of botany* **109**: 429–41.

531 **Médail F, Quézel P. 1997.** Hot-spots analysis for conservation of plant biodiversity in the  
532 Mediterranean basin. *Annals of the Missouri Botanical Garden* **84**: 112–127.

533 **Médail F, Diadema K. 2009.** Glacial refugia influence plant diversity patterns in the  
534 Mediterranean Basin. *Journal of Biogeography* **36**: 1333–1345.

535 **Médail F, Verlaque R. 1997.** Ecological characteristics and rarity of endemic plants from  
536 southeast France and Corsica: implications for biodiversity conservation. *Biological*  
537 *Conservation* **80**: 269–281.

538 **Medrano M, Herrera CM. 2008.** Geographical structuring of genetic diversity across the  
539 whole distribution range of *Narcissus longispathus*, a habitat-specialist, Mediterranean narrow  
540 endemic. *Annals of botany* **102**: 183–94.

541 **Moeller DA, Geber MA, Tiffin P. 2011.** Population genetics and the evolution of geographic  
542 range limits in an annual plant. *The American naturalist* **178 Suppl** : S44–57.

543 **Molins A, Bacchetta G, Rosato M, Rosselló JA, Mayol M. 2011.** Molecular  
544 phylogeography of *Thymus herba-barona* ( Lamiaceae ): Insight into the evolutionary history  
545 of the flora of the western Mediterranean islands. *Taxon* **60**: 1295–1305.

546 **Moritz C. 2002.** Strategies to protect biological diversity and the evolutionary processes that  
547 sustain it. *Systematic biology* **51**: 238–254.

548 **Nicoletti F, De Benedetti L, Airò M, Ruffoni B, Mercuri A, Minuto L, Casazza G. 2012.**  
549 Spatial Genetic Structure of *Campanula sabatia*, a Threatened Narrow Endemic Species of  
550 the Mediterranean Basin. *Folia Geobotanica* **47**: 249–262.

551 **Nieto Feliner G. 2011.** Southern European glacial refugia: A tale of tales. *Taxon* **60**: 365–372.

552 **Pavoine S, Ollier S, Pontier D, Chessel D. 2008.** Testing for phylogenetic signal in  
553 phenotypic traits: new matrices of phylogenetic proximities. *Theoretical population biology*  
554 **73**: 79–91.

555 **Peakall R, Smouse PE. 2006.** GenAlEx 6: genetic analysis in Excel. Population genetic  
556 software for teaching and research. *Molecular Ecology Notes* **6**: 288–295.

557 **Petit J, Aguinagalde I, De Beaulieu J-L, Bittkau C, Brewer S, Cheddadi Rachid, Ennos**  
558 **R, Fineschi Silvia, Grivet D, Lascoux M, Mohanty A, Mu G, Demesure-musch B, Palme**  
559 **A, Marti JP, Rendell S, Vendramin Giovanni G. 2003.** Glacial refugia: hotspots but not  
560 melting pot of genetic diversity. *Science* **300**: 1563–1565.

561 **Posada D. 2008.** jModelTest: phylogenetic model averaging. *Molecular biology and*  
562 *evolution* **25**: 1253–6.

563 **R Development Core Team. 2011.** R: A language and environment for statistical computing.  
564 R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL  
565 <http://www.R-project.org/>.

566 **Sagarin RD, Gaines SD. 2002.** The “abundant centre” distribution: to what extent is it a  
567 biogeographical rule? *Ecology Letters* **5**: 137–147.

568 **Sexton JP, McIntyre PJ, Angert AL, Rice KJ. 2009.** Evolution and ecology of species range  
569 limits. *Annual Review of Ecology, Evolution, and Systematics* **40**: 415–436.

570 **Sexton JP, Strauss SY, Rice KJ. 2011.** Gene flow increases fitness at the warm edge of a  
571 species’ range. *Proceedings of the National Academy of Sciences of the United States of*  
572 *America* **108**: 11704-11709

573 **Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS. 2006.** Comparative  
574 phylogeography of unglaciated eastern North America. *Molecular ecology* **15**: 4261–93.

575 **Stebbins GL, Major J. 1965.** Endemism and speciation in the california flora. *Ecological*  
576 *Monographs* **35**: 1–35.

577 **Stewart JR, Lister AM, Barnes I, Dalén L. 2010.** Refugia revisited: individualistic  
578 responses of species in space and time. *Proceedings of the Royal Society / Biological sciences*  
579 **277**: 661–71.

580 **Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F. 1998.** Comparative phylogeography  
581 and postglacial colonization. *Molecular Ecology* **7**: 453–464.

582 **Thompson JD, Lavergne S, Affre L, Gaudeul M. 2005.** Ecological differentiation of  
583 Mediterranean endemic plants. **54**: 967–976.

584 **Véla E, Auda P, Léger JF, Gonçalves V, Baumel A. 2008.** Exemple d’une nouvelle  
585 évaluation du statut de menace suivant les critères de l’UICN version 3.1. : le cas de  
586 l’endémique provençale *Arenaria provincialis* Chater & Halliday (= *Gouffeia arenarioides*  
587 DC., Caryophyllaceae). *Acta Botanica Gallica* **155** : 547-562.

588 **Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J,**  
589 **Peleman J, Kuiper M. 1995.** AFLP: a new technique for DNA fingerprinting. *Nucleic acids*  
590 *research* **23**: 4407–14.

591 **Westergaard KB, Alsos IG, Popp M, Engelskjøn T, Flatberg KI, Brochmann C. 2010.**  
592 Glacial survival may matter after all: nunatak signatures in the rare European populations of  
593 two west-arctic species. *Molecular ecology* **20**: 376–93.

594 **Winkler M, Tribsch A, Schneeweiss GM, Brodbeck S, Gugerli F, Holderegger R, Abbott**  
595 **RJ, Schönswetter P. 2012.** Tales of the unexpected: phylogeography of the arctic-alpine  
596 model plant *Saxifraga oppositifolia* (Saxifragaceae) revisited. *Molecular ecology* **21**: 4618–  
597 4630.

598 **Youssef S, Baumel A, Véla E, Juin M, Dumas E, Affre L, Taton T. 2011.** Factors  
599 Underlying the Narrow Distribution of the Mediterranean Annual Plant *Arenaria provincialis*  
600 (Caryophyllaceae). *Folia Geobotanica* **46**: 327–350.

601



602 FIGURES AND TABLES

603

604 Fig. 1. (a) Geographical distribution of the 28 haplotypes detected for the 94 individuals of  
605 *Arenaria provincialis* combining four cpDNA regions (internal part of *matK* gene, *trnK* intron  
606 plus 5' part of *matK*, *trnL-trnF* and *trnT-trnL* intergenic spacer for a total of 2,150 bp).  
607 Haplotypes are clustered by geographical groups (from A to E). For population codes see  
608 Table 1. Crosses represent the 156 occurrences of *Arenaria provincialis* used in Youssef *et al.*  
609 (2011).

610 (b) Abundance of *Arenaria provincialis* estimated by an isotropic gaussian kernel function.  
611 591 occurrences were used, at least 100 metres apart (see Materials and Methods for more  
612 details). The maximum abundance was 17 occurrences within a 500x500 metre plot.

613 Fig. 2. NeighborNet diagram generated using SplitsTree4 connecting the 28 haplotypes  
614 detected for the 94 individuals of *Arenaria provincialis* combining the polymorphism of four  
615 cpDNA locus (internal part of *matK* gene, *trnK* intron and 5' part of *matK*, *trnL-trnF* and *trnT-*  
616 *trnL* intergenic spacer for a total of 2,150 bp). Each circle represents a haplotype, with size  
617 proportional to number of individuals (see table 1).

618 Fig. 3. Maximum clade credibility tree based on sequence variation at four cpDNA loci of  
619 *Arenaria provincialis* and reconstruction of ancestral areas according to BSSVS method.  
620 Nodes are named according to the highest posterior probability of location state. Node  
621 histograms for the three principal nodes show posterior probability distributions of location  
622 state.

623 Fig. 4. Migration patterns in three steps were mapped using the ancestral distributions  
624 reconstructed at each node via the BSSVS method. Arrows indicate direction of migration.

625 Fig. 5. (a) Projection of the first two axes of the OMI analysis following Youssef *et al.* (2011)  
626 realized on 14 variables (described in the text) and 254 stations.

627 (b) Position of the 156 stations in presence of *A. provincialis* inside the two axes.

628 (c) Dotchart figure represents the MCC tree (Fig. 2) and the position of each haplotype within  
629 OMI axes relative to the mean (values above the mean on the right and values below the mean  
630 on the left); individuals are named by geographical group.

631

632 Table 1. Chloroplast DNA haplotypes detected (with the number of individuals carrying a  
633 given haplotype in brackets) for each of the 40 populations of *Arenaria provincialis* sampled  
634 relative to geographical group memberships, codes, GPS coordinates (decimal degrees) and  
635 altitude (geographical location of haplotype in Fig. 1).

636

637 Table 2. Characterization and diversity of the four chloroplast DNA sequences separately and  
638 for the combined matrix (Total) examined for *Arenaria provincialis*.

639

640 Table 3. Summary statistics of AFLP data for the geographical groups of populations of  
641 *Arenaria provincialis*. Number of investigated individuals (N), number of AFLP fragments  
642 ( $\text{Frag}_{\text{tot}}$ ) per population, number of fragments that only occur in one population (private),  
643 percentage of fragments exhibiting intrapopulational polymorphism (%poly), Shannon  
644 diversity index (Shan), and genetic differentiation (Diff) are listed.

645

646

647 Table 1.

Group	Code	Latitude, Longitude	Altitude (m)	Haplotypes
A (South-West)	M16	43.2091 ; 5.36656	39	H20 (3)
	M3A	43.2232 ; 5.36799	365	H20 (3)
	MCC	43.2251 ; 5.36003	361	H20 (3)
	MG4	43.2208 ; 5.35103	143	H26 (3)
	MHM	43.2184 ; 5.38723	395	H20 (3)
	MS	43.2231 ; 5.36841	408	H20 (3)
B (South)	Car8	43.2555 ; 5.52511	303	H18 (1)
	Carp	43.2408 ; 5.48494	325	H21 (2) ; H22 (1)
	Cer	43.2467 ; 5.46236	239	H21 (3)
	Gar10	43.2154 ; 5.48339	276	-
	Gar14	43.2066 ; 5.47836	196	H23 (1)
	Gar2	43.2278 ; 5.48654	398	H21 (2)
	Gar4	43.2254 ; 5.48099	438	-
	Gar8	43.2212 ; 5.47848	282	-
	Gard9	43.2309 ; 5.50711	169	H18 (3)
	LMO	43.2113 ; 5.44408	35	H26 (2) ; H24 (1)
	LSO	43.2122 ; 5.42633	91	H25 (1) ; H27 (2)
	Vau	43.2503 ; 5.43408	200	H27 (2) ; H28 (1)
C (Centre)	Cap	43.2044 ; 5.56306	316	H16 (3)
	Cap10	43.2101 ; 5.57739	194	H19 (2)
	Cap13	43.2384 ; 5.58016	284	H18 (3)
	Cap3	43.1954 ; 5.56473	282	H16 (2) ; H17 (1)
	Car3	43.2637 ; 5.55819	299	H5 (2)
	Cey5	43.2430 ; 5.61752	465	-
	Cey12	43.2229 ; 5.61626	215	H15 (3)
	FB6	43.2374 ; 5.65623	413	H11 (3)
D (East)	Cas3	43.2468 ; 5.73241	409	H6 (3)
	ColB	43.3114 ; 5.68911	1017	H7 (2) ; H8 (1)
	Cu1	43.2799 ; 5.71646	296	H12 (2)
	FB4	43.2450 ; 5.67466	433	H9 (2) ; H10 (1)
	Gem18	43.3193 ; 5.65844	735	H12 (3)
	Gem2	43.3149 ; 5.62552	277	H13 (3)
	Jou	43.3326 ; 5.78356	1089	H14 (3)
	Pcab	43.3239 ; 5.75241	954	H14 (2)
E (North)	Cad11	43.3798 ; 5.51971	361	H1 (2)
	Cad15	43.3531 ; 5.49952	345	-
	G10	43.3433 ; 5.53942	587	H3 (2)
	Roq3	43.3667 ; 5.60136	294	H5 (3)
	Roq5	43.3595 ; 5.60011	388	H5 (2) ; H4 (1)
	Sav2	43.4028 ; 5.52514	582	H1 (1) ; H2 (2)

648

649

650 Table 2.

	fragment length	haplotypes number	polymorphic sites	haplotype diversity (h)	nucleotide diversity ( $\pi$ )
MatK	711	14	19	0.90	0.00787
trnK-matK	830	16	12	0.88	0.00297
trnL-trnF	267	9	12	0.85	0.01099
trnT-trnL	359	13	8	0.77	0.00613
total	2150	28	51	0.94	0.00626

651

652 Table 3.

	A	B	C	D	E
N	11	24	12	16	10
Frag <sub>tot</sub>	133	212	215	186	160
Private	1	9	13	2	4
% poly	50,2	81,67	80,48	72,11	61,35
Shan	0,20±0,02	0,28±0,01	0,33±0,02	0,26±0,01	0,25±0,02
Diff	0,697±0,1	0,703±0,1	0,644±0,08	0,716±0,09	0,701±0,1

653

654

655

656

657

## SUPPLEMENTARY INFORMATION

658 Figure S1. Posterior trace of Bayesian analysis in BEAST (Tracer software); sampled values  
659 are plotted against the step in the MCMC chain.

660 Table S1. Primers and references for amplification of chloroplast loci.

661 Table S2. Genbank accession number of the 28 haplotypes of *Arenaria provincialis* and  
662 Outgroup.

663 Table S3. Cycling condition of the PCR.

664 Table S4. Phylogenetic autocorrelation results from Abouheif tests showing significance on  
665 the two ecological gradients provided by the OMI analysis for all haplotypes and the two  
666 phylogroups independently.

667

668 Table S1.

	Primer names and sequences		references
MatK	390F	GATCTATTCATTCAATATTTTC	Cuenoud <i>et al.</i> , 2002
	1440R	GTGTTTACGAGCYAAAGTTC	
TrnK-MatK	3914F	GGGGTTGCTAACTCAACGG	Schmitz-Linneweber, 2001
	1110R	TATTCTGTTGATACATTCG	
TrnL-trnF	e	GGTTCAAGTCCCTCTATCCC	Taberlet <i>et al.</i> , 1991
	f	ATTTGAACTGGTGACACGAG	
TrnT-L	a	CATTACAAATGCGATGCTCT	Taberlet <i>et al.</i> , 1991
	b	TCTACCGATTTGCCATATC	

669

670

671

672 Table S2.

Haplotypes	Accession numbers of the four chloroplastic regions			
H1	JX475969	JX475952	JX475928	JX475938
H2	JX475969	JX475952	JX475936	JX475938
H3	JX475978	JX475952	JX475928	JX475938
H4	JX475971	JX475956	JX475931	JX475950
H5	JX475971	JX475956	JX475931	JX475942
H6	JX475974	JX475958	JX475933	JX475945
H7	JX475976	JX475960	JX475933	JX475945
H8	JX475976	JX475960	JX475933	JX475946
H9	JX475976	JX475960	JX475933	JX475945
H10	JX475976	JX475960	JX475933	JX475945
H11	JX475977	JX475960	JX475933	JX475945
H12	JX475976	JX475958	JX475933	JX475945
H13	JX475976	JX475964	JX475933	JX475945
H14	JX475980	JX475958	JX475934	JX475945
H15	JX475975	JX475959	JX475933	JX475945
H16	JX475970	JX475953	JX475929	JX475939
H17	JX475970	JX475954	JX475929	JX475939
H18	JX475970	JX475953	JX475929	JX475941
H19	JX475970	JX475955	JX475930	JX475940
H20	JX475982	JX475966	JX475935	JX475949
H21	JX475972	JX475957	JX475932	JX475943
H22	JX475973	JX475957	JX475932	JX475944
H23	JX475979	JX475957	JX475932	JX475943
H24	JX475981	JX475957	JX475932	JX475947
H25	JX475981	JX475957	JX475932	JX475948
H26	JX475981	JX475957	JX475932	JX475943
H27	JX475972	JX475965	JX475932	JX475943
H28	JX475972	JX475967	JX475932	JX475943
<i>Arenaria cinerea</i>	JX475968	JX475951	JX475927	JX475937
<i>Arenaria ciliata</i>	KC524727	KC524729	KC524730	KC524732
<i>Arenaria hispida</i>	GU732828	KC524728	KC524731	KC524733

673

674

675

676 Table S3.

Chloroplast sequence	matK	trnT-trnL	trnK-matK	trnL-trnF
	94°C 2 min	94°C 2 min	94°C 30s	94°C 2 min
	94°C 45s	94°C 45s	94°C 30s	94°C 45s
	48°C 45s	52°C 45s	48°C 30s	52°C 45s
	72°C 2 min	72°C 2 min	72°C 2 min	72°C 2 min
	go to 2 rep29	go to 2 rep31	go to 2 rep29	go to 2 rep31
	72°C 5 min	72°C 5 min	72°C 5 min	72°C 5 min

677

678 Table S4

		Obs	Std.Obs	Alter	Pvalue
All haplotype	Axis1	0.745	10.615	greater	0.001
	Axis2	0.571	7.982	greater	0.001
phylogroup 1	Axis1	0.525	5.539	greater	0.001
	Axis2	0.587	6.284	greater	0.001
phylogroup 2	Axis1	0.659	6.434	greater	0.001
	Axis2	0.575	5.576	greater	0.001

679

680

681

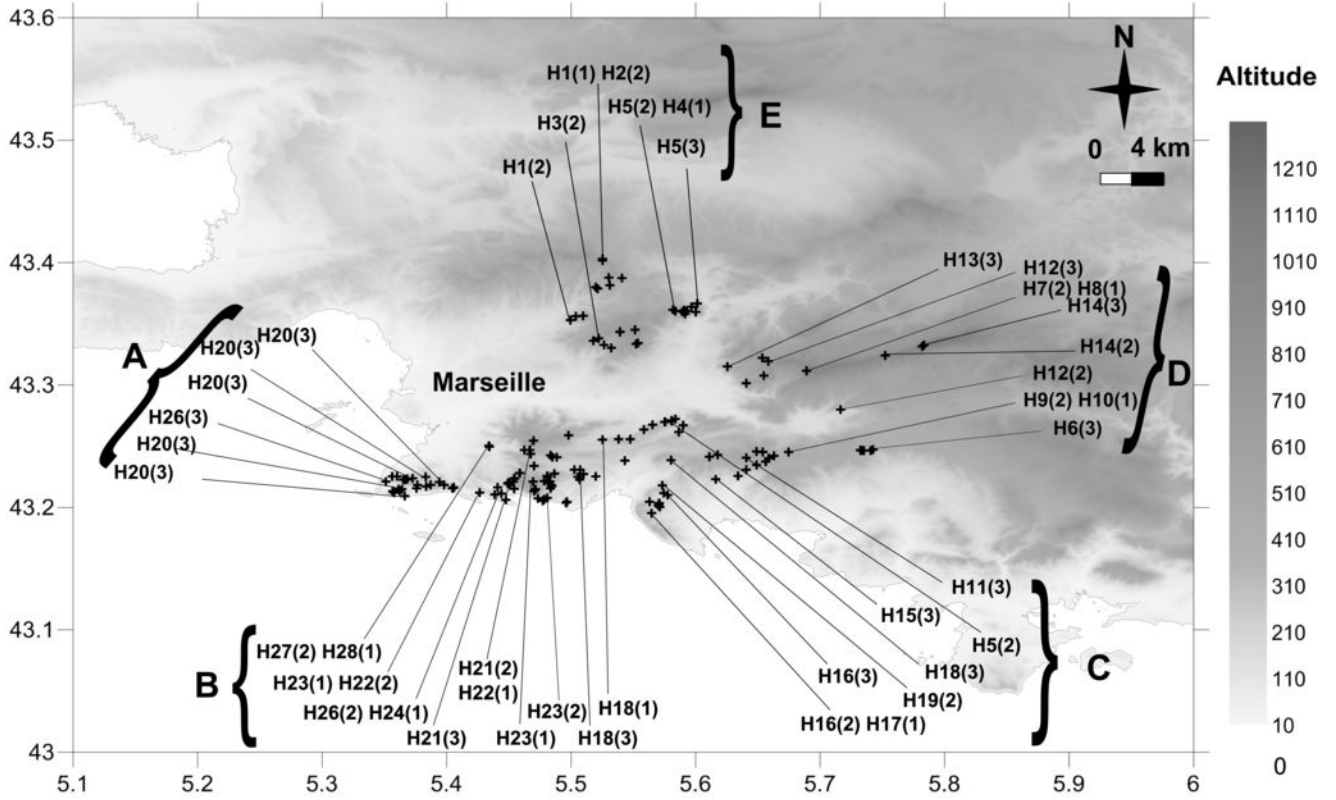
SUPPLEMENTARY LITERATURE

- 682 **Cuenoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW. 2002.**  
683 Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL,  
684 atpB, and matK DNA sequences. *American Journal of Botany* **89**: 132-144.
- 685 **Taberlet P, Gielly L, Patou G, Bouvet J. 1991.** Universal primers for amplification of three  
686 non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**:1105–1109.
- 687 **Schmitz-Linneweber C, Maier RM, Alcaraz JP, Cottet A, Hermann RG, Marche R. 2001.**  
688 The plastid chromosome of spinach (*Spinacia oleracea*): complete nucleotide sequence and  
689 gene organization. *Plant Molecular Biology* **45**:307–315

690

691



**A.****B.**