Plant budding speciation predominant by ecological and geographical differentiation: an 'evolutionary snapshot' in *Iberodes* (Boraginaceae).

Ana Otero[1], Pablo Vargas[1], Virginia Valcárcel[2,3], Mario Fernández-Mazuecos[1], Pedro Jiménez-Mejías[2,3] and Andrew L. Hipp[4,5]

[1] Departamento de Biodiversidad, Real Jardín Botánico, CSIC. Pza. de Murillo, 2, 28014 Madrid, Spain

[2] Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM), Universidad Autónoma de Madrid, 28049 Madrid, Spain

[3] Departamento de Biología (Botánica), Universidad Autónoma de Madrid, C/Darwin, 2, 28049 Madrid, Spain

[4] The Morton Arboretum, 4100 Illinois Route 53, Lisle, IL 60532, USA

[5] The Field Museum, 1400 S Lake Shore Drive, Chicago, IL 60605, USA

ABSTRACT

Traditional classification of speciation modes has been focused on physical barriers to gene flow. Allopatry—geographic separation of populations so that gene flow between them is more or less severed—has as a consequence been viewed as the most common mechanism of speciation (allopatric speciation). By contrast, parapatry and sympatry—no physical barrier is found separating populations (adjacent or overlapped) so that speciation takes place in the face of ongoing gene flow— is more difficult to demonstrate. *Iberodes* (Boraginaceae) is a primarily Iberian genus that comprises a relatively small number of recently derived species (five) with contrasting diaspore dispersal traits, habitats and

distribution patterns. In particular, species distributions of *Iberodes* offer an ideal system to study drivers of lineage divergence and differentiation in the course of evolution. To reconstruct the evolutionary history of the genus, we addressed an integrative study combining a suitable set of tools: (i) phylogenetics based on restriction-site associated DNA sequencing (RAD-seq), (ii) morphometrics, and (iii) climatic niche modelling. Our robust phylogenetic reconstruction of *Iberodes* reveals a predominantly budding speciation pattern. Climatic niche analyses, together with the morphometric data and species distributions, suggests that ecological and geographical differentiation interact in concert, in such a way that both allopatric and parapatric have been operating in *Iberodes*.

Keywords: Restriction-site associated DNA sequencing (RAD-seq), molecular dating, species distribution modelling, niche overlap, paraphyly, allopatry, parapatry.

INTRODUCTION

The classification of modes of speciation has been one of the more contentious topics in evolutionary biology for several decades (e.g. Templeton, 1981; Rieseberg and Brouillet, 1994; Butlin *et al.*, 2008; Horandl and Stuessy, 2010). Reproductive barriers are the most frequently used criteria for classifying three main speciation modes (Futuyma, 2009). Speciation of adjacent populations in absence of clear physical or reproductive barriers is frequently explained under the process of parapatric speciation, which entails the evolution of reproductive isolation in the face of limited but ongoing gene flow (Coyne and Orr, 2004, p. 112). Parapatric speciation can be considered intermediate between allopatric and sympatric

50 speciation. In allopatry, gene flow is prevented by geographic barriers between
51 populations. In sympatry, partial or complete reproductive isolation arises
52 between subsets of a single population without spatial segregation (keeping
53 overlapping ranges). Sympatric speciation often entails significant levels of gene
54 flow at the early stages of differentiation (Gavrilets, 2003; Futuyma, 2009).
55 Whereas allopatry is widely viewed to be the most common mode of speciation
56 (Coyne and Orr, 2004; Vargas *et al.*, 2018), the relative importance of parapatry
57 and sympatry has been the subject of considerable debate (Fitzpatrick *et al.*, 2008,
58 2009; Arnold, 2015) in large part due to the difficulty of (1) resolving fine-scale
59 phylogenetic relationships among very closely related populations that may still be
60 exchanging genes at a low level, and (2) quantifying the amount of gene flow
61 between populations during population divergence (Barluenga *et al.*, 2006;
62 Fontdevilla, 2014). In the last decade, new genomic, phylogenetic, and ecological
63 tools have made non-allopatric modes of speciation increasingly amenable to
64 study, allowing us to explore with statistical rigor a wider range of speciation
65 scenarios (e.g. Savolainen *et al.*, 2006; Chozas *et al.*, 2017; Zheng *et al.*, 2017). In
66 particular, there are some critical points of evidence that would support a
67 parapatric scenario: (1) phylogenetic relationships usually reflect a progenitor-
68 derivative (budding) pattern in which the most widely distributed species is the
69 living progenitor of the most restricted one; (2) the absence of any physical barrier
70 to gene-flow in both present and past times, i.e. distribution ranges are adjacent;
71 (3) clear interspecific genetic clustering that putatively discards processes of
72 secondary contact when testing for ancient and contemporary gene flow; (4) the
73 need of alternative factors (differential climatic niches, reproductive exclusion,

74 ploidy variation, among others) of isolation other than physical barriers are
75 claimed to be in play.

76 Ecological differentiation is often attributed to non-allopatric speciation (Sobel
77 *et al.* 2010; Nosil 2012; Stankowsky *et al.* 2015). Indeed, ecological speciation is
78 claimed to be more important in plants than in animals because a sessile habit
79 makes plants more sensitive to fine-scale environmental heterogeneity (Anacker
80 and Strauss, 2014). Both the influence of niche conservatism (i.e. retention of
81 ancestral ecological characteristics of species over time, Peterson *et al.* 1999) and
82 niche evolution (i.e. adaptation of lineages to changes in the environment,
83 Donoghue and Edwards 2014) have been argued to play a role in speciation and
84 lineage diversification (Wiens and Graham, 2005; Cavender-Bares, 2019).
85 Nevertheless, the contribution of ecological factors as primary drivers of
86 speciation remains more elusive (Mayr, 1954; Pyron and Burbrink, 2010; Yin et al.,
87 2016).

88 The advent of inexpensive phylogenomic approaches in the last decade have
89 made the genetic side of testing recent speciation scenarios tractable (Mc Cormack
90 *et al.*, 2013; Mc Vay *et al.*, 2017a). High Throughput Sequencing (HTS) approaches
91 help to analyze loci sampled from the entire genome in reconstructing species
92 trees (Fernández-Mazuecos *et al.*, 2018) and speciation in the face of gene flow
93 (Leroy *et al.*, 2017; Folk *et al.*, 2018; Crowl *et al.*, 2019). In particular, restriction-
94 site associated DNA (RAD-seq) is considerably contributing to our understanding
95 of recent evolutionary processes, being especially helpful for non-model organisms
96 since a reference genome is not needed for accurate phylogenetic inference (e.g.
97 Fitz-Gibbson *et al.*, 2017). RAD-seq has as a consequence been useful in inferring
98 complex speciation histories: repeated cycles of island connectivity and isolation

99 (the pump hypothesis; Papadopoulou and Knowles, 2015), incipient sympatric
100 speciation (Kautt et al., 2016), allopatric speciation despite historical gene flow
101 (Maguilla *et al.*, 2017), ancient introgression among now-extinct species (Mc Vay *et*
102 *al.* 2017b).

103 The Mediterranean subendemic genus *Iberodes* M.Serrano, R.Carbajal &
104 S.Ortiz (Boraginaceae, Cynoglossoideae; Serrano *et al.*, 2016) provides an excellent
105 system for studying recent speciation, morphological and ecological differentiation
106 because it is clearly demonstrated to be a recently-derived genus (Chacón *et al.*
107 2017, Otero *et al.* 2019a). *Iberodes* comprises five species and one subspecies that
108 differ in geographic ranges (including allopatric and parapatric species) and
109 habitat (ranging from forest understories to scrublands and coastal dunes).
110 Moreover, all species are endangered and narrowly endemic, except for the
111 widely-distributed *I. linifolia* of the Iberian Peninsula and southeastern France
112 (Talavera *et al.*, 2012). Sanger sequencing has failed to resolve phylogenetic
113 relationships among *Iberodes* species because of scarcely information of traditional
114 markers (Otero *et al.*, 2014; Holstein *et al.*, 2016ab, Otero *et al.*, 2019a). The
115 question remains as to what is the relative importance of extrinsic geographical
116 and ecological barriers in the complex speciation pattern observed for *Iberodes*.
117 Given the narrow endemicity in most *Iberodes* species, we hypothesize that
118 geographic barriers rather than ecological factors have been predominant in the
119 speciation history of *Iberodes*. To test this hypothesis, we aimed to: (1) reconstruct
120 phylogenetic relationships of the five species; (2) infer speciation patterns and
121 lineage diversification during the last geological epochs comparing diverse
122 morphologies and ecologies; and (3) evaluate the relative contribution of
123 geographical and ecological processes in speciation.

124 ## MATERIAL AND METHODS

125 ### Taxon sampling

126 Based on species distributions, between one and four individuals from one to nine
127 populations were sampled for the five species of *Iberodes* including ssp. *gallaecica*
128 and ssp. *littoralis* of *I. littoralis*) (Fig. 1, Table S1, in Supporting Information).
129 Samples of *I. littoralis*, *I. kuzinskyana* and *I. brassicifolia* were collected in the field
130 during 2015 and stored in silica gel until extraction. Since these species are
131 considered threatened according to IUCN criteria (Lopes and Carvalho, 1990;
132 Serrano and Carbajal, 2003; ICNB, 2007; Moreno, 2008; IUCN, 2019; INPN, 2019),
133 permits were obtained prior to collecting. Mature seeds of *I. commutata* were
134 collected in 2015 and germinated in a glasshouse to obtain green leaves for DNA
135 extraction. Both field and herbarium materials were used to represent the large
136 distribution of *I. linifolia* (Table S1). Three species (17 individuals) of the tribe
137 Omphalodeae were included as the outgroup (Otero *et al.*, 2019b): *Omphalodes*
138 *nitida*, *Myosotidium hortensia*, and *Gyrocaryum oppositifolium* (Table S1).

139 ### DNA extraction and RAD library preparation

140 DNA was extracted from leaf tissue using a modified CTAB protocol from Doyle &
141 Doyle (1987) and Shepherd & McLay (2011), including a precipitation step in
142 isopropanol and ammonium acetate 7.5M at -20ºC overnight to improve the yield.
143 DNA extractions were visualized and quantified with NanoDrop 2000/2000c
144 (Thermo-Fisher, Waltham, Massachusetts) and Qubit Fluorometric Quantification
145 (Thermo-Fisher, Waltham, Massachusetts) in the *Instituto de Investigaciones*
146 *Biomédicas* (IIBm, CSIC-UAM). Final DNA sample concentrations were
147 standardized to 10 ng/ul. Preparation of single-end RAD-seq libraries using
148 restriction enzyme PstI from genomic DNA was conducted at Floragenex Inc.

149 (Eugene, Oregon) following Baird *et al.* (2008) and sequenced as single-end, 100-
150 bp reactions on an Illumina HiSeq 2500 at the University of Orgeon Genomics &
151 Cells Characterization Core Facility (library C606). Processed data were returned
152 in the Illumina 1.3 variant of the FASTQ format, with Phred quality scores for all
153 bases.

154 **Data clustering**

155 Sequences from Illumina were analyzed through *PyRAD* v. 3.0.6, a pipeline with
156 seven main steps that filter and cluster putatively orthologous loci from RAD
157 sequencing reads (Eaton, 2014). This pipeline is able to cluster highly divergent
158 sequences, taking into account indel variation and nucleotide (SNP) variation. It is
159 thus well-suited to the inter- and intraspecific scales addressed in our study. As no
160 evidence of polyploidy was found from flow cytometry, we did not assess
161 sensitivity of paralog detection to sequencing depth or number of heterozygotic
162 positions allowed per loci or shared among individuals. Base calls with a Phred
163 Quality score <20 were replaced by the ambiguous base code (N). Three similarity
164 percentages were set for the clustering (c) threshold of reads: 85%, 90% and 95%.
165 Four levels of minimum coverage (m) of a sample in a final locus (4, 12, 20, 28)
166 were tested for each of the three percentages of clustering similarity, resulting in a
167 fully factorial set of 12 matrices from the combination of filtering parameters, so
168 we could assess sensitivity of our phylogenetic inferences to clustering
169 assumptions. Consensus base calls were made for clusters with a minimum depth
170 of coverage >5. From each parameter combination we obtained three kinds of
171 output: (1) a matrix with all loci retrieved for each individual (concatenated
172 matrix), (2) a matrix with single nucleotide polymorphisms (SNP matrix), and (3) a
173 matrix with one SNP per locus (unlinked SNP matrix). The loci matrix was

174 explored with the RADami (Hipp, 2014) and vegan (Oksanen *et al.*, 2008) packages
175 of R version 3.3.2 (R Core Team, 2013).

176 **Phylogenetic analysis**

177 Maximum likelihood analyses of concatenated matrices obtained in *PyRAD* were
178 performed in RAxML v8.2.10 (Stamatakis, 2014) through the Cipres Portal (Miller
179 *et al.*, 2010). We used rapid bootstrapping, GTRGAMMA model, and automatic-stop
180 bootstrapping with the majority rule criterion. We conducted a maximum
181 likelihood analysis for each of the 12 *PyRAD* concatenated matrices. Finally, we
182 compared the topologies based on average bootstrap support (BS). Between the
183 two topologies with average BS>98%, we chose the one maximizing the number of
184 loci (HBS tree) for the remaining phylogenetic-based analyses. Moreover, we used
185 the unlinked SNP matrix from the HBS tree (HBS matrix) to perform a coalescent-
186 based phylogeny using SVDquartets (Chifman and Kubatko, 2014) in PAUP
187 (Swofford, 2001) through the Cipres portal (Miller *et al.*, 2010). For SVDquartets
188 we grouped individuals based on main clades obtained from concatenated analysis
189 in RAxML and evaluated all possible quartets with 100 bootstrap replicates.

190 **Estimates of divergence times**

191 We used penalized likelihood as implemented in TreePL (Smith and O'Meara,
192 2012) to estimate a time-calibrated tree for all bootstrap trees (with branch
193 lengths) obtained in RAxML from the HBS tree. TreePL is suitable for divergence
194 time estimation when dealing with large amounts of data, as those yielded by high
195 throughput sequencing such as RAD-sequencing (Zheng and Wiens, 2015). Two
196 calibration points were used: (1) tribe Omphalodeae (root) (minimum age
197 (minAge) = 9.062; maximum age (maxAge) = 24.4097) and (2) clade *Iberodes*
198 (ingroup) (minAge = 0.6591; maxAge = 3.7132) based on the averaged ages and

199 standard deviation inferred in Otero *et al.* (2019a). Average node ages over all
200 posterior trees from the latter study were calculated using *treeStat* from the
201 BEAST v 1.8.2 package (Drummond *et al.*, 2012). We first conducted an analysis
202 under the "prime" option to select the optimal set of parameter values and we set
203 random subsample and replicate cross-validation (RSRCV) to identify the best
204 value for the smoothing parameter lambda. For this first analysis we used the
205 thorough analysis option and set 200,000 iterations for penalized likelihood and
206 5,000 iterations for cross validation. Consequently, we repeated the same
207 thorough analysis and iterations but setting the smoothing value of lambda to 10,
208 gradient based (opt) and auto-differentiation based optimizers (optad) to five, and
209 auto-differentiation cross-validation-based optimizers to two (optcvad). Although
210 TreePL does not draw a prior distribution for the phylogenetic tree, we took into
211 account the bootstrapped variance in topology and branch length by running the
212 same TreePL analysis for each bootstrap tree and then a maximum clade
213 credibility tree using the mean heights was reconstructed in *TreeAnnotator* from
214 the BEAST v. 1.8.2 package (Drummond *et al.*, 2012).

215 Phylogenetic inferences showed two cases of paraphyly within the lineage
216 of *I. linifolia* that involved three other taxa (*I. kuzinskyana*, *I. littoralis* ssp. *littoralis*
217 and ssp. *gallaecica*; see results). To evaluate these cases of paraphyly, a set of
218 analyses were done to analyze the degree of genetic, morphological and ecological
219 differentiation among the four taxa involved. Therefore, the following analyses
220 were performed on a subset of data only including these four taxa.

221 **Genetic structure**

222 The HBS matrix was used to perform a Discriminant Analysis of Principal
223 Components (DAPC) to identify and describe clusters of genetically related

224 individuals (Jombart *et al.*, 2010). The format of the HBS matrix was converted
225 from *vcf* to *genlight* with vcfR (Knaus and Grünwald, 2017). DAPC analysis was
226 performed in adegenet (Jombart, 2008). We fixed the number of clusters to four,
227 based on the natural groups of four taxonomic species and subspecies (the
228 Linifolia clade, see results). We used the function 'optim.a.score' to estimate the
229 optimal number of principal components for discriminant analysis. We
230 represented genetic differentiation using scatterplots of the principal components
231 and a barplot representing assignment probabilities of individuals to groups. In
232 addition, analysis of molecular variance (AMOVA) was performed in Arlequin v.
233 3.5.2.1 (Excoffier and Lischer, 2010). AMOVA was run to evaluate the level of
234 genetic differentiation of the paraphyletic *I. linifolia* as a whole with respect to the
235 other two species (*I. kunzinskyana* and *I. littoralis*) of the Linifolia clade. Therefore,
236 we considered only two groups, *I. linifolia* on one hand, and the two coastal species
237 together (*I. kusinskyana* and *I. littoralis*) on the other. Finally, to evaluate a
238 potential role of historical introgression in determining paraphyly, we conducted
239 D-statistic tests in PyRAD (see Methods S1 for details).

240 **Morphometric evaluation**

241 We performed a morphometric exploration using principal component analysis
242 (PCA). We consulted the taxonomic literature to identify diagnostic characters in
243 *Iberodes* (Brand, 1921; Ferguson, 1972; Fernández and Talavera, 2012; Neto *et al.*,
244 2015). We measured 12 diagnostic macromorphological vegetative and
245 reproductive characters for 111 individuals from 58 vouchers (two different
246 individuals per voucher except for 5 vouchers with only one individual). We
247 averaged measures from each two individuals of the same voucher. As a result, we
248 obtained a final matrix of 58 accessions from 47 locations (74 individuals from 35

249 locations of *I. linifolia*, 14 from five locations of *I. kuzinskyana*, and 23 from seven
250 locations of *I. littoralis* (including the two subspecies) (Fig. 1, Table S2). Ten
251 morphological characters were quantitatively continuous: four vegetative
252 characters (length of stem, leaves and pedicels, and width of leaves) and six
253 reproductive characters (length of fruiting calix, width and length of nutlet, length
254 of nutlet margin, length of the nutlet margin teeth, and hair length). In addition,
255 one binomial character (presence or absence of bracts), and one quantitative,
256 discrete character (hair number per mm^2 on the abaxial surface of the nutlet) were
257 also considered.

258 PCA was run using the function 'prcomp' from Stats R package (R Core Team,
259 2013). Results were visualized with the function 'ggbiplot' in the ggbiplot R
260 package (Vu, 2011).

261 **Climatic niche differentiation and distribution modelling**

262 Occurrences of the four taxa were collected from GBIF (https://www.gbif.org/).
263 We filtered the resulting dataset by removing points suspected to be incorrect,
264 such as those placed clearly outside of the known species distribution range (based
265 on taxonomic literature and databases; Fernández and Talavera, 2012; Muséum
266 national d'Histoire Naturelle, 2003-2019). Additional geographic coordinates were
267 obtained from herbarium material and our own fieldwork data. We also estimated
268 the geographic coordinates of non-georeferenced herbarium specimens, whenever
269 the locality provided could be pinpointed with an accuracy of ± 5 km (Table S3, Fig.
270 1). To reduce sampling bias, we filtered the dataset by randomly removing points
271 that were 0.2° of each other for *I. linifolia* and 0.1° for *I. littoralis* ssp. *littoralis* and
272 ssp. *gallaecica*. Filtering was not applied for *I. kuzinskyana* because of the low
273 number of occurrences. The final dataset after filtering included 115 occurrences

274 of *I. linifolia*, 11 of *I. kuzinskyana*, 16 of *I. littoralis* ssp. *littoralis* and 11 of ssp.
275 *gallaecica*.

276 We obtained 19 bioclimatic variables (resolution of 30") from WorldClim 1.4
277 (http://www.worldclim.org/; Hijmans et al, 2006) for the study region (34° to 50°
278 N; 11° W to 7° E). To avoid collinearity among bioclimatic variables, first we
279 excluded variables displaying a high correlation coefficient with other variables
280 (|r| > 0.7) in the study area (Dormann *et al.*, 2013). Then, we calculated the
281 variance inflation factor (VIF) for each remaining variable using the HH package in
282 R (Heiberger, 2017) and selected variables with VIF <5 following Benitez-Benitez
283 *et al.* (2018). As a result, the following six variables were selected: bio1, bio3, bio4,
284 bio8, bio12 and bio14. We assessed climatic niche overlap among the four taxa by
285 pairwise comparisons. The E-space (i.e. environmental space resulting from the six
286 climatic variables) of each taxon was analyzed in a Principal Component Analysis
287 (PCA) as implemented in the R package ecospat (Di Cola *et al.*, 2017). To visualize
288 niche differences along each of the first two principal components, the result was
289 plotted using the niceOverPlot function (Fernández-López and Villa-Machío, 2017),
290 which relies on the ggplot2 package (Wickham, 2016). We calculated values of
291 Schoener's index (D) as a measure of niche overlap (Schoener, 1968; Warren *et al.*,
292 2008). Then, we conducted tests of niche equivalency and niche similarity using
293 ecospat (Di Cola *et al.*, 2017). The niche equivalency test evaluates whether the
294 observed niche overlap is significantly different from a null simulated by randomly
295 reallocating the occurrences of both entities between their ranges (Broennimann
296 *et al.*, 2012; Warren *et al.*, 2008). The niche similarity test checks whether the
297 overlap between two niches is significantly different from the overlap obtained if
298 random shifts within each environmental space are allowed (Schoener, 1968;

299 Warren *et al.*, 2008). In both cases, we tested for niche divergence by using the
300 alternative="lower" option. All tests were based on 100 iterations. In similarity
301 tests, both ranges were randomly shifted (rand.type=1).

302 In addition, we modeled the potential range of the four taxa of the Linifolia
303 clade under present conditions, which was then projected to last interglacial (LIG,
304 c. 120–140 kya) and last glacial maximum (LGM, c. 21 kya) conditions. To this end,
305 we used the same six WorldClim bioclimatic variables employed for niche
306 differentiation analyses, including present layers, LIG layers from Otto-Bliesner *et*
307 *al.* (2006) and LGM layers based on the CMIP5 project (three global climate models
308 (GCMs): CCSM4, MIROC-ESM and MPI-ESM-P). Resolution was 30″ for current and
309 LIG, and 2.5′ for LGM layers. We performed species distribution modeling (SDM)
310 using the maximum entropy algorithm, as implemented in Maxent v3.4 (Phillips *et*
311 *al.*, 2006). We used the same occurrences for each taxon employed for niche
312 differentiation analyses (see above). 80% of occurrences for each taxon were used
313 for model training and 20% for model evaluation. For each taxon, ten subsample
314 replicates were run by randomly partitioning the data into a training set and
315 evaluation set, and a mean model was calculated. For the LGM, an average of the
316 three GCMs was calculated. Logistic outputs were converted into
317 presence/absence using the maximum training sensitivity plus specificity logistic
318 threshold.

319 **Ploidy level estimation**

320 Flow cytometry was used to estimate genome size, explore the possible role of
321 polyploidization in speciation, and assess the risk of clustering paralogs in RAD-
322 seq *de novo* clustering. Earlier cytogenetic studies for three *Iberodes* species
323 showed them as diploids (2n=28 for *I. linifolia*, *I. commutata* and *I. kuzinskyana*

324 (Franco, 1984; Saly, 1997; Talavera *et al.*, 2012), and aneuploidy was detected for *I.*
325 *littoralis* with 2n=24 (Fernández-Casas, 1975). Nevertheless, some of these
326 references do not cite the information sources. Individuals of the five *Iberodes*
327 species were cultivated at the glasshouse and five individuals per species were
328 sampled. In addition, samples of the diploid *Solanum lycopersicum* of known
329 genome size (2n=24, Rick, 1979; 2C=1.88-2.07 pg, Grandillo *et al.*, 2011) were also
330 cultivated to compare standardized results with our study species and estimate
331 genome sizes (Doležel and Bartoš, 2005). We performed a two-step nuclear DNA
332 Content Analysis using Partec Buffer (de Laat *et al.*, 1987) and DAPI fluorochrome
333 through a Cell Lab Quanta SC flow cytometer (Beckman Coulter, Fullerton, CA,
334 USA) equipped with a mercury lamp following the protocol of Doležel et al. (2007).
335 Samples were visualized through the Cell Lab Quanta SC Software package. FL1
336 detector (530 nm / 28 nm; DAPI emission maximum=461) and FL1-Area was used
337 as directly correlated to DNA content. A minimum number of 1000 nuclei for G1
338 peak were analyzed. Only histograms with a coefficient of variation (CV) lower
339 than 10% for the G1 peak were accepted.
340

341 **RESULTS**

342 **Phylogenetic analysis**

343 After filtering and processing all samples under the different parameter
344 combinations in *PyRAD*, 12 matrices of 54 taxa with different numbers of loci and
345 SNPs were analysed (see Table S4). The same topology was obtained in RAxML
346 from each of the 12 matrices, differing only in bootstrap support (BS) values for
347 the nodes (Fig. S1). The highest BS supports and highest number of loci were
348 obtained for the matrix c95m20 (95% similarity within clusters and 20 minimum

349 coverage of samples for a final locus). Consequently, this matrix was preferred for
350 the remaining phylogenetic analyses (HBS matrix). *Iberodes brassicifolia* is the
351 earliest-diverging lineage, sister to a clade that contains *I. commutata* sister to
352 another subclade with the remaining species (Fig. 2). Paraphyly was inferred for *I.*
353 *linifolia* since a clade including *I. kuzinskyana* and *I. littoralis* is nested within.
354 Likewise, paraphyly was inferred for *I. littoralis* ssp. *gallaecica* since ssp. *littoralis*
355 is nested within populations of subsp. *gallaecica* (Fig. 2).

356 Taxon groups for coalescent-based phylogeny (SVDquartets) were based on
357 the eight main lineages of RAxML (*Iberodes brassicifolia*, *I. commutata*, populations
358 of *I. linifolia* from Seville, core clade of *I. linifolia*, populations of *I. linifolia* from SW
359 Spain, populations of *I. linifolia* from central Portugal, *I. kusinskyana*, and *I.*
360 *littoralis*, Fig. S2). The species tree topology was mostly congruent with the one
361 obtained from the concatenated data-based phylogeny. The only difference was the
362 placement of *I. linifolia* populations from Seville and SW Spain, which were nested
363 within the core of *I. linifolia* (Fig. S2). In any case, paraphyly of *I. linifolia* is still
364 inferred because of the nested position of *I. kuzinskyana* and *I. littoralis* sister to
365 populations of *I. linifolia* from Portugal (Fig. S2).

366 **Estimate of divergence times**

367 Node ages for main lineages of *Iberodes* are represented in Fig. 2. The stem age
368 estimated for *Iberodes* was 17.56 myr (17.26-17.86 Bootstrapped Variance, BV)
369 with a crown age of 5.32 myr (5.32-5.32 BV). Crown ages for *I. brassicifolia* and *I.*
370 *commutata* were 1.29 (1.21-1.33 BV) and 2.22 myr (2.17-2.28 BV), respectively.
371 The crown age of the Linifolia clade (*I. linifolia* s.l., *I. kuzinskyana* and *I. littoralis*)
372 was inferred in 4.33 myr (4.25-4.59 BV). Divergence time between Portuguese
373 populations of *I. linifolia* and ancestor of the coastal species (*I. kuzinskyana* and *I.*

374 *littoralis*) was 3.67 myr (3.60-3.85 HPD). The origin of *I. littoralis* ssp. *littoralis*,
375 possibly from an *I. littoralis* ssp. *gallaecica* ancestor is inferred 10,000 years ago
376 (Fig. 2).

377 **Genetic structure**

378 We set four as the *a priori* number of clusters (k=4) based on the four taxa
379 circumscribed within the Linifolia clade. DAPC scatterplot revealed two well-
380 differentiated clusters: *I. linifolia* (LIN) on the one side, and the coastal *I.*
381 *kuzinskiana* (KUZ) and *I. littoralis* (LIT_LIT, LIT_GAL) on the other (Fig. 3c,). Within
382 the coastal species cluster, *I. kuzinskyana* occupied a position closer to *I. linifolia*,
383 while the two subspecies of *I. littoralis* were intermingled (Fig. 3c). Compoplot
384 assigned all individuals of *I. linifolia* to a separate cluster, without admixture (Fig.
385 3c). The other three clusters were shared among coastal species with different
386 membership probability (Fig. 3c). One of the two individuals of *I. kuzinskyana* was
387 uniquely assigned to one of the clusters, whereas the other individual showed a
388 mixed ascription between the three genetic groups (Fig. 3c). Most individuals
389 (nine) of both subspecies of *I. littoralis* were placed in the same two clusters, while
390 two of them showed also admixture with the same cluster shared by the two
391 samples of *I. kuzinskyana* (Fig. 3c). Given the interdigitated pattern of the two
392 subspecies of *I. littoralis*, we repeated the analysis by considering the two
393 subspecies as part of a single group (K=3) to evaluate the genetic clustering of *I.*
394 *littoralis* as a genetic group. Compoplot for K=3 assigned most individuals (nine) of
395 *I. littoralis* to the same cluster, while two samples had mixture assignment
396 probability to the cluster of *I. kuzinskyana* (Fig. S3). Likewise, AMOVA of the two
397 groups (*I. linifolia* vs coastal species) showed significant among-group variance
398 (58.88%) higher than within-population variance (38.09%) or variance within

399 populations among groups (3.03%), with a fixation index of F_{ST}= 0.619 (Table 1).
400 D-statistic tests did not conclusively support a role of historical introgression in
401 determining paraphyletic pattern since ancient introgression was inferred
402 indiscriminately among lineages of *I. linifolia* and both coastland species (see
403 Methods S1).
404
405 **Morphometric analysis**
406 Principal Component Analysis (PCA) revealed morphological differentiation
407 between the three taxa. *I. linifolia* held the greatest morphological variation
408 ranging from the two extremes of both PC1 and PC2 compared to the more
409 restricted variation of coastal species that is concentrated in higher scores of PC1
410 and lower of PC2. Three species of the Linifolia clade were almost completely
411 differentiated, which supports the previous taxonomic observations. *I. linifolia* and
412 *I. littoralis* occupied the two extremes of variation ranging from higher length of
413 vegetative characters, higher nutlet sizes and nutlet teeth for *I. linifolia* to higher
414 length and density of nutlet trichomes and presence of bracts for *I. littoralis* (Fig.
415 3a). *Iberodes kuzinskyana* occupied an intermediate position between *I. littoralis*
416 and *I. linifolia*, overlapping very slightly with the latter. The two subspecies of *I.*
417 *littoralis* were intermingled at the highest scores of PC1 and the lowest of PC2 (Fig.
418 3a).
419 **Climatic niche differentiation and distribution modelling**
420 The two main principal components of the 6 climatic variables explained 73.59%
421 (PC1=48.86% and PC2=24.73%) of the climatic variability in the study region. The
422 variables that contributed most to PC1 were: bio1 (annual mean temperature),
423 bio8 (mean temperature of wettest quarter), bio12 (annual precipitation), and

424 bio14 (precipitation of the driest month). Bio3 (isothermality) and bio4
425 (temperature seasonality) were the two variables most related to PC2. The
426 visualization of the e-space showed little overlap among taxa (Fig. 3b). The
427 geographically widespread *Iberodes linifolia* showed the widest niche, whereas the
428 more narrowly distributed coastal species are more restricted in climatic niche
429 space (Fig. 3b). Among coastal species, *I. kuzinskyana* differs from *I. littoralis* in PC1
430 (less precipitation and higher temperatures for *I. kuzinskyana*) and overlaps with *I.*
431 *littoralis* ssp. *gallaecica* but not with *I. littoralis* ssp. *littoralis* in PC2 (higher
432 isothermality in *I. kuzinskyana* and *I. littoralis* ssp. *gallaecica*). The two subspecies
433 of *I. littoralis* have well-differentiated niches particularly along PC2. *I. kuzinskyana*
434 occupies an intermediate position between *I. linifolia* and *I. littoralis* subsp.
435 *gallaecica*. Thus, *I. kuzinskyana* marginally overlapped with *I. linifolia* for those
436 parts of the niche of *I. linifolia* with higher isothermality, lower temperature and
437 higher precipitation. Indeed, minimum niche overlap was inferred between both
438 subspecies of *I. littoralis* and *I. linifolia* due to lower temperature and higher
439 precipitation for coastal taxa. Consequently, although tests of similarity indicated
440 that observed values of niche overlap are not significantly lower than random
441 overlap, all pairwise equivalency tests significantly rejected the equivalency of
442 niches and this is supported by the low values of D obtained (Table 2).

443 The potential ranges for the species of the Linifolia clade in the present are
444 broadly congruent with known geographic distributions, with overlapping
445 distributions of *I. linifolia* and *I. kuzinskyana* (Fig. 4). In contrast, a contraction of
446 distribution ranges was inferred for the LIG, including the disappearance of the
447 potential range of *I. littoralis* ssp. *gallaecica* (Fig. 4), but maintaining the overlap
448 and absence of physical barriers between *I. linifolia* and *I. kuzinskyana*. Projections

449 to the LGM showed a range similar to the present one for *I. linifolia* and expanded
450 ranges for both *I. kuzinskyana* and *I. littoralis* subsp. *gallaecica*. The range of *I.*
451 *littoralis* subsp. *gallaecica* is inferred to have occupied most of the current range of
452 both subspecies in the LGM, before subsp. *littoralis* differentiated from subsp.
453 *gallaecica*.

454 **Ploidy level estimation**

455 Mean peaks of fluorescence, CV values, and genome sizes obtained for each species
456 are summarized in Fig. S4 and Table S5. Fluorescence peaks were similar among all
457 *Iberodes* species, ranging from 45.2 to 62.6 nm for G1. No evidence of duplicated
458 DNA content was observed, suggesting no ploidy differences among taxa or
459 populations. Likewise, similar genome sizes were obtained for the five species
460 ranging from 1.788 to 1.980 pg. *I. kuzinskyana* and *I. littoralis* ssp. *littoralis* were
461 the *Iberodes* species with the largest and the smallest genome sizes, respectively.
462 The five species are similar in genome size to the *S. lycopersicum* standard (1.88-
463 2.07 pg, Grandillo *et al.*, 2011).

464

465 ## DISCUSSION

466 Phylogenetic relationships among *Iberodes* species are robustly inferred by
467 RAD-sequencing, which unveiled the monophyly of populations into species except
468 for *I. linifolia*. Two early-diverging monophyletic species (*I. brassicifolia* and *I.*
469 *commutata*) contrast with paraphyly in the Linifolia clade of three species (Fig. 1).
470 The paraphyly of *I. linifolia* with respect to *I. kuzinskyana* and *I. littoralis* suggests a
471 progenitor-derivative relationship between the species, with the widespread
472 species giving rise to the coastal endemics in the course of a northward
473 colonization from the coast of Portugal to the west coast of France. Our results

474 provide a multiple lines of evidence that support parapatric divergence of the
475 coastland lineage (*I. kuzinskyana*-*I. littoralis* clade): (1) lack of any physical barrier
476 to gene flow between inland and coastland in the present or the past; (2) low level
477 of current gene flow (despite historical introgression) that support budding
478 differentiation rather secondary contact; (3) recent niche differentiation from
479 inland habitats with higher temperature seasonality to coastland ones with higher
480 precipitation; (4) paraphyly of *I. linifolia*, which can be interpreted as the extant
481 inland ancestor of the coastland lineage; and (5) incipient morphological
482 differentiation between *I. linifolia* and the coastal lineage. Subsequently,
483 differentiation in allopatry of the two coastal species during the mid-Pleistocene
484 led to the modern-day species *I. kuzinskyana* and *I. littoralis*, most likely mediated
485 by geographic isolation and niche differentiation. Thus ecological and geographic
486 differentiation contributed to both parapatric and allopatric speciation in the
487 clade.

488 **A predominant paraphyletic pattern during *Iberodes* evolution**

489 Our RAD-seq analyses revealed a pattern of predominant paraphyly within the
490 recently diverged Linifolia clade (Fig. 2). While the populations of *I. brassicifolia*
491 and those of *I. commutata* form monophyletic groups, paraphyly is inferred for *I.*
492 *linifolia* populations with respect to the two coastal species. While apparent
493 paraphyly of species may be a consequence of introgression dragging populations
494 toward the species to which they have introgressed (e.g., Eaton *et al.*, 2015), our
495 analyses of historical introgression do not conclusively support a role of this
496 process in determining the paraphyletic topology (Methods S1) and genetic
497 structure shows no evidence of recent admixture involving *I. linifolia* (Fig. 3c).
498 Despite the tendency in taxonomy to only consider monophyletic species,

499 paraphyletic ones have been reported to be particularly common in groups
500 showing a recent local speciation pattern, i.e. groups composed of a 'mother'
501 species that coexists in time and space with more restricted derivative species
502 (Rieseberg and Brouillet, 1994).

503 **Recent ecological speciation in parapatric *Iberodes***

504 Although parapatry has been argued to be common in plants (Rieseberg and
505 Brouillet, 1994), few empirical examples are found on the literature (Table 3). Our
506 study points to demonstrate parapatry through a process of budding (or
507 progenitor-derivative) speciation between *I. linifolia* and the ancestor of *I.*
508 *kuzinskyana* and *I. littoralis*. Indeed, the projection of distribution models to past
509 climatic periods during the Quaternary revealed a similar or greater degree of
510 geographic overlap between the potential ranges of *I. linifolia* and *I. kuzinskyana*
511 (Fig. 4), which makes allopatry unlikely (Zheng *et al.*, 2017). Peripheral
512 populations of the widespread *I. linifolia* close to the coast of Portugal are the most
513 closely related and possible living progenitor of coastland species (*I. kuzinskyana*
514 and *I. littoralis*; Fig. 2), which seem to have differentiated in an adjacent area with
515 no apparent physical barrier. Congruently, paraphyly underlies such budding
516 speciation event (Stuessy and Hörandl, 2013). In addition, despite the fact that
517 historical introgression was inferred among *I. linifolia* and coastland species
518 (Methods S1), it does not seem to underlie the phylogenetic gradation since
519 introgression is found indiscriminately with all lineages of *I. linifolia*. Moreover,
520 our results of AMOVA and DAPC pointed to lack of recent gene flow between the
521 nearby species (*I. kuzinskyana* and *I. linifolia*). This pattern supports strong
522 isolation rather than secondary contacts as suggested in other cases of ecological
523 differentiation (e.g. *Mimulus*, Stankowski *et al.*, 2015; *Stauracanthus*, Chozas *et al.*,

524 2017; Table 3). Nevertheless, a deeper population genetics study focused on the
525 adjacent populations of Portugal is necessary to obtain more detailed estimates of
526 gene flow.

527 In the earliest stages of parapatric speciation mediated by ecological
528 differentiation, reproductive barriers are expected to evolve in response to
529 selection towards locally adapted, divergent ecotypes (Sakaguchi *et al.*, 2019). The
530 niche identity tests points to a role of ecologically-based selection in the speciation
531 of the coastal lineage from *I. linifolia.* The coastal lineage is distributed in niches
532 with less temperature seasonality and higher precipitation means than *I. linifolia*
533 (Fig. 3b). Interestingly, although niches are not equivalent and overlap is close to
534 zero, similarity tests indicated a certain degree of similarity within the E-space,
535 which points to recent divergence and suggests ongoing differentiation or some
536 degree of niche conservatism. Besides, edaphic factors such as the specialization to
537 coastal dunes substrates can have also contributed to ecological differentiation,
538 although this is a fine-scale soil variable that is difficult to include in niche models.
539 The differentiation of coastal lineages (*I. kuzinskyana* and *I. littoralis*) from inland
540 populations of *I. linifolia* seems to have started in western Iberian Peninsula during
541 the Mid Pliocene (3.85-3.60 myr, Fig. 2), when increasing seasonality led to the
542 origin of Mediterranean climate with its summer droughts (Suc, 1984). The onset
543 of Mediterranean climate favored the differentiation and configuration of modern
544 Mediterranean flora (Vargas *et al.*, 2018), which agrees with recent *Iberodes*
545 speciation (Fig. 2). Posterior glacial/interglacial periods during the Late Pliocene
546 promoted the expansion and contraction of geographic ranges and the formation
547 of sources and sinks of genetic diversity for most Mediterranean species (Médail
548 and Diadema, 2009). Interestingly, the overlapping area inferred for *I. linifolia* and

549 *I. kuzinskyana* in the past is congruent with the location of glacial refugia
550 (Estremadura and Beira litoral) proposed by Médail and Diadema (2009) and plant
551 species previously studied as glacial refugial species (e.g. *Drosophyllum*
552 *lusitanicum*, Müller and Deil, 2001; *Cistus monspeliensis*, Fernández-Mazuecos and
553 Vargas, 2010; *Linaria elegans*, Fernández-Mazuecos and Vargas, 2013).

554 In addition, ecological differentiation in parapatric speciation involves
555 divergent natural selection towards contrasting environments that promotes
556 adaptation to different environments and subsequent barriers to gene flow
557 (Rundle and Nosil, 2005; Schluter, 2001). The strength of divergent selection can
558 modulate the degree of completeness of a speciation process (Nosil *et al.*, 2009). In
559 our case, both the high interspecific genetic differentiation and clear
560 morphological and ecological divergence (Fig. 3) between inland and coastal
561 species may reflect strong divergent selection promoting rapid speciation
562 (Pliocene-Pleistocene, Fig. 2). Indeed, our study reinforces the common trend for
563 Mediterranean plants in which narrow endemics arise from ecological
564 differentiation of peripheral populations of more widespread progenitor species
565 (Papuga *et al.*, 2018). Moreover, this trend highlights that this peripheral
566 populations may set the scene for diversification, thus revealing their conservation
567 value (Debussche and Thompson, 2003).

568

569 **Allopatric differentiation along the Atlantic coast**

570 An allopatric pattern is also found in the Linifolia clade. Given the current
571 separation between ranges of the three coastal taxa (Fig. 1), allopatry seems seems
572 to have triggered speciation process. On the one hand, the widely separated ranges
573 of *I. kuzinskyana* and *I. littoralis* subsp. *gallaecica* both in current times and in

574 projections to LIG (despite probable intermittent contact during LGM) suggest
575 either a colonization mediated by LDD followed by geographic isolation and niche
576 shift (Fig. 3, 4) or a vicariant process mediated by an early expansion and
577 contraction of distribution ranges. We find similar examples of recent allopatric
578 speciation for both hypothesis: by LDD among coastal Mediterranean flora in the
579 same period (e.g. Jakob *et al.*, 2007; Carnicero *et al.*, 2017; Herrando-Moraira *et al.*,
580 2017); and expansion/contraction of distribution ranges (e.g. Ortiz *et al.*, 2007; Lo
581 Presti and Oberprieler, 2011; Fernández-Mazuecos and Vargas, 2011). On the
582 other hand, the allopatric differentiation of the two subspecies of *I. littoralis* was
583 likely preceded by an expansion of *I. littoralis* during the LGM, followed by
584 geographic isolation as its range contracted and niche shifted, leading to the
585 differentiation of *I. littoralis* subsp. *littoralis* in more recent times (Fig. 3, 4). In this
586 case, the probable role of long distance dispersal (LDD) is further supported by the
587 fact that the expanded area during the LGM does not seem to have formed a
588 continuous range (Fig. 4). Indeed, unlike inland species of *Iberodes*, the three
589 coastal taxa seem to have LDD specializations in the form of uncinated hooks (Fig.
590 1) in the nutlet, related to the attachment to feathers and fur in other closely
591 related Boraginaceae genera (Selvi *et al.*, 2011).

592 **CONCLUSIONS**

593 Our integrative approach has let us overcome previous barriers to understanding
594 parapatric speciation. The combination RAD-sequencing, morphometry and
595 climatic niche has provided notable evidences of parapatry revealing this
596 workflow as a powerful resource for scenarios of non-allopatric processes in other
597 groups of plants. In particular, RAD-sequencing once again is revealed as a useful
598 tool to disentangle recent speciation events in non-model organisms.

Consequently, we have empirically shown the recurrence of progenitor-derivative speciation in plants as previously suggested. Processes such us budding speciation bring to light the need to extend traditional species concepts and advocate the consideration of paraphyletic species as natural units of taxonomical classification.

ACKNOWLEDGEMENTS

We would like to thank those botanists and colleagues that help us with material sampling and particularly, Diana Íñigo, Paola Pérez, Celia Otero, David Vallecillo, Yurena Arjona, Lua López, Joaquín Ramírez, George Hinton, Javier Morente, Manuel Joao Pinto, Ruben Retuerto, Miguel Serrano, Antonio Rivas, Claude Dauge, Asli Koca, Peter Heenan, Enrique Rico, Ester Vega, Santiago Martín-Bravo, Íñigo Pulgar, and Juan Carlos Zamora. Likewise, we would also appreciate the worthy loans provided by numerous herbaria mentioned at supplementary tables. We gratefully acknowledge the work of the lab technicians of Real Jardín Botánico de Madrid, Yolanda Ruiz, Emilio Cano, Jose Gónzalez, Olga Popova, and Lucía Sastre. We also thank Lucía Checa and Miriam Pérez for the cytometry lab work. We are grateful for the many methodological comments of John McVay, Irene Villa-Machío, Yurena Arjona, Marcial Escudero, Tamara Villaverde, Javier Fernández-López and Alberto Coello. Finally, we would like to thank the researchers of the Herbarium and Center for Tree Science at The Morton Arboretum (Lisle, Illinois, USA), in particular Marlene Hahn for lab assistance, and the Botany Department of the Field Museum (Chicago, Illinois, USA) for accommodating me at the Field Museum and handling the hosting during the lab stay in Chicago. These data were initially analyzed during a research stay by the first author in the Herbarium of The Morton Arboretum and The Field Museum. This study was supported by the Fundación General CSIC and Banco Santander as part of the project titled “Do all endangered

624 species hold the same value?: origin and conservation of living fossils of flowering
625 plants endemic to Spain."

626 AUTHOR CONTRIBUTIONS

627 AO, PJM, VV and PV contributed to developing the question and experimental
628 design. AO led the material collection. AO, MFM and ALH conducted the data
629 analysis. AO wrote the manuscript with the contribution of all authors in both
630 interpretation and writing.

631 REFERENCES

632 Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. Baird, N. Bierne, J. Boughman, *et*
633 *al.* 2013. Hybridization and speciation. *Journal of evolutionary biology* 26:
634 229-246.

635 Anacker, B. L., and S. Y. Strauss. 2014. The geography and ecology of plant
636 speciation: range overlap and niche divergence in sister species.
637 *Proceedings of the Royal Society B: Biological Sciences* 281: 20132980.

638 Andrew, R. L., and L. H. Rieseberg. 2013. Divergence is focused on few genomic
639 regions early in speciation: incipient speciation of sunflower ecotypes.
640 *Evolution* 67: 2468-2482.

641 Arnold, M. L. 2015. Divergence with genetic exchange. Oxford University Press.
642 New York, USA.

643 Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U.
644 Selker, *et al.* 2008. Rapid SNP discovery and genetic mapping using
645 sequenced RAD markers. *PLoS One* 3: 1-7.

646 Barluenga, M., K. N. Stölting, W. Salzburger, M. Muschick, and A. Meyer. 2006.
647 Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439: 719.

648 Barton, N. H. 2013. Does hybridization influence speciation? *Journal of evolutionary*
649 *biology* 26: 267-269.
650 Benítez-Benítez, C., M. Escudero, F. Rodríguez-Sánchez, S. Martín-Bravo, and P.
651 Jiménez-Mejías. 2018. Pliocene–Pleistocene ecological niche evolution
652 shapes the phylogeography of a Mediterranean plant group. *Molecular*
653 *ecology* 27: 1696-1713.
654 Brand, A. 1921. Omphalodes Moench. *In* A. Engler [ed.], Das Pflanzenreich, vol. 78,
655 96-112. Verlag von Wilhelm Engelman, Leipzig.
656 Brayshaw, D. J., C. M. Rambeau, and S. J. Smith. 2011. Changes in Mediterranean
657 climate during the Holocene: Insights from global and regional climate
658 modelling. *The Holocene* 21: 15-31.
659 Broennimann, O., M. C. Fitzpatrick, P. B. Pearman, B. Petitpierre, L. Pellissier, N. G.
660 Yoccoz, W. Thuiller, *et al.* 2012. Measuring ecological niche overlap from
661 occurrence and spatial environmental data. *Global Ecology and*
662 *Biogeography* 21: 481-497.
663 Butlin, R. K., J. Galindo, and J. W. Grahame. 2008. Review. Sympatric, parapatric or
664 allopatric: the most important way to classify speciation? *Philosophical*
665 *Transactions of the Royal Society B: Biological Sciences* 363: 2997-3007.
666 Carnicero, P., L. Sáez, N. Garcia-Jacas, and M. Galbany-Casals. 2017. Different
667 speciation types meet in a Mediterranean genus: The biogeographic history
668 of *Cymbalaria* (Plantaginaceae). *Taxon* 66: 393-407.
669 Cavender-Bares, J. 2019. Diversification, adaptation, and community assembly of
670 the American oaks (*Quercus*), a model clade for integrating ecology and
671 evolution. *New Phytologist* 221: 669–692.

672 Crowl, A. A., P. S. Manos, J. D. Mc Vay, A. R. Lemmon, E. M. Lemmon, and A. L. Hipp.
673 2019. Uncovering the genomic signature of ancient introgression between
674 white oak lineages (*Quercus*). New Phytologist.
675 https://doi.org/10.1111/nph.15842.
676 Chacón, J., F. Luebert, and M. Weigend. 2017. Biogeographic Events Are Not
677 Correlated with Diaspore Dispersal Modes in Boraginaceae. *Frontiers in*
678 *Ecology and Evolution* 5.
679 Chifman, J., and L. Kubatko. 2014. Quartet inference from SNP data under the
680 coalescent model. *Bioinformatics* 30: 3317-3324.
681 Chozas, S., R. M. Chefaoui, O. Correia, R. Bonal, and J. Hortal. 2017. Environmental
682 niche divergence among three dune shrub sister species with parapatric
683 distributions. *Annals of botany* 119: 1157-1167.
684 Coyne, J. A., and H. A. Orr. 2004. Speciation. Sunderland, MA. Sinauer Associates,
685 Inc.
686 Crawford, D. J. 2010. Progenitor-derivative species pairs and plant speciation.
687 *Taxon* 59: 1413-1423.
688 de Laat, A., W. Gohde, and M. Vogelzakg. 1987. Determination of ploidy of single
689 plants and plant populations by flow cytometry. *Plant breeding* 99: 303-307.
690 Debussche, M. and J. D. Thompson. 2003. Habitat differentiation between two
691 closely related Mediterranean plant species, the endemic *Cyclamen*
692 *balearicum* and the widespread *C. repandum*. *Acta Oecologica* 24: 35–45.
693 Di Cola, V., O. Broennimann, B. Petitpierre, F. T. Breiner, M. D'amen, C. Randin, R.
694 Engler, *et al.* 2017. ecospat: an R package to support spatial analyses and
695 modeling of species niches and distributions. *Ecography* 40: 774-787.

696 Doležel, J., and J. Bartoš. 2005. Plant DNA flow cytometry and estimation of nuclear
697 genome size. *Annals of botany* 95: 99-110.
698 Doležel, J., J. Greilhuber, and J. Suda. 2007. Estimation of nuclear DNA content in
699 plants using flow cytometry. *Nature protocols* 2: 2233.
700 Donoghue, M. J., and E. J. Edwards. 2014. Biome Shifts and Niche Evolution in
701 Plants. *Annual Review of Ecology, Evolution, and Systematics* 45: 547-572.
702 Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G.,
703 Gruber, B., Lafourcade, B., Leitão, P.J., and Münkemüller, T. 2013.
704 Collinearity: a review of methods to deal with it and a simulation study
705 evaluating their performance. *Ecography* 36: 27-46.
706 Doyle, J., and J. Doyle. 1987. A rapid DNA isolation procedure for small quantities of
707 fresh leaf tissue. *Phytochemistry Bulletin* 19: 11—15., and AHD BROWN.
708 1990. Chlo roplast DNA polymorphism and phylogeny in the B genome of
709 Glycine subgenus Glycine (Le guminosae). *Amer. J. Bot* 77: 772-782.
710 Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian
711 phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and*
712 *evolution* 29: 1969-1973.
713 Eaton, D. A. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic
714 analyses. *Bioinformatics* 30: 1844-1849.
715 Eaton, D. A., A. L. Hipp, A. González-Rodríguez, and J. Cavender-Bares. 2015.
716 Historical introgression among the American live oaks and the comparative
717 nature of tests for introgression. Evolution 69: 2587–2601.
718 Excoffier, L., and H. E. Lischer. 2010. Arlequin suite ver 3.5: a new series of
719 programs to perform population genetics analyses under Linux and
720 Windows. *Molecular Ecology Resources* 10: 564-567.

721 Feder, J. L., and P. Nosil. 2009. Chromosomal inversions and species differences:
722 when are genes affecting adaptive divergence and reproductive isolation
723 expected to reside within inversions? *Evolution: International Journal of*
724 *Organic Evolution* 63: 3061-3075.
725 Ferguson, L. 1972. Omphalodes Miller. *In* T. G. Tutin, V. H. Heywood, N. A. Burges,
726 D. M. Moore, D. H. Valentine, S. M. Walters, AND D. A. Webb [eds.], Flora
727 Europaea, vol. 3, 118-119. Cambridge University Press, Cambridge.
728 Fernández, I., and S. Talavera. 2012. Omphalodes Mill. *In* S. Talavera, C. Andrés, M.
729 Arista, M. Fernández Piedra, M. Gallego, P. Ortiz, C. Romero Zarco, *et al.*
730 [eds.], Flora Iberica, vol. 11, 471-479. Real Jardín Botánico, CSIC, Madrid.
731
732 Fernández-Casas, J. 1975. "Números cromosomicos de plantas españolas."
733 [Chromosome numbers of Spanish plants] *Anales del Instituto Botánico*
734 *Cavanilles* 32: 301–307.
735 Fernández-López, J., and I. Villa-Machío. 2017. NiceOverPlot, or when the number
736 of dimensions does matter (R script) [online]. Website
737 https://allthiswasfield.blogspot.com.es/2017/05/niceoverplot-or-when-
738 number-of.html [accessed 20-12-2018].
739 Fernandez-Mazuecos, M., and Vargas, P. 2011. Historical isolation versus recent long-
740 distance connections between Europe and Africa in bifid toadflaxes (Linaria sect.
741 Versicolores). *PLoS One* 6: e22234.
742 Fernández-Mazuecos, M., and P. Vargas. 2013. Congruence between distribution
743 modelling and phylogeographical analyses reveals Quaternary survival of a
744 toadflax species (Linaria elegans) in oceanic climate areas of a mountain
745 ring range. *New Phytologist* 198: 1274-1289.

746 Fernández-Mazuecos, M., G. Mellers, B. Vigalondo, L. Sáez, P. Vargas, and B. J.
747 Glover. 2018. Resolving recent plant radiations: power and robustness of
748 genotyping-by-sequencing. *Systematic biology* 67: 250-268.
749 Fernández-Mazuecos, M., and P. Vargas. 2010. Ecological rather than geographical
750 isolation dominates Quaternary formation of Mediterranean Cistus species.
751 *Molecular ecology* 19: 1381-1395.
752 Fitz-Gibbon, S., A. L. Hipp, K. K. Pham, P. S. Manos, and V. Sork. 2017. Phylogenomic
753 inferences from reference-mapped and de novo assembled short-read
754 sequence data using RADseq sequencing of California white oaks (*Quercus*
755 subgenus *Quercus*). *Genome* 60: 743–755.
756 Fitzpatrick, B., J. Fordyce, and S. Gavrilets. 2008. What, if anything, is sympatric
757 speciation? *Journal of evolutionary biology* 21: 1452-1459.
758 Fitzpatrick, B. M., J. A. Fordyce, and S. Gavrilets. 2009. Pattern, process and
759 geographic modes of speciation. *Journal of evolutionary biology* 22: 2342-
760 2347.
761 Folk R. A., P. S. Soltis, D. E. Soltis, and R. Guralnick 2018. New prospects in the
762 detection and comparative analysis of hybridization in the tree of life.
763 *American Journal of Botany* 105: 364–375.
764 Fontdevila, A. 2014. Speciation. *In* P. Vargas AND R. Zardoya [eds.], The tree of life,
765 445-456. Sinauer Associates, Sunderland, MA, USA.
766 Franco, J.A. 1984. Nova Flora de Portugal (Continente e Açores). [New Flora of
767 Portugal (Mainland and Azores)]: Vol. 2, Clethraceae-Compositae. Lisboa:
768 Instituto Superior de Agronomia.
769 Friesen, V. L. 2015. Speciation in seabirds: why are there so many species…and
770 why aren't there more? *Journal of Ornithology* 156: 27-39.

Futuyma, D. J. 2009. Evolution. Sinauer, Sunderland, MA, USA.

Gavrilets, S. 2003. Perspective: models of speciation: what have we learned in 40 years? *Evolution* 57: 2197–2215.

Grandillo, S., R. Chetelat, S. Knapp, D. Spooner, I. Peralta, M. Cammareri, O. Pérez, *et al.* 2011. Solanum sect. Lycopersicon. *In* C. Kole [ed.], Wild Crop Relatives: Genomic and Breeding Resources, 129-215. Springer, Berlin, Heidelberg.

Haeckel, E. 1866. Generelle Morphologie der Organismen allgemeine Grundzuge der organischen Formen-Wissenschaft, mechanisch begrundet durch die von Charles Darwin reformirte Descendenz-Theorie von Ernst Haeckel: Allgemeine Entwickelungsgeschichte der Organismen kritische Grundzuge der mechanischen Wissenschaft von den entstehenden Formen der Organismen, begrundet durch die Descendenz-Theorie. Verlag von Georg Reimer, Berlin.

Harrison, R. G. 1998. Linking Evolutionary Pattern and Process. The Relevance ofSpecies Concepts for the Study ofSpeciation. *In* D. J. Howard AND S. H. Berlocher [eds.], Endless forms: species and speciation, 19–31 Oxford University Press, New York, USA.

Heiberger, R. M. (2017). HH: Statistical analysis and data display: Heiberger and Holland. R package v. 3.1-34.

Heleno, R., and P. Vargas. 2015. How do islands become green? *Global Ecology and Biogeography* 24: 518-526.

Hennig, W. 1966. Phylogenetic systematics. University of Illinois Press, Urbana.

Herrando-Moraira, S., P. Carnicero, J. M. Blanco-Moreno, L. Sáez, E. Véla, R. Vilatersana, and M. Galbany-Casals. 2017. Systematics and phylogeography of the Mediterranean *Helichrysum pendulum* complex (Compositae) inferred

796 from nuclear and chloroplast DNA and morphometric analyses. *Taxon* 66:
797 909-933.
798 Hijmans, R. J., S. Cameron, J. Parra, P. Jones, A. Jarvis, and K. Richardson. 2006.
799 World-Clim version 1.4. *Museum of Vertebrate Zoology of the University of*
800 *California, CIAT, and Rainforest CRC.*
801 Hipp, A. 2014. RADami: R package for phylogenetic analysis of RADseq data. *R*
802 *package version*: 1.0-3.
803 Holstein, N., J. Chacon, H. H. Hilger, and M. Weigend. 2016a. No longer
804 shipwrecked—Selkirkia (Boraginaceae) back on the mainland with generic
805 rearrangements in South American "Omphalodes" based on molecular data.
806 *Phytotaxa* 270: 231-251.
807 Holstein, N., J. Chacon, A. Otero, P. Jiménez-mejías, and M. Weigend. 2016b.
808 Towards a monophyletic Omphalodes—or an expansion of North American
809 Mimophytum. *Phytotaxa* 288: 131-144.
810 Hörandl, E., and T. F. Stuessy. 2010. Paraphyletic groups as natural units of
811 biological classification. *Taxon* 59: 1641–1653.
812 Huang, C. L., C. W. Ho, Y. C. Chiang, Y. Shigemoto, T. W. Hsu, C. C. Hwang, X. J. Ge, *et*
813 *al.* 2014. Adaptive divergence with gene flow in incipient speciation of
814 Miscanthus floridulus/sinensis complex (Poaceae). *Plant Journal* 80: 834-
815 847.
816 ICNB. 2007. Plano Nacional de Conservaçao da Flora em Perigo (1º Fase). Relatorio
817 Final. Volume IV. Instituto da Conservaçao da Naturaleza.
818 INPN. 2019. Liste Rouge de la Flore Vasculaire de Poitou-Charentes [online].
819 Website

820 https://inpn.mnhn.fr/espece/listerouge/RG/LRR_flore_vasculaire_Poitou_C
821 harentes_2018
822 IUCN. 2019. IUCN Red List of Threatened Species. Version 2019.1. [online].
823 Website www.iucnredlist.org.
824 Jakob, S. S., A. Ihlow, and F. R. Blattner. 2007. Combined ecological niche modelling
825 and molecular phylogeography revealed the evolutionary history of
826 Hordeum marinum (Poaceae)--niche differentiation, loss of genetic
827 diversity, and speciation in Mediterranean Quaternary refugia. *Molecular*
828 *ecology* 16: 1713–1727.
829 Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic
830 markers. *Bioinformatics* 24: 1403-1405.
831 Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal
832 components: a new method for the analysis of genetically structured
833 populations. *BMC genetics* 11: 94.
834 Kautt, A. F., G. Machado-Schiaffino, J. Torres-Dowdall, and A. Meyer. 2016. Incipient
835 sympatric speciation in Midas cichlid fish from the youngest and one of the
836 smallest crater lakes in Nicaragua due to differential use of the benthic and
837 Knaus, B. J., and N. J. Grünwald. 2017. VCFR: a package to manipulate and visualize
838 variant call format data in R. *Molecular Ecology Resources* 17: 44-53.
839 Leroy, T., C. Roux, L. Villate, C. Bodénès, J. Romiguier, J. A. P. Paiva, C. Dossat, J-M.
840 Aury, C. Plomion and A. Kremer. 2017. Extensive recent secondary contacts
841 between four European white oak species. *New Phytologist* 214: 865–878.
842 Lopes, M. H. R., M. L. S. Carvalho, and A. R. P. Silva. 1990. Lista de espécies
843 botânicas a proteger em Portugal continental. SNPRCN. Lisboa.

844 Lo Presti, R. M., and Oberprieler, C. 2011. The central Mediterranean as a
845 phytodiversity hotchpotch: phylogeographical patterns of the Anthemis
846 secundiramea group (Compositae, Anthemideae) across the Sicilian
847 Channel. *Journal of biogeography* 38: 1109-1124.
848 Maguilla, E., M. Escudero, A. L. Hipp, and M. Luceño. 2017. Allopatric speciation
849 despite historical gene flow: Divergence and hybridization in Carex furva
850 and C. lucennoiberica (Cyperaceae) inferred from plastid and nuclear
851 RAD-seq data. *Molecular ecology* 26: 5646-5662.
852 Mayr, E. 1954. Change of genetic environment and evolution. Pages 157-180 in J.
853 Huxley, A. C. Hardy, and E. B. Ford, eds. Evolution as a process. Allen &
854 Unwin, London.
855 Médail, F., and K. Diadema. 2009. Glacial Refugia Influence Plant Diversity Patterns
856 in the Mediterranean Basin. *Journal of Biogeography* 36: 1333-1345.
857 Mc Cormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens and R. T. Brumfield. 2013.
858 Applications of next-generation sequencing to phylogeography and
859 phylogenetics. *Molecular Phylogenetics and Evolution* 66: 526–538.
860 Mc Vay, J. D., D. Hauser, A. L. Hipp, and P. S. Manos. 2017a. Phylogenomics reveals a
861 complex evolutionary history of lobed-leaf white oaks in western North
862 America. *Genome* 60: 733–742.
863 Mc Vay, J. D., A. L. Hipp, and P.S. Manos. 2017b. A genetic legacy of introgression
864 confounds phylogeny and biogeography in oaks. *Proceedings of the Royal*
865 *Society B* 284: 20170300.
866 Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES Science
867 Gateway for inference of large phylogenetic trees. *Proceedings of the*
868 *Gateway Computing Environments Workshop*: 1-8.

869 Mishler, B. D., and M. J. Donoghue. 1982. Species Concepts: A Case for Pluralism.
870 *Systematic Zoology* 31: 491-503.
871 Moreno, J.C., coord. 2008. Lista Roja 2008 de la flora vascular española. Dirección
872 General de Medio Natural y Política Forestal (Ministerio de Medio
873 Ambiente, y Medio Rural y Marino) y Sociedad Española de Biología de la
874 Conservación de Plantas, Madrid.
875 Müller, J., and U. Deil. 2001. Ecology and structure of *Drosophyllum lusitanicum* (L.)
876 link populations in the southwestern of the Iberian Peninsula. *Acta Botanica*
877 *Malacitana* 26: 47-68.
878 Muséum national d'Histoire naturelle. 2003-2019. Omphalodes littoralis Lehm.
879 Inventaire National du Patrimoine Naturel [online]. Website
880 https://inpn.mnhn.fr/espece/cd_nom/110070.
881 Neto, C., J. P. Fonseca, J. C. Costa, and F. Bioret. 2015. Ecology and phytosociology of
882 endangered psammophytic species of the *Omphalodes* genus in western
883 Europe. *Acta Botanica Gallica* 162: 37-54.
884 Nosil, P. 2012. Ecological speciation. Oxford University Press, New York, USA.
885 Nosil, P., L. J. Harmon, and O. Seehausen. 2009. Ecological explanations for
886 (incomplete) speciation. *Trends in ecology and evolution* 24: 145-156.
887 Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. O'hara, G. L.
888 Simpson, *et al.* 2013. Package 'vegan'. *Community ecology package, version* 2.
889 Orr, H. A. 1991. Is single-gene speciation possible? *Evolution* 45: 764-769.
890 Ortiz, M. A., Tremetsberger, K., Talavera, S., Stuessy, T., & García-Castaño, J. L. 2007.
891 Population structure of Hypochaeris salzmanniana DC.(Asteraceae), an
892 endemic species to the Atlantic coast on both sides of the Strait of Gibraltar,
893 in relation to Quaternary sea level changes. *Molecular Ecology* 16: 541-552

894 Otero, A., P. Jiménez-Mejías, V. Valcárcel, and P. Vargas. 2014. Molecular
895 phylogenetics and morphology support two new genera (*Memoremea* and
896 *Nihon*) of Boraginaceae s.s. *Phytotaxa* 173: 241-277.
897 Otero, A., P. Jiménez-Mejías, V. Valcárcel, and P. Vargas. 2019a. Being in the right
898 place at the right moment? Parallel diversification bursts related to
899 persistence of ancient epizoochorous traits and hidden factors in
900 Cynoglossoideae. *American Journal of Botany* 106: 438-452.
901 Otero, A., P. Jiménez-Mejías, V. Valcárcel, and P. Vargas. 2019b. Worldwide long
902 distance dispersal favored by epizoochorous traits in the biogeographic
903 history of Omphalodeae (Boraginaceae). *Journal of Systematics and*
904 *Evolution* https://doi.org/10.1111/jse.12504.
905 Otto-Bliesner, B. L., E. C. Brady, G. Clauzet, R. Tomas, S. Levis, and Z. Kothavala.
906 2006. Last glacial maximum and Holocene climate in CCSM3. *Journal of*
907 *Climate* 19: 2526-2544.
908 Papadopoulou, A., and L. L. Knowles. 2015. Genomic tests of the species-pump
909 hypothesis: recent island connectivity cycles drive population divergence
910 but not speciation in Caribbean crickets across the Virgin Islands. *Evolution*
911 69: 1501-1517.
912 Papuga, G., P. Gauthier, V. Pons, E. Farris, J. D. Thompson. 2018. Ecological niche
913 differentiation in peripheral populations: a comparative analysis of eleven
914 Mediterranean plant species. *Ecography* 41: 1650-1664.
915 Paule, J., N. D., Wagner, K. Weising, and G. Zizka. 2017. Ecological range shift in the
916 polyploid members of the South American genus *Fosterella* (Bromeliaceae).
917 *Annals of Botany* 120: 233-243.

Peterson, A. T., S. J., and V. Sánchez-Cordero. 1999. Conservatism of Ecological Niches in Evolutionary Time. *Science* 285: 1265-1267.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological modelling* 190: 231-259.

Pyron, A. R., and F. T. Burbrink. 2010. Hard and soft allopatry: physically and ecologically mediated modes of geographic speciation. *Journal of Biogeography* 10.1111/j.1365-2699.2010.02336.x DOI: no-no.

RCoreTeam. 2013. R: a Language and Environment for Statistical Computing, version Version 3.2.3., website: https://www.r-project.org/ [accessed 12 June 2013].

Richards, T. J., and Ortiz-Barrientos, D. 2016. Immigrant inviability produces a strong barrier to gene flow between parapatric ecotypes of *Senecio lautus*. *Evolution*, 70:1239-1248.

Rieseberg, L. H., and L. Brouillet. 1994. Are many plant species paraphyletic? *Taxon* 43: 21-32.

Rosenberg, N. H. 2003. The shapes of neutral gene genealogies in two species: Probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. *Evolution* 57: 1465–1477.

Rundle, H. D., and P. Nosil. 2005. Ecological speciation. *Ecology letters* 8: 336-352.

Sakaguchi, S., K. Horie, N. Ishikawa, S. Nishio, J. R. P. Worth, K. Fukushima, M. Yamasaki, *et al.* 2019. Maintenance of soil ecotypes of Solidago virgaurea in close parapatry via divergent flowering time and selection against immigrants. *Journal of Ecology* 107: 418-435.

941 Saly, F. 1997. "Étude botanique, cytogénétique et pédologique de l'arc dunaire
942 Gavres-Quiberon. Incidences sur la conservation du patrimoine végétal
943 sauvage." PhD diss., Muséum national d'histoire naturelle de Paris.
944 Savolainen, V., M.-C. Anstett, C. Lexer, I. Hutton, J. J. Clarkson, M. V. Norup, M. P.
945 Powell, *et al.* 2006. Sympatric speciation in palms on an oceanic island.
946 *Nature* 441: 210.
947 Schluter, D. 2001. Ecology and the origin of species. *Trends in ecology and evolution*
948 16: 372-380.
949 Schoener, T. W. 1968. The Anolis lizards of Bimini: resource partitioning in a
950 complex fauna. *Ecology* 49: 704-726.
951 Schulte, L. J., J. L. Clark, S. J. Novak, S. K. Jeffries, and J. F. Smith. 2015. Speciation
952 within *Columnea* section *Angustiflora* (Gesneriaceae): islands, pollinators
953 and climate. *Molecular phylogenetics and evolution* 84: 125-144.
954 Selvi, F., A. Coppi, and L. Cecchi. 2011. High epizoochorous specialization and low
955 DNA sequence divergence in mediterranean *Cynoglossum* (Boraginaceae):
956 Evidence from fruit traits and ITS region. *Taxon* 60: 969-985.
957 Serrano, M., R. Carbajal. 2003. *Omphalodes littoralis* subsp. *gallaecica* M. Laínz. *In*
958 Bañares, Á., G. Blanca, J. Güemes, J.C. Moreno AND S. Ortiz, [eds.] Atlas y
959 Libro Rojo de la Flora Vascular Amenazada de España: 274-275. Dirección
960 General de Conservación de la Naturaleza, Madrid.
961 Serrano, M., R. Carbajal, A. Pereira-Coutinho, and S. Ortiz. 2016. Two new genera in
962 the *Omphalodes* group (Cynoglosseae, Boraginaceae). *Nova Acta Científica*
963 *Compostelana (Bioloxía)* 23: 1-14.

964 Shepherd, L. D., and T. G. Mc Lay. 2011. Two micro-scale protocols for the isolation
965 of DNA from polysaccharide-rich plant tissue. *Journal of Plant Research* 124:
966 311-314.
967 Smith, S. A., and B. C. O'meara. 2012. treePL: divergence time estimation using
968 penalized likelihood for large phylogenies. *Bioinformatics* 28: 2689-2690.
969 Sobel, J. M., and M. A. Streisfeld. 2015. Strong premating reproductive isolation
970 drives incipient speciation in Mimulus aurantiacus. *Evolution* 69: 447-461.
971 Sobel, J. M., G. F. Chen, L. R. Watt, and D. W. Schemske. 2010. The biology of
972 speciation. *Evolution* 64: 295-315.
973 Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-
974 analysis of large phylogenies. *Bioinformatics* 30: 1312-1313.
975 Stankowski, S., J. M. Sobel, and M. A. Streisfeld. 2015. The geography of divergence
976 with gene flow facilitates multitrait adaptation and the evolution of
977 pollinator isolation in Mimulus aurantiacus. *Evolution* 69: 3054-3068.
978 Stuessy, T. F., and E. Hörandl. 2013. The importance of comprehensive
979 phylogenetic (evolutionary) classification—a response to Schmidt-Lebuhn's
980 commentary on paraphyletic taxa. *Cladistics*: 1-3.
981 Stuessy, T. F., and E. Hörandl. 2014. Evolutionary Systematics and Paraphyly:
982 Introduction. *Annals of the Missouri Botanical Garden* 100: 2-5.
983 Swofford, D. L. 2001. PAUP*: Phylogenetic analysis using parsimony (and other
984 methods) 4.0. B5.
985 Talavera, S., C. Andrés, M. Arista, M. Fernández Piedra, M. Gallego, P. Ortiz, C.
986 Romero Zarco, *et al.* 2012. Boraginaceae. *In* S. Talavera, C. Andrés, M. Arista,
987 M. Fernández Piedra, M. Gallego, P. Ortiz, C. Romero Zarco, *et al.* [eds.], Flora
988 Iberica, vol. 11, 324-532. Real Jardín Botánico, CSIC, Madrid.

989 Templeton, A. R. 1981. Mechanisms of speciation- a population genetic approach.
990 *Annual Review of Ecology and Systematics* 12: 23-48.
991 Thompson, J. D. 2005. Plant evolution in the Mediterranean. Oxford University
992 Press, New York, USA.
993 Vargas, P., and R. Zardoya. 2014. The tree of life. Sinauer Associates Sunderland,
994 MA, USA.
995 Vargas, P., M. Fern, and R. Heleno. 2018. Phylogenetic evidence for a Miocene
996 origin of Mediterranean lineages : species diversity , reproductive traits and
997 geographical isolation 20: 157-165.
998 Vu, V. Q. 2011. ggbiplot: A ggplot2 based biplot. *R package*.
999 Warren, D. L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency
1000 versus conservatism: quantitative approaches to niche evolution. *Evolution*
1001 62: 2868-2883.
1002 Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer,
1003 Switzerland.
1004 Wiens, J. J., and C. H. Graham. 2005. Niche Conservatism: Integrating Evolution,
1005 Ecology, and Conservation Biology. *Annual Review of Ecology, Evolution, and*
1006 *Systematics* 36: 519-539.
1007 Yin, H., X. Yan, W. Zhang, Y. Shi, C. Qian, C. Yin, F. Tian, *et al.* 2016. Geographical or
1008 ecological divergence between the parapatric species Ephedra sinica and E.
1009 intermedia? *Plant Systematics and Evolution* 302: 1157-1170.
1010 Zhao, J. L., P. F. Gugger, Y. M. Xia, and Q. J. Li. 2016. Ecological divergence of two
1011 closely related *Roscoea* species associated with late Quaternary climate
1012 change. *Journal of biogeography* 43: 1990-2001.

1013 Zheng, H., L. Fan, R. I. Milne, L. Zhang, Y. Wang, and K. Mao. 2017. Species
1014 Delimitation and Lineage Separation History of a Species Complex of Aspens
1015 in China. *Fronters in Plant Science* 8: 375.
1016 Zheng, X. M., and S. Ge. 2010. Ecological divergence in the presence of gene flow in
1017 two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular*
1018 *Ecology* 19: 2439-2454.
1019 Zheng, Y., and J. J. Wiens. 2015. Do missing data influence the accuracy of
1020 divergence-time estimation with BEAST? *Molecular Phylogenetics and*
1021 *Evolution* 85: 41-49.
1022

1023 **FIGURE LEGENDS**

1024 **Fig. 1** Locations of the *Iberodes* species sampled for the different analyses
1025 performed. The star symbols indicate locations for RAD-sequencing, squares
1026 represent the locations for the morphometric analysis, and circles indicate
1027 locations for the ecological niche modelling. Color legend for each taxa of *Iberodes*
1028 is shown in the figure. Images of the six taxa are framed following each taxa color
1029 in the map.
1030 **Fig. 2** Time-calibrated phylogeny obtained from TreePL by using maximum clade
1031 credibility from 1000 bootstrap trees of the maximum likelihood analysis of
1032 c95m20 matrix. Branch thickness indicates the ML bootstrap supports above 98.
1033 Node bars in blue indicate the node age ranges taking into account the branch
1034 length variance along the 1000 bootstrap trees. No branch length variation was
1035 inferred for nodes without blue bar. The international stratigraphic scale is
1036 included from 26 myr until present. Text codes for each individual location are
1037 detailed in the Table S1. The color for each lineage of the Linifolia clade is

1038 geographycally represented through the colored ellipses of the map. Grey ellipses
1039 group the different lineages of the four taxa within the Linifolia clade.
1040 **Fig. 3** Morphometric, niche overlap and genetic clustering analyses of the Linifolia
1041 clade. Codes for each taxa are: LIN for *I. linifolia*, KUZ for *I. kuzinskyana*, LIT-GAL
1042 for *I. littoralis* ssp. *gallaecica*, and LIT-LIT for *I. littoralis* ssp. *littoralis*. (a)
1043 Morphometric analysis. Principal Component Analaysis (PCA) of the 12
1044 morphologic chartacters: LS, length of the stem; LL, length of the leaf; WL, wide of
1045 the leaf; PINFL, pedicel of inflorescence; LCFR, length of fruit calix; DN, diameter of
1046 the nutlet; LN, length of the nutlet; LM, length of the margin nutlet; LT, length of
1047 the margin teeth; DH, density of trichomes; LH, length of trichome; BR: absence=0,
1048 presence=1 of flowering bracts). The contribution of each trait to the two principal
1049 components (PC1, PC2) as well as the positive or negative sense of the relationship
1050 is represented through the length of each arrow and the directionality of the
1051 arrow, respectively. The percentages of the variability explained by the two PCs
1052 are indicated close to the axes. (b) Niche overlap analysis. PCA of the six climatic
1053 variables obtained from Worldclim. BIO1, Annual Mean Temperature; BIO3,
1054 Isothermality; BIO4, Temperature Seasonality; BIO8, Mean Temperature of
1055 Wettest Quarter; BIO12, Annual Precipitation; BIO14, Precipitation of Driest
1056 Month. The strength of contribution of each variable to PCs and the sense of the
1057 relationship is represented, as explained above, through the length and the
1058 directionality of the arrows. The percentages of the variability explained by the
1059 two PCs are indicated close to the axes. Values of the variables are also
1060 represented for each PC separately following Fernández-López and Villa-Machío
1061 (2017). (c) Genetic clustering. Scatterplot from the DAPC analysis showing the
1062 genetic differentiation retained in one PC as the optimal number of PC for the

discriminant function. Compoplot for K=4 is also represented indicating membership probability to each genetic cluster for each taxa.

Fig. 4 Species distribution modelling obtained through the entropy algorithm, as implemented in Maxent for the four taxa of the Linifolia clade. Codes for each taxa are: LIN for *I. linifolia*, KUZ for *I. kuzinskyana*, and LIT for *I. littoralis* (LIT-GAL for *I. littoralis* ssp. *gallaecica*, LIT-LIT for *I. littoralis* ssp. *littoralis*). Analyses are based on WorldClim and projected to past conditions: last interglacial (LIG, c. 120–140 kya) and last glacial maximum (LGM, c. 21 kya). The presence/absence was estimated using the maximum training sensitivity plus specificity logistic threshold. The projection to the past is only presented according to the inferred times of divergence of each taxa.

TABLES

Table 1 Analysis of molecular variance (AMOVA) results for individuals from two contrasted groups: (1) mainland, *I. linifolia* and (2) coastal, *I. kuzinskyana* and *I. littoralis.*

Source of Varitation	d.f.	Sum of squares	Variance components		Percentage of variation
Among groups	1	1602.666	106.33509	Va	58.88
Among populations within groups	2	181.372	5.47321	Vb	3.03
Within populations	24	1651.033	68.79306	Vc	38.09
Total	27	3435.071	180.60135		

Note: Variance components are as in Excoffier *et al*. (1992). Fst = 0.61909; Fsc = 0.07370; Fct = 0.58878.

Table 2 Pairwise statistical test for comparison of ecological niche overlap between the different taxa within the Linifolia clade. The statistical significance is represented by p-values (Warren *et al.*, 2008). Asterisk (*) indicate significant values for *p*-values < 0.01.

Taxa comparison	Schoener's D	**Niche equivalency p-value**	**Niche similarity p-value**
I. linifolia vs *I. kuzinskyana*	0.026	0.0099*	0.94059
I. linifolia vs *I. littoralis* ssp. *gallaecica*	0.003	0.003	0.60396
I. linifolia vs *I. littoralis* ssp. *littoralis*	0.004	0.0099*	0.69307
I. kuzinskyana vs *I. littoralis* ssp. *gallaecica*	0.000	0.0099*	0.64356
I. kuzinskyana vs *I. littoralis* ssp. *littoralis*	0.000	0.0099*	0.71287
I. littoralis ssp. *gallaecica* vs *I. littoralis* ssp. *littoralis*	0.000	0.0099*	0.67327

Table 3 Examples of parapatric speciation in plants.

Genus	**Family**	**Event suggested**	**Reference**
Ephedra	Ephedraceae	Ecological divergence mediated by bioclimatic niche	Yin *et al.* (2016)
Fosterella	Bromeliaceae	Ecological divergence mediated by polyploidization and	Paule *et al.* (2017)
Helianthus	Asteraceae	Ecological divergence mediate by bioclimatic niche	Andrew & Rieseberg (2013)

Mimulus	Phrymaceae	Ecological divergence mediated by pollinator shift	Sobel & Streisfeld (2015), Stankowski *et al.* (2015)
Miscanthus	Poaceae	Ecological divergence mediated by bioclimatic niche	Huang *et al.* (2014)
Oryza	Poaceae	Ecological divergence mediated by bioclimatic niche	Zheng & Ge (2010)
Populus	Salicaceae	Ecological divergence mediated by bioclimatic niche	Zheng *et al.* (2017)
Roscoea	Zingiberaceae	Ecological divergence mediated by bioclimatic niche	Zhao *et al.* (2016)
Senecio	Asteraceae	Ecological divergence mediated by ecological niche	Richards and Ortiz-Barrientos (2016)
Stauracanthus	Fabaceae	Ecological divergence mediated by bioclimatic niche	Chozas *et al.* (2017)
Columnea	Gesneriaceae	Ecological divergence mediated by bioclimatic niche and pollinator shift	Schulte *et al.* (2015)

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 RAxML trees of *Iberodes* using three different similarity percentages (85%, 90%, 95%) and four levels of minium coverage (m4, m12, m20, m28). Bootstrap support are indicated when is lower than 100.

1097 **Fig. S2.** SVDquartets species tree of *Iberodes* using the eight main lineages
1098 obtained in the RAxML analyses.

1099 **Fig. S3** Compoplot showing individual assignment probability to different species-
1100 groups of the Linofolia clade of *Iberodes*, considering the two subspecies of *I.*
1101 *littoralis* as one group (k=3). Codes for each taxa are: LIN for *I. linifolia*, KUZ for *I.*
1102 *kuzinskyana*, LIT-GAL for *I. littoralis* ssp. *gallaecica*, and LIT-LIT for *I. littoralis* ssp.
1103 *littoralis*.

1104 **Fig. S4** Flow cytometry peaks for each of five samples used for each of the five
1105 species of *Iberodes* (including the two subspecies of *I. littoralis*. Peak of the
1106 standard *Solanum lycopersicum* is included. FL1 in the X axis indicates the signal
1107 intensity (530 nm / 28 nm; DAPI emission maximum=461). Y axis indicates the
1108 number of events found. Mean and median values of FL and percentage of
1109 variation coefficient is shown for each colored peak.

1110

1111 **Table S1** Data information and Genbank numbers of all individuals sampled for
1112 RADseq.

1113 **Table S2** Morphological characters analyzed. Values represent the mean of two
1114 individuals measured per herbarium sheet. LS: length of the stem, LL: length of the
1115 leaf, WL: wide of the leaf, PINFL: pedicel of inflorescence, LCFR: length of fruit
1116 calix, DN: diameter of the nutlet, LN: length of the nutlet, LM: length of the margin
1117 nutlet, LT: length of the margin teeth, DH: density of the hair, LH: length of
1118 trichome.

1119 **Table S3** Data points for environmental analysis

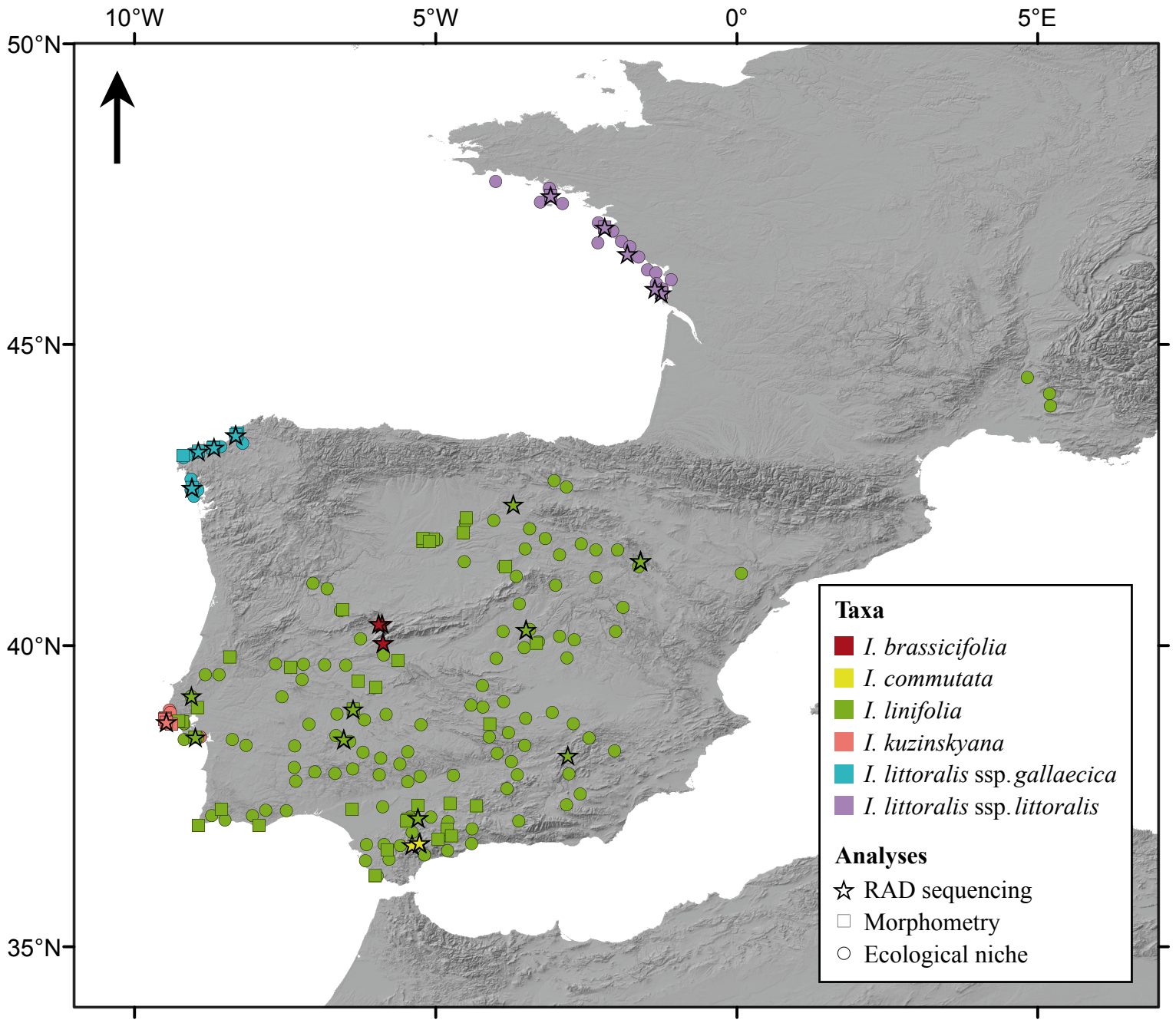
1120 **Table S4** Number of loci, SNPs retrieved for each of the 12 parameter
1121 combinations. Parameter abreviations indicate: c, clustering threshold; m, minimum
1122 taxon coverage to consider a locus.

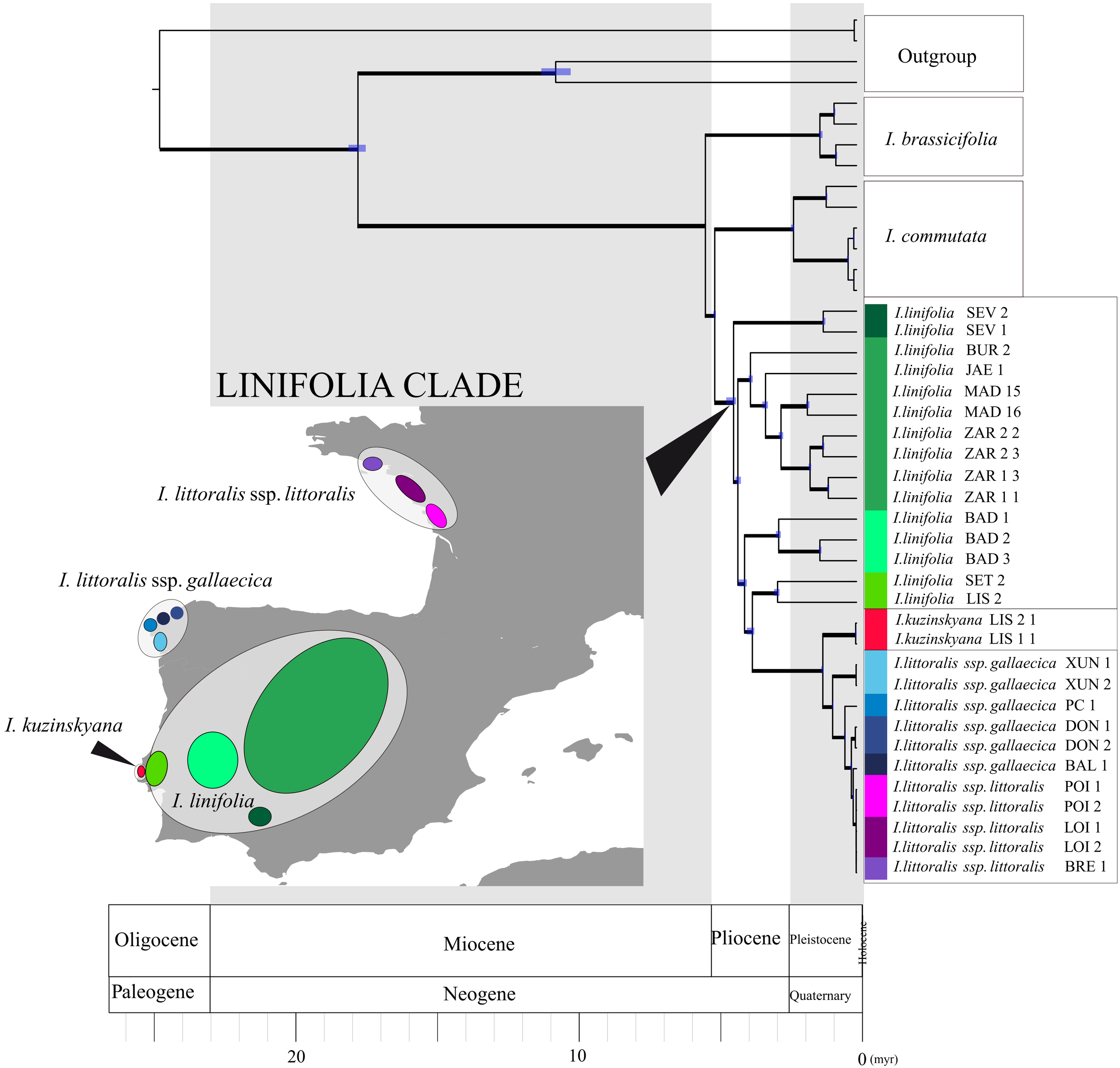
1123 **Table S5** Genome size estimation for the five species of *Iberodes* (including the two
1124 subspecies of *I. littoralis*.

1125 **Methods S1** Introgression analysis.

1126

1127

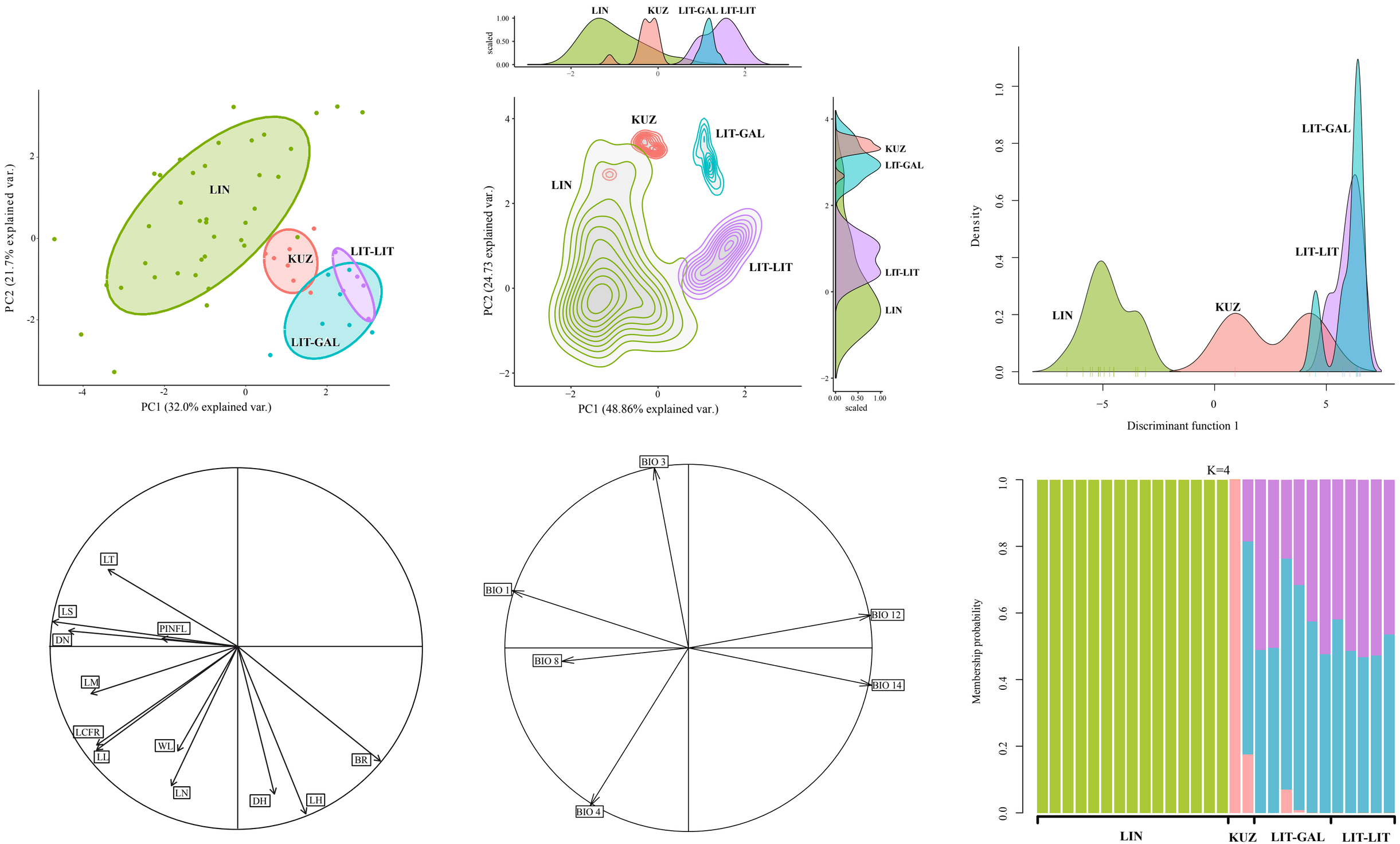

10°W
5°W
0°
5°E
50°N
45°N
40°N
35°N
Taxa
I. brassicifolia
I. commutata
I. linifolia
I. kuzinskyana
I. littoralis ssp. gallaecica
I. littoralis ssp. littoralis
Analyses
RAD sequencing
Morphometry
Ecological niche

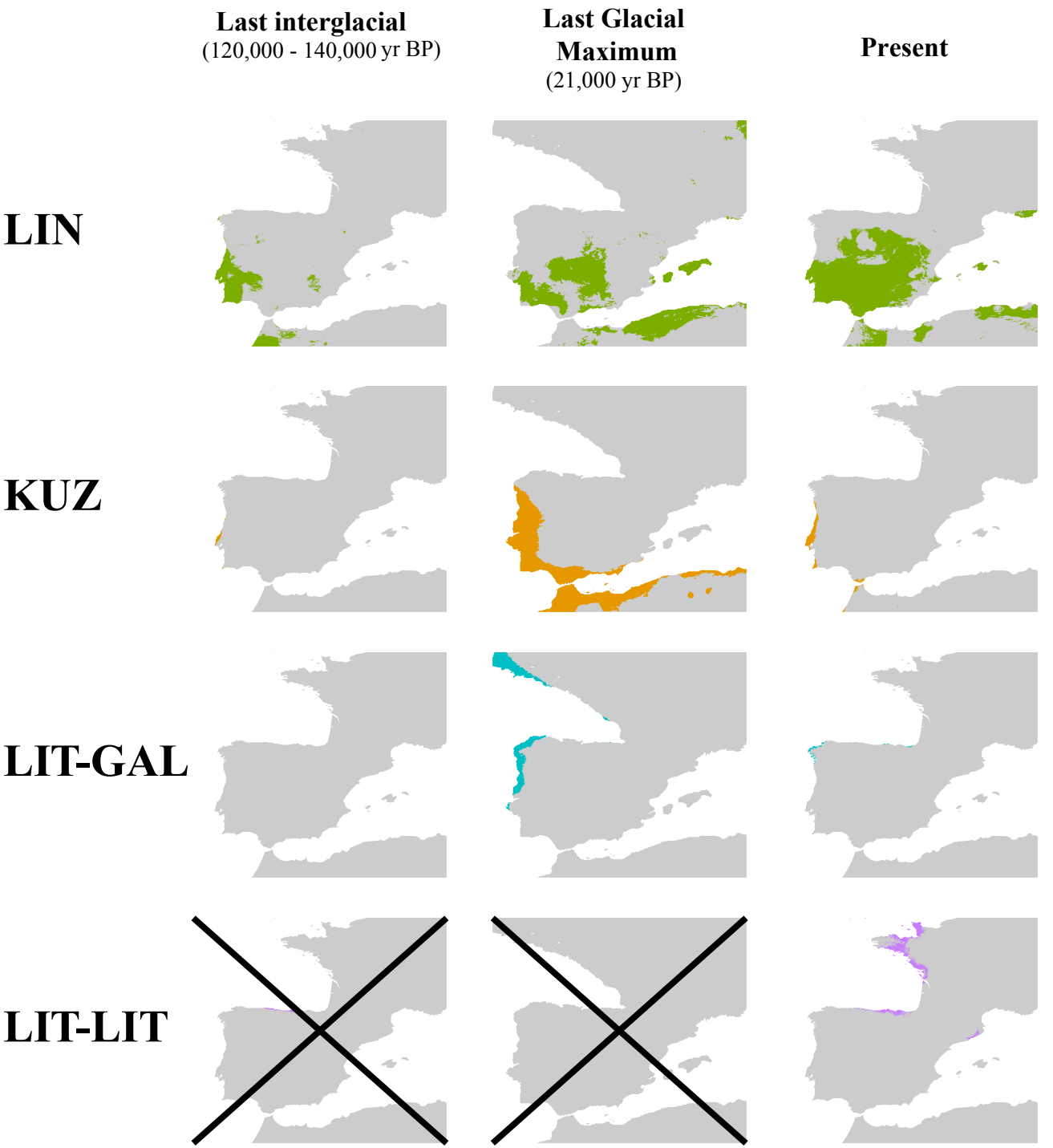

Outgroup
I. brassicifolia
I. commutata
LINIFOLIA CLADE
I. littoralis ssp. littoralis
I. littoralis ssp. gallaecica
I. kuzinskyana
I. linifolia
I.linifolia SEV 2
I.linifolia SEV 1
I.linifolia BUR 2
I.linifolia JAE 1
I.linifolia MAD 15
I.linifolia MAD 16
I.linifolia ZAR 2 2
I.linifolia ZAR 2 3
I.linifolia ZAR 1 3
I.linifolia ZAR 1 1
I.linifolia BAD 1
I.linifolia BAD 2
I.linifolia BAD 3
I.linifolia SET 2
I.linifolia LIS 2
I.kuzinskyana LIS 2 1
I.kuzinskyana LIS 1 1
I.littoralis ssp. gallaecica XUN 1
I.littoralis ssp. gallaecica XUN 2
I.littoralis ssp. gallaecica PC 1
I.littoralis ssp. gallaecica DON 1
I.littoralis ssp. gallaecica DON 2
I.littoralis ssp. gallaecica BAL 1
I.littoralis ssp. littoralis POI 1
I.littoralis ssp. littoralis POI 2
I.littoralis ssp. littoralis LOI 1
I.littoralis ssp. littoralis LOI 2
I.littoralis ssp. littoralis BRE 1
Oligocene
Miocene
Pliocene
Pleistocene
Holocene
Paleogene
Neogene
Quaternary
20
10
0 (myr)

(a) Morphometric analysis

(b) Niche overlap

(c) Genetic clustering




Last interglacial
(120,000 - 140,000 yr BP)
Last Glacial Maximum
(21,000 yr BP)
Present
LIN
KUZ
LIT-GAL
LIT-LIT