1 ORIGINAL ARTICLE

2

3 Evidence for niche similarities in the allopatric sister species *Lepus castroviejo* and
4 *L. corsicanus*

5 Pelayo Acevedo^{1,2,3*}, José Melo-Ferreira¹, Raimundo Real², Paulo C. Alves^{1,4,5}

6

7 ¹ CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBio
8 Laboratório Associado. Universidade do Porto, 4485-661 Vairão, Portugal.

9 ² Biogeography, Diversity and Conservation Research Team, University of Malaga,
10 29071 Malaga, Spain.

11 ³ SaBio-IREC (UCLM- CSIC-JCCM), 13071 Ciudad Real, Spain.

12 ⁴ Dpto. de Biologia, Faculdade de Ciências, Universidade do Porto, 4169-007 Porto,
13 Portugal.

14 ⁵ University of Montana, Wildlife Biology Program, College of Forestry and
15 Conservation, Missoula, MT 59812-0004 Montana, USA.

16

17 * Correspondence: Pelayo Acevedo, CIBIO, Centro de Investigação em Biodiversidade
18 e Recursos Genéticos, InBio Laboratório Associado. Universidade do Porto, 4485-661
19 Vairão, Portugal. E-mails: pacevedo@irec.csic.es, pelayo.acevedo@gmail.com

20

21 **Running title:** Niche similarities in allopatric sister species

22 **Word count:** 5750 (2 pages for tables and figures)

ABSTRACT

Aim *Lepus castroviejo* and *L. corsicanus* are sister species with allopatric distributions that share extensive phenotypic and genetic variation. Under the framework of niche conservatism, we assessed whether these species have similar ecological niches, which could provide insights on their mode of divergence, conservation, and taxonomic status.

Location The distribution range of *L. castroviejo*, in the Northern Iberian Peninsula, and of *L. corsicanus* in mainland Italy and Sicily.

Methods We developed spatially explicit ecological models to characterize the niches of the species, modelling the individual species separately and together. Individual models were transferred onto the territory of the sister species to explore their niche relationships. Predictions were assessed in terms of discrimination and calibration in a cross-assessment procedure.

Results The model trained for *L. castroviejo* was not able to predict the range of *L. corsicanus*. The model trained on *L. corsicanus* was able to discriminate the *L. castroviejo* distribution better than chance alone (AUC=0.814) but the reliability of the predictions was limited. However, the model trained on *L. corsicanus* in Italy's mainland (excluding range in Sicily), discriminated *L. castroviejo* presences/absences (AUC=0.788) and accurately predicted its probability of occurrence. Furthermore, a well-calibrated model, which was able to discriminate the species distributions (AUC=0.828 and AUC=0.956; for *L. castroviejo* and *L. corsicanus*, respectively), was obtained when species were considered together.

Main conclusions Our results suggest that *L. castroviejo* and *L. corsicanus* share extensive niche properties, which reinforces their possible conspecific status. The ecological niche of their ancestor may have resembled the present one in mainland Italy since this model was able to accurately predict the distribution range of both species.

48 Finally, ecological evidences suggest that niche conservatism may explain the
49 fragmentation of the distribution range of their ancestor, and thus may have been the
50 driver of the initial stages of divergence.

51

52 **Keywords:** allopatric speciation, Apennine Hare, Broom Hare, conservation
53 biogeography, ecological niche, lagomorphs, macroecology, species distribution
54 modelling.

55

For Peer Review

INTRODUCTION

The general tendency of species to maintain similar ecological requirements over evolutionary time scales is commonly known as ‘niche conservatism’ (Peterson *et al.*, 1999; see also Wiens & Graham, 2005). However, whether species retain their ecological niches throughout time is currently being highly debated (see e.g. Peterson, 2011). While some authors argued that speciation is not typically accompanied by ecological innovation, i.e., niches are conserved (e.g. Peterson *et al.*, 1999), others have found evidence of the contrary (e.g. Graham *et al.*, 2004; Schluter, 2009). The main discrepancies among studies may derive from methodological artefacts (Warren *et al.*, 2008) and, to a certain extent, from the evolutionary age of the community (Losos *et al.*, 2003). Aside from the debate of whether niche conservatism exists, which should probably be addressed in terms of identifying degrees of conservatism/divergence, the concept of niche conservatism is highly valuable for evolutionary biology and conservation biogeography; most importantly, it allows for the combination of inference methods from both disciplines.

Among other topics in evolutionary biology and conservation biogeography, the conceptual framework behind niche conservatism allows for the exploration of the ecological relationships within and between phylogenetically related taxa (see Wiens & Graham, 2005). Initial studies on this topic suggested that closely related species have more similar niches (e.g. Peterson *et al.*, 1999). However, some studies have examined the degree to which niche similarities among taxa is a direct function of the phylogenetic similarity, and have concluded that patterns of niche evolution beyond sister taxa are inconsistent (Losos *et al.*, 2003; Knouft *et al.*, 2006; Kalkvik *et al.*, 2012). While some closely related taxa have widely divergent niches, other distantly related taxa can maintain similar niches. Knouft *et al.* (2006) suggested that some

species – the anole *Anolis sagrei* group – in a given territory are segregated by environmental characteristics, and thus phylogenetically related species can be ecologically distant due to disruptive ecological adaptation. In such cases, interspecific interactions are stronger among closely related species, and niches tend to diverge from that of near relatives (Losos *et al.*, 2003). This is the expectation for parapatric species, closely related taxa with separate but contiguous distributions without physical barriers between them, which only co-occur in a narrow contact zone (Bull, 1991). Parapatric species usually display divergent ecological niches, that are forced by asymmetric competitive exclusion (Acevedo *et al.*, 2012a), otherwise the distributions could tend to coalesce (Bull & Possingham, 1995). Niche conservatism along phylogenetic trees cannot therefore be considered a rule in nature, but it depends on whether the divergence is caused by differential ecological adaptation. For example, niche conservatism in allopatric sister species would imply little ecological differentiation despite geographic separation, which can be tested empirically. Thus, integrating phylogenetic and macroecological studies within the niche conservatism conceptual framework allows the testing of different evolutionary hypothesis regarding the onset of species divergence (e.g. Kozak & Wiens, 2006; Kalkvik *et al.*, 2012).

The genus *Lepus* (hares) is a promising model for studying questions related to niche conservatism. Hares are distributed worldwide and include over 30 presently recognized species that are thought to have resulted from recent rapid radiation, presumably from North America (Matthee *et al.*, 2004; Melo-Ferreira *et al.*, 2012). Many taxa have parapatric distributions (e.g. Acevedo *et al.*, 2012a), which occupy distinct ecological niches even if they are phylogenetically closely related (Melo-Ferreira *et al.*, 2012). On the contrary, *L. castroviejoi* and *L. corsicanus* display high morphological (Palacios, 1996) and genetic similarities (Alves *et al.*, 2008a; Melo-Ferreira *et al.*, 2012), but

1
2
3 106 currently have allopatric ranges (Figure 1). While *L. castroviejo* is endemic to the
4
5 107 Cantabrian Mountains in the Iberian Peninsula (Palomo *et al.*, 2007), *L. corsicanus* is
6
7 108 endemic to the Italian Apennines and Sicily (Angelici & Luiselli, 2001; Angelici *et al.*,
8
9 109 2008), and it was rediscovered in Corsica (Scalera & Angelici, 2003; see also Pietri *et*
10
11 110 *al.*, 2011). Furthermore, Melo-Ferreira *et al.* (2012) suggested that ancient hybridization
12
13 111 between the ancestors of *L. castroviejo*-*L. corsicanus* and *L. timidus*, an arctic or boreal
14
15 112 taxa currently distributed in the northern Palaearctic, must have occurred. These events
16
17 113 led to the complete replacement of the original mitochondrial DNA (mtDNA) variants
18
19 114 of *L. castroviejo*-*L. corsicanus* by that of *L. timidus*, and explain the close genetic
20
21 115 similarity between the mtDNA haplotypes of *L. timidus*, *L. corsicanus* and *L.*
22
23 116 *castroviejo*, which have been previously described (Alves *et al.*, 2003; Wu *et al.*,
24
25 117 2005). The data suggest that *L. corsicanus* and *L. castroviejo* are sister species which
26
27 118 have shared a very recent common ancestor during the Late Pleistocene.
28
29 119 In this work, we assessed whether *L. castroviejo* and *L. corsicanus* conserve similar
30
31 120 ecological niches, which could provide key insights on the mode of divergence,
32
33 121 ecological requirements, conservation, and taxonomic status of these species. The work
34
35 122 specifically tested whether the niche properties of one of these species is capable of
36
37 123 predicting the distribution range of the other, and whether the species distribution
38
39 124 ranges of both could be accurately predicted by a unique ecological model. Analyses of
40
41 125 the niches were fully integrated into an evolutionary context.
42
43
44
45
46
47
48

127 **METHODS**

128 **Study areas and species distribution**

129 As the extent of the geographical background has substantial effects on the outputs of
130 species distribution modelling (see Anderson & Raza, 2010; Barve *et al.*, 2011;

1
2
3 131 Acevedo *et al.*, 2012b), first we delimited the arena of each species model. This was
4
5 132 done by modelling the species distribution where the third-degree polynomial of the
6
7 133 spatial coordinates were considered as predictors; the geographical background was
8
9 134 delimited by localities which had a predicted favorability higher than the minimum
10
11 135 value assigned to a presence (for further details see Acevedo *et al.*, 2012b; Figure 1).
12
13 136 Species distribution data were obtained on UTM 10 km × 10 km squares (our territorial
14
15 137 unit for modelling purposes) from Palomo *et al.* (2007) for *L. castroviejo* (n=69
16
17 138 presences), and for *L. corsicanus* the information was provided by Boitani *et al.* (2002),
18
19 139 and was adapted to these territorial units (n=367; of them n=298 are concentrated in
20
21 140 Sicily). Although *L. corsicanus* is also present in Corsica, its presence is considered
22
23 141 testimonial (Angelici *et al.*, 2008); thus information for Corsica was not considered for
24
25 142 modelling.
26
27
28
29
30
31

32 144 **Ecological variables and spatially explicit modelling**

33
34 145 Previous studies reported some of the ecological requirements of the studied species. *L.*
35
36 146 *castroviejo* has been described as typically occupying shrub lands in which mosaics
37
38 147 with pastures and broadleaved woods are frequent (Acevedo *et al.*, 2007). *L. corsicanus*
39
40 148 inhabits a broader spectrum of habitats in Mediterranean environments ranging from
41
42 149 coastal dune to high mountain habitats, as well as grassland, pastures and/or cultivated
43
44 150 areas (Angelici & Luiselli, 2007; Angelici *et al.*, 2008). Based on previous knowledge
45
46 151 and the availability of environmental predictors for the scale required in this study, a set
47
48 152 of 19 environmental predictors were preselected to account for the niches of these
49
50 153 species (see Table 1). The predictors were related to two factors: topoclimate (8
51
52 154 predictors) and land uses (11 predictors). Bioclimatic variables and altitude (~1000 m
53
54
55
56 155 spatial resolution) were obtained from the Worldclim project database (see Hijmans *et*
57
58
59
60

1
2
3 156 *al.*, 2005). Land use data was obtained from Global Land Cover 2005, which accurately
4
5 157 assigns the land uses for each pixel of ~300 m (see Bicheron *et al.*, 2008).
6
7 158 It is recommended, when conclusions about niche conservatism want to be extracted, to
8
9 159 control overfitting by managing the number of predictors (dimensions) in the models
10
11 160 (Peterson, 2011). In general, the probability of detecting a conserved niche increases as
12
13 161 the number of dimensions in the model decreases (Peterson & Nakazawa, 2008).
14
15 162 Following the recommendations provided by Peterson (2011), different strategies were
16
17 163 explored to deal with problems deriving from dimensionality (for details Appendix S1
18
19 164 of the Supplementary Material). Given that similar interpretations for niche
20
21 165 conservatism could be obtained from all the sets of predictors, the results of the most
22
23 166 parsimonious models are presented; they were obtained when both topoclimatic and
24
25 167 land use predictors were considered together.
26
27 168 Using an inductive approach, the macroecological requirements of the studied species –
28
29 169 “ecological niche” – based on the locations in which they occurred were determined.
30
31 170 Two statistical stages were carried out. To control the increase in type I errors as the
32
33 171 number of predictors increased, in a first stage the false discovery rate (FDR) using the
34
35 172 procedure proposed by Benjamini & Hochberg (1995) was evaluated. Only predictors
36
37 173 that were significantly related to the species distribution under a FDR of $q < 0.05$ were
38
39 174 selected. Predictors selected after controlling the FDR were considered in a multiple
40
41 175 logistic regression analysis in a second stage, and the final models were obtained using
42
43 176 a forwards-backwards stepwise procedure based on the Akaike Information Criterion
44
45 177 (AIC). Statistical analyses were carried out in R 2.15.2 (R Core Team 2012).
46
47 178 Four different models were developed: a) an overall model for both sister species, b)
48
49 179 one for *L. castroviejoi*, c) one for *L. corsicanus*, and d) one for *L. corsicanus*, which
50
51 180 excluded its distribution range in Sicily. To assess the niche similarities between
52
53
54
55
56
57
58
59
60

181 species, a model for one species, which was trained on its specific study area, was
182 projected onto the territory of the other species. To project the model developed for one
183 species in another territory (spatial transferability), the similarity in environmental
184 gradients and the correlation structure among independent variables between study
185 areas should be maintained (e.g. Jiménez-Valverde *et al.*, 2011). We successfully
186 checked for model transferability requirements between our study areas (for details see
187 Appendix S2).

188
189 **Niche conservatism: a cross-assessment based on discrimination and calibration**

190 Warren *et al.* (2008) pointed out the relevance of the null hypotheses in the study of
191 niche conservatism, since both niche similarity and niche identity can be measured, but
192 the results can yield opposite conclusions regarding niche conservatism. Niche
193 similarity tests whether one species' niche model predicts the occurrences of a second
194 species better than expected by chance (e.g. Peterson *et al.*, 1999). The area under the
195 receiver operating characteristic (ROC) curve (AUC), the most popular discrimination
196 capacity measure in species distribution modelling, can be interpreted as the probability
197 that a presence chosen at random will be assigned a higher probability than an absence
198 chosen at random (for details see Krzanowski & Hand, 2009). Therefore, the AUC can
199 be a suitable parameter to quantify niche similarity when the predictions of the model
200 for one species are assessed on the territory of a second species such as it is used in this
201 study. It should be noted that the discrimination performance based only on Sicily
202 cannot be assessed since *L. corsicanus* is present in all territorial units of this island.
203 If we view niche conservatism as a continuum gradient (ranging from niches that are
204 identical to niches that are more similar than random), the AUC can provide
205 information about one extreme. For instance, values significantly higher than 0.5

1
2
3 206 indicate niches that are more similar than expected by chance. Yet, the AUC does not
4
5 207 provide information on niche identity, since a model can adequately discriminate the
6
7 208 distribution of a species even when it is not highly informative of its distribution. This is
8
9 209 because the AUC is affected by the distribution of the territory units along the modeled
10
11 210 environmental gradient (Lobo *et al.*, 2008; Jiménez-Valverde *et al.*, 2013). Thus
12
13 211 calibration, i.e., the degree to which the observed proportion of presences equate to the
14
15 212 model estimated probabilities (Pearce & Ferrier, 2000), is useful for informing on niche
16
17 213 identity; if the niches of two related species are identical, the niche model for one of
18
19 214 them is able to accurately predict the probability of occurrence for the second species.
20
21 215 Calibration plots were developed to visually assess the relatedness of the niches. Graphs
22
23 216 were constructed with 'ggplot2' R package (Wickham, 2009) by plotting the proportion
24
25 217 of occupied evaluation sites of a given species against the predicted probability of
26
27 218 presence of the other species (for the ten equal-size probability intervals); points located
28
29 219 along the 45° line indicate identical niches. Hosmer–Lemeshow goodness-of-fit statistic
30
31 220 (H-L; Lemeshow & Hosmer, 1982) was used for the statistical assessment of the
32
33 221 calibration plots. H-L tests assessed whether the observed values matched the expected
34
35 222 ones in the calibration plots. Good calibration is obtained when the expected and
36
37 223 observed presence rates are not significantly different ($p > 0.05$). Statistical analyses
38
39 224 were carried out in R 2.15.2 (R Core Team 2012).
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

226 RESULTS

227 Spatially explicit ecological models

228 After controlling the FDR, not all predictors were considered for inclusion in the final
229 models: 14 were considered for an overall model for both species (BIO0, BIO2, BIO4,
230 BIO9, BIO15, BIO19, T11, T20, T30, T70, T110, T120, T130 and T150), 7 for the *L.*

231 *castroviejo* model (BIO2, BIO9, BIO15, T14, T30, T120 and T150), 15 for the *L.*
232 *corsicanus* model (BIO0, BIO2, BIO4, BIO9, BIO15, BIO19, T11, T14, T20, T30, T70,
233 T110, T120, T130 and T150), and 9 for the *L. corsicanus* model excluding Sicily
234 (BIO0, BIO2, BIO15, BIO19, T14, T20, T70, T100 and T210).
235 The forward-backward stepwise procedure allowed one to obtain the final logistic
236 regression models for all species/territories (Table 2, see also Figure 2). In the training
237 territories, models achieved good predictive performance both in terms of
238 discrimination (AUC=0.948, AUC=0.934, AUC=0.954, and AUC=0.813; for the
239 overall model, *L. castroviejo* model, *L. corsicanus* model, and for the model of *L.*
240 *corsicanus* in mainland Italy, respectively) and calibration (see Figure 3).

242 **Niche conservatism**

243 A unique ecological model for both sister species was fitted (Table 2), and it was able to
244 accurately predict their distribution ranges (Figures 2a). When the discrimination
245 performance of this model was assessed, a high discrimination capacity was obtained on
246 *L. castroviejo* territory (AUC=0.828) and on that of *L. corsicanus* (AUC=0.956).
247 Calibration plots indicated that the model predicted similarly the probability of presence
248 for the two species, *L. castroviejo* and *L. corsicanus* (Figure 3a).
249 When the model trained on *L. castroviejo* was transferred to the territory of *L.*
250 *corsicanus*, it did not discriminate better than chance the distribution of the latter
251 species (AUC=0.232). On the contrary, the model trained on *L. corsicanus* data was
252 able to predict the territory of *L. castroviejo* (AUC=0.814). Furthermore, the model
253 trained on *L. corsicanus* mainland distribution also discriminated better than chance the
254 *L. castroviejo* distribution (AUC=0.789), and *L. corsicanus* overall distribution (AUC=
255 0.853). Calibration plots for the model trained on *L. castroviejo*, partially supported the

1
2
3 256 interpretations obtained by assessing the AUC (Figure 3b). However, calibration plots
4
5 257 showed that the predictions of the model for *L. corsicanus* did not fit accurately the
6
7 258 observed frequency of *L. castroviejo* (Figure 3c), but it did fit the predictions of the
8
9 259 model for *L. corsicanus* in mainland Italy (Figure 3d).
10
11
12 260

13 14 261 **DISCUSSION**

15
16 262 Multidisciplinary approaches are fundamental for understanding the past, present and
17
18 263 future evolutionary and ecological features of species, and to gather relevant
19
20 264 information that can be used to implement adequate conservation measures. Analyses
21
22 265 using morphological and genetic data have previously suggested that two allopatric
23
24 266 species of hares, *L. castroviejo* and *L. corsicanus*, are recent descendants of a common
25
26 267 species (Palacios, 1996; Alves *et al.*, 2008a; Melo-Ferreira *et al.*, 2012). Here, these
27
28 268 taxa were studied from a novel perspective, and extensive similarities were also found
29
30 269 in their ecological traits. These results have important implications on the understanding
31
32 270 of the evolutionary history of these species and their ecological requirements.
33
34 271 Moreover, it contributes to the discussion of the relationship between niche
35
36 272 conservatism and evolutionary history.
37
38
39
40
41 273

42 43 274 **On the methodological approach**

44
45 275 Various metrics have been proposed to quantify potential niche overlap, and some were
46
47 276 comparatively assessed in the context of niche conservatism (e.g. Warren *et al.*, 2008;
48
49 277 Rödder & Engler, 2011). Each showed advantages and drawbacks on the desirable
50
51 278 characteristics for a metric of potential niche overlap (see Rödder & Engler, 2011).
52
53 279 Based on the concept behind the pioneer study on niche conservatism (Peterson *et al.*,
54
55 280 1999), a simple and more in-depth exploration of niche similarities between two sister
56
57
58
59
60

species was used here. The approach was based on a cross-assessment focused on the two characteristic of the model's predictive performance, i.e., discrimination and calibration. Assessing the discrimination with the AUC, and according to the model based on the sister species, the probability for a presence chosen at random for one species to be assigned a higher probability than an absence chosen at random was determined. In other words, niche similarity was assessed ($AUC > 0.5$). To determine if two niches are (or are not) identical, the interspecific niche relationships should also be assessed in terms of calibration. In the cross-assessment context, a perfectly calibrated model shows that two niches are identical. The calibration plot informs about the uncertainty of the predictions, and how they are distributed along the gradient of probabilities predicted by the model, which is also needed for a proper understanding of the AUC values (Jiménez-Valverde *et al.*, 2013). Thus, in a cross-assessment context, discrimination and calibration allow for a more in-depth exploration of niche similarities; hence, additional conclusions about niche relationships can be extracted. For instance, a situation in which the model overpredicts species occurrence under a given threshold of probability and underpredicts over that threshold can be identified (see Figure 3a in the *L. corsicanus*'s territory). That situation could be informing about the presence of an unconsidered factor (e.g., a competitor, a resource, human interference, etc.), which is preventing the species from occupying some favorable localities.

Ecological similarity of taxa: implications for taxonomy, evolutionary history and conservation

The spatial resolution of the analysis was suggested as a factor affecting the conclusions about niche conservatism (e.g., Peterson & Nyári, 2007). Therefore, the grain size of the analyses is important to properly understand the interpretations extracted from the

1
2
3 307 results. Results obtained in this study – from an ecological perspective and at 10km ×
4
5 308 10km spatial resolution – support the close evolutionary relationship between *L.*
6
7 309 *castroviejo* and *L. corsicanus*. This similarity adds to the morphological analyses
8
9 310 performed by Palacios (1996), and to the molecular inferences made by Alves *et al.*
10
11 311 (2008a) and Melo-Ferreira *et al.* (2012), which led these authors to suggest that these
12
13 312 taxa could eventually be considered conspecific. If parapatric groups of species (as
14
15 313 *Lepus* spp. in Europe) are naturally conceived to be ecologically divergent (Bull &
16
17 314 Possingham, 1995), then niche similarities between related species such as *L.*
18
19 315 *castroviejo* and *L. corsicanus* (even in allopatry), can be used as arguments to discuss a
20
21 316 possible conspecific status. Even if species concepts vary and the classification of *L.*
22
23 317 *castroviejo* and *L. corsicanus* as conspecific will depend on the applied criterion, the
24
25 318 similarities of these taxa is remarkable and shows how morphology, genes and ecology
26
27 319 may be conserved in the initial stages of divergence.
28
29 320 According to genetic evidences, *L. castroviejo* and *L. corsicanus* shared a common
30
31 321 ancestor possibly at least until the Upper Pleistocene (Alves *et al.*, 2008a; Melo-Ferreira
32
33 322 *et al.*, 2012). Similarly, as it has been described for other taxa such as the chamois
34
35 323 (*Rupicapra* spp., Masini & Lovari, 1988), it could be hypothesized that their common
36
37 324 ancestor occupied a larger range in Europe during the Pleistocene and the subsequent
38
39 325 climatic changes split it into two allopatric refugia in the two Southern European
40
41 326 peninsulas, Iberia and Italy (Angelici & Luiselli, 2007). Our results suggest that this
42
43 327 range of separation was not driven by disruptive adaptation to different ecological
44
45 328 settings, but must have resulted from the fragmentation of favorable habitat.
46
47 329 Apparently, *L. castroviejo* and *L. corsicanus* did not expand from these refugia, which
48
49 330 may have been due to i) competition with other hare species, ii) niche conservatism, or
50
51 331 most probably iii) by both factors to certain extent (see below). *L. castroviejo* could
52
53
54
55
56
57
58
59
60

1
2
3 332 have initially faced competition by *L. granatensis* and likely by *L. timidus* when this
4
5 333 species was present in northern Iberian Peninsula in Late Pleistocene (Altuna, 1970).
6
7 334 The presence of *L. timidus* mtDNA type in *L. castroviejo* due to hybridization (Alves *et*
8
9 335 *al.*, 2008b; Melo-Ferreira *et al.*, 2012) is congruent with the hypothesis of contact
10
11 336 between these species. Also, when the niche expressed by *L. corsicanus* was projected
12
13 337 onto Iberia, in general, a high environmental potential for *L. castroviejo* in more
14
15 338 southern areas of the Iberian study area was obtained (Figure 2). These areas are not
16
17 339 currently occupied by *L. castroviejo* but are highly favourable for *L. granatensis*
18
19 340 (Acevedo *et al.*, 2012c). This suggests that competition with *L. granatensis* may have
20
21 341 been an important factor impeding colonization of these southern territories by *L.*
22
23 342 *castroviejo*. Similarly, in the Italian Peninsula, the northwards expansion of *L.*
24
25 343 *corsicanus* could have been prevented by competition with *L. timidus* and *L. europaeus*
26
27 344 since the past range of *L. corsicanus* reached more northern latitudes (Angelici &
28
29 345 Luiselli, 2007; Angelici *et al.*, 2008).
30
31
32 346 Niche conservatism could have also limited the capacity of these species to colonize
33
34 347 new habitats since they are currently occupying “ecological islands” (e.g. Figure 1),
35
36 348 separated by intervening landscapes unsuitable for these species, i.e., by an ecological
37
38 349 barrier. In this situation, niche conservatism can limit adaptation to suboptimal
39
40 350 ecological conditions (see Wiens & Graham, 2005); hence, it restricts the expansion
41
42 351 ability of the species. Following this line of thought, niche conservatism may have
43
44 352 further implications on the evolution of these sister species. If the species cannot adapt
45
46 353 to suboptimal environmental conditions, then the barrier will continue to prevent gene
47
48 354 flow between these species, and they will inevitably continue to diverge (Kozak &
49
50 355 Wiens, 2006), according to the allopatric speciation model. At this point, it should be
51
52
53
54
55
56
57
58
59
60

1
2
3 356 highlighted that these species offer a fascinating opportunity for a natural experiment in
4
5 357 which, for instance, the genetic bases of recent local adaptations can be explored.
6
7 358 Finally, but not least, our results suggest that the niche expressed by *L. corsicanus* in
8
9 359 mainland Italy retained ecological traits able to accurately predict the global distribution
10
11 360 of these sister species. Interestingly, some genetic differences between hares from
12
13 361 mainland Italy and Sicily were found using mtDNA (Pierpaoli *et al.*, 1999; Alves *et al.*,
14
15 362 2008a; Pietri *et al.*, 2011), which may have originated from a drift during a founding
16
17 363 event of the Sicilian *L. corsicanus* population. The ecological divergence can be related
18
19 364 to the fact that *L. corsicanus* is the only hare species in Sicily despite the intensive
20
21 365 release of *L. europaeus* during the last decades (Angelici *et al.*, 2008). Therefore, the
22
23 366 populations in Sicily are evolving in absence of potential competitors, and thus
24
25 367 displaying something that could closely resemble the part of its fundamental niche that
26
27 368 is present in Sicily. On the contrary, continental populations evolved in contact with
28
29 369 competitors species, and the pattern of distribution accounted by the models is closer to
30
31 370 the species realized niche (e.g. Soberón & Peterson, 2005; Angelici & Luiselli, 2007).
32
33 371 Thus, different parts of the niche could have been occupied in mainland Italy and Sicily
34
35 372 which could be related to rather different expressions of the species ecological niche
36
37 373 (see Randin *et al.*, 2006). Moreover, the fact that the model trained on *L. corsicanus*
38
39 374 mainland population was able to predict the range of *L. castroviejoi* but not vice-versa,
40
41 375 suggest that *L. corsicanus* populations in mainland Italy have retained the ecological
42
43 376 traits of their ancestor, and thus, these Italian populations have an impressive value for
44
45 377 ecological and evolutionary studies. Also, due to the characteristics described above, *L.*
46
47 378 *castroviejoi* and *L. corsicanus* are significant evolutionary units that should be
48
49 379 preserved.
50
51
52
53
54
55
56
57
58
59
60

380 Similarly to the seminal studies of Peterson *et al.* (1999) and Wiens & Graham (2005),
381 among others, our study shows that macroecology can be useful to explore diverse
382 topics in evolution. Therefore, the integration of macroecology and phylogeography
383 would contribute decisively in the strengthening of the geographic components of
384 evolutionary biology, which is often omitted (Kidd & Ritchie, 2006).

385

386 **ACKNOWLEDGMENTS**

387 We are grateful to Miles Silman and three anonymous referees for their useful
388 comments and suggestions on previous versions of the manuscript. Lucía D. Maltez
389 kindly reviewed the English. P. A. had a Jose Castillejo fellowship (2010-2011) in
390 Portugal awarded by the Ministerio de Ciencia e Innovación to conduct this study, and
391 was also partially supported by the CGL2009-11316/BOS Spanish Government and
392 FEDER project. Currently, P. A. and J. M.-F. hold post-doctoral grants from Fundação
393 para a Ciência e a Tecnologia (FCT) funded by Programa Operacional Potencial
394 Humano (POPH) – Quadro de Referência Estratégico Nacional (QREN) from the
395 European Social Fund and by the Portuguese Ministério da Educação e Ciência
396 (SFRH/BPD/90320/2012 and SFRH/ BPD/43264/2008, respectively). Financial support
397 was partially obtained from research project grants PTDC/BIA-EVF/111931/2009 and
398 PTDC/BIA-EVF/115069/2009 funded by FEDER through the COMPETE program and
399 Portuguese national funds through the FCT.

400

401 **REFERENCES**

402 Acevedo, P., Alzaga, V., Cassinello, J. & Gortázar, C. (2007) Habitat suitability
403 modelling reveals a strong niche overlap between two poorly known species, the

- 1
2
3 404 broom hare and the Pyrenean grey partridge, in the north of Spain. *Acta*
4
5 405 *Oecologica*, **31**, 174-184.
6
7 406 Acevedo, P., Jiménez-Valverde, A., Lobo, J.M. & Real, R. (2012b) Delimiting the
8
9 407 geographical background in species distribution modelling. *Journal of*
10
11 408 *Biogeography*, **39**, 1383-1390.
12
13 409 Acevedo, P., Jiménez-Valverde, A., Melo-Ferreira, J., Real, R. & Alves, P.C. (2012a)
14
15 410 Parapatric species and the implications for climate change studies: a case study
16
17 411 on hares in Europe. *Global Change Biology*, **18**, 1509-1519.
18
19 412 Acevedo, P., Melo-Ferreira, J., Real, R. & Alves, P.C. (2012c) Past, Present and Future
20
21 413 Distributions of an Iberian Endemic, *Lepus granatensis*: Ecological and
22
23 414 Evolutionary Clues from Species Distribution Models. *PLoS ONE* 7(12):
24
25 415 e51529.
26
27 416 Altuna, J. (1970) Hallazgo de una liebre ártica (*Lepus timidus*) en el yacimiento
28
29 417 prehistorico de Urtiga (Guipuzcoa). *Munibe*, **22**, 165-168.
30
31 418 Alves, P.C., Ferrand, N., Suchentrunk, F. & Harris, D.J. (2003) Ancient introgression of
32
33 419 *Lepus timidus* mtDNA into *L. granatensis* and *L. europaeus* in the Iberian
34
35 420 Peninsula. *Molecular Phylogenetics and Evolution*, **27**, 70-80.
36
37 421 Alves, P.C., Melo-Ferreira, J., Branco, M., Suchentrunk, F., Ferrand, N. & Harris, D.J.
38
39 422 (2008a) Evidence for genetic similarity of two allopatric European hares (*Lepus*
40
41 423 *corsicanus* and *L. castroviejoii*) inferred from nuclear DNA sequences.
42
43 424 *Molecular Phylogenetics and Evolution*, **46**, 1191-1197.
44
45 425 Alves, P.C., Melo-Ferreira, J., Freitas, H. & Boursot, P. (2008b) The ubiquitous
46
47 426 mountain hare mitochondria: multiple introgressive hybridization in hares,
48
49 427 genus *Lepus*. *Philosophical Transactions of the Royal Society, Biological*
50
51 428 *Sciences*, **363**, 2831-2839.
52
53
54
55
56
57
58
59
60

Anderson, R.P. & Raza, A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, **37**, 1378–1393.

Angelici, F.M. & Luiselli L. (2001) Distribution and status of the critically endangered Apennine hare *Lepus corsicanus* De Winton 1898, in continental Italy and Sicily. *Oryx*, **35**, 245-249.

Angelici, F.M. & Luiselli, L.(2007) Range, dynamic biogeography and ecological interactions of two species: *Lepus corsicanus* and *Lepus europaeus* in Italy. *Wildlife Biology*, **13**, 251-257.

Angelici, F.M., Randi, E., Riga, F. & Trocchi, V. (2008) *Lepus corsicanus*. In: IUCN 2011. IUCN Red List of Threatened Species. Version 2011.2. <www.iucnredlist.org>. Downloaded on 05 June 2012.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810-1819.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, **57**, 289-300.

Bicheron, P., Defourny, P., Brockmann, C. *et al.* (2008) *GlobCover 2005 – Products description and validation report, Version 2.1, 2008*. Available on the ESA IONIA website (<http://ionial.esrin.esa.int/>).

Boitani, L., Corsi, F., Falcucci, A., Marzetti, I., Masi, M., Montemaggiori, A., Ottaviani, D., Reggiani, G. & Rondinini, C. (2002) Rete ecologica nazionale: un

- 454 approccio alla conservazione dei vertebrati italiani. Relazione Finale. DCN –
455 Ministero Ambiente, Dep. BAU - University of Rome “La Sapienza”.
- 456 Bull, C.M. & Possingham, H. (1995) A model to explain ecological parapatry. *The*
457 *American Naturalist*, **145**, 935-947.
- 458 Bull, C.M. (1991) Ecology of parapatric distributions. *Annual Review of Ecology and*
459 *Systematics*, **22**, 19-36.
- 460 Graham, C.H., Ron, S.R., Santos, J.C., Schneider, C.J. & Moritz, C. (2004) Integrating
461 phylogenetics and environmental niche models to explore speciation
462 mechanisms in dendrobatid frogs. *Evolution*, **58**, 1781-1793.
- 463 Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high
464 resolution interpolated climate surfaces for global land areas. *International*
465 *Journal of Climatology*, **25**, 1965-1978.
- 466 Jiménez-Valverde, A., Acevedo, P., Barbosa, A.M., Lobo, J.M., Real, R. (2013)
467 Discrimination capacity in species distribution models depends on the
468 representativeness of the environmental domain. *Global Ecology and*
469 *Biogeography*, **22**, 508-516.
- 470 Jiménez-Valverde, A., Decae, A.E. & Arnedo, M.A. (2011) Environmental suitability of
471 new reported localities of the funnelweb spider *Macrothele calpeiana*: an
472 assessment using potential distribution modelling with presence-only
473 techniques. *Journal of Biogeography*, **38**, 1213-1223.
- 474 Jovani, R. & Tella, J.L. (2006) Parasite prevalence and sample size: misconceptions and
475 solutions. *Trends in Parasitology*, **22**, 214-218.
- 476 Kalkvik, H.M., Stout, I.J., Doonan, T.J. & Parkinson, C.L. (2012) Investigating niche
477 and lineage diversification in widely distributed taxa: phylogeography and

ecological niche modeling of the *Peromyscus maniculatus* species group.
Ecography, **35**, 54-64.

Kidd, D.M. & Ritchie, M.G. (2006) Phylogeographic information systems: putting the
geography into phylogeography. *Journal of Biogeography*, **33**, 1851-1865.

Knouft, J.H., Losos, J.B., Glor, R.E. & Kolbe, J.J. (2006) Phylogenetic analysis of the
evolution of the niche in lizards of the *Anolis sagrei* group. *Ecology*, **87**, S29-
S38.

Kozak, K.H. & Wiens, J.J. (2006) Does niche conservatism promote speciation? A case
study in North American salamanders. *Evolution*, **60**, 2604-2621.

Krzanowski, W.J. & Hand, D.J. (2009) ROC curves for continuous data. CRC/Chapman
and Hall, New York.

Lemeshow, S. & Hosmer, D.W. (1982) A review of goodness of fit statistics for use in
the development of logistic regression models. *American Journal of
Epidemiology*, **115**, 92-106.

Lobo, J., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the
performance of predictive distribution models. *Global Ecology and
Biogeography*, **17**, 145-151.

Losos, J.B., Leal, M., Glor, R.E., de Queiroz, K., Hertz, P.E., Rodriguez Schettino, L.,
Chamizo Lara, A., Jackman, T.D. & Larson, A. (2003) Niche lability in the
evolution of a Caribbean lizard community. *Nature*, **424**, 542-545.

Masini, F. & Lovari, S. (1988) Systematics, phylogenetic and dispersal of the chamois
(*Rupicapra* spp.). *Quaternary Research*, **30**, 339-349.

Matthee, C.A, van Vuuren, B.J, Bell, D. & Robinson, T.J (2004) A molecular
supermatrix of the rabbits and hares (Leporidae) allows for the identification of

- 502 five intercontinental exchanges during the Miocene. *Systematic Biology*, **53**,
503 433-447.
- 504 Melo-Ferreira, J., Boursot, P., Carneiro, M., Esteves, P.J., Farelo, L. & Alves, P.C.
505 (2012) Recurrent Introgression of Mitochondrial DNA Among Hares (*Lepus*
506 spp.) revealed by Species-tree Inference and Coalescent Simulations.
507 *Systematics Biology*, **61**, 367-381.
- 508 Palacios, F. (1996) Systematics of the indigenous hares of Italy traditionally identified
509 as *Lepus europaeus* Pallas, 1778 (Mammalia: Leporidae). *Bonner Zoologische*
510 *Beitraege*, **5**, 59-91.
- 511 Palomo, L.J., Gisbert, J. & Blanco, J.C. (2007) *Atlas y Libro Rojo de los Mamíferos*
512 *Terrestres de España*. Ministerio de Medio Ambiente, Madrid.
- 513 Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models
514 developed using logistic regression. *Ecological Modelling*, **133**, 225-245.
- 515 Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological
516 niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*.
517 *Global Ecology and Biogeography*, **17**, 135-144.
- 518 Peterson, A.T. & Nyári, A.S. (2007) Ecological niche conservatism and Pleistocene
519 refugia in the thrush-like mourner, *Schiffornis* sp., in the neotropics. *Evolution*,
520 **62-1**, 173-183
- 521 Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of
522 evidence. *Journal of Biogeography*, **38**, 817-827.
- 523 Peterson, A.T., Soberón, J. & Sánchez-Cordero, V. (1999) Conservatism of ecological
524 niches in evolutionary time. *Science*, **285**, 1265–1267

1
2
3 525 Pierpaoli, M., Riga, F., Trocchi, V. & Randi, E. (1999) Species distinction and
4
5 526 evolutionary relationships of the Italian hare (*Lepus corsicanus*) as described by
6
7 527 mitochondrial DNA sequencing. *Molecular Ecology*, **8**, 1805-1817.
8
9
10 528 Pietri, C., Alves, P. & Melo-Ferreira, J. (2011) Hares in Corsica: high prevalence of
11
12 529 *Lepus corsicanus* and hybridization with introduced *L. europaeus* and *L.*
13
14 530 *granatensis*. *European Journal of Wildlife Research*, **57**, 313-321.
15
16 531 R Core Team (2012) R: A language and environment for statistical computing. R
17
18 532 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0,
19
20 533 URL <http://www.R-project.org/>.
21
22
23 534 Randin, C.F., Dirnbock, T., Dullinger, S., Zimmermann, N., Zappa, M. & Guisan, A.
24
25 535 (2006) Are niche-based models transferable in space? *Journal of Biogeography*,
26
27 536 **33**, 1689-1703.
28
29 537 Rödder, D. & Engler, J.O. (2011) Quantitative metrics of overlaps in Grinnellian niches:
30
31 538 advances and possible drawbacks. *Global Ecology and Biogeography*, **20**, 915-
32
33 539 927.
34
35
36 540 Scalera, R. & Angelici, F.M. (2003) Rediscovery of the Apennine hare *Lepus*
37
38 541 *corsicanus* in Corsica. *Bollettino del Museo Regionale Scienze Naturali Torino*,
39
40 542 **20**, 161-166.
41
42
43 543 Schluter, D. (2009) Evidence for Ecological Speciation and Its Alternative. *Science*,
44
45 544 **323**, 737-741.
46
47 545 Soberón, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological
48
49 546 niches and species' distribution areas. *Biodiversity Informatics*, **2**, 1-10.
50
51
52 547 Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus
53
54 548 conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868-
55
56 549 2883.
57
58
59
60

- 1
2
3 550 Wickham, H. (2009) ggplot2: elegant graphics for data analysis. Springer New York.
4
5 551 Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology,
6
7 552 and conservation biology. *Annual Review of Ecology, Evolution, and*
8
9 553 *Systematics*, **36**, 519-539.
10
11 554 Wu, C., Wu, J., Bunch, T.D., Li, Q., Wang, Y. & Zhang, Y. (2005) Molecular
12
13 555 phylogenetics and biogeography of *Lepus* in Eastern Asia based on
14
15 556 mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **37**, 45-
16
17 557 61.
18
19
20
21 558
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

559 **SUPPORTING INFORMATION**

560 Additional Supporting Information may be found in the online version of this article:
561 **Appendix S1** Controlling dimensionality effects in niche conservatism interpretations.
562 **Appendix S2** Checking for models transferability between study areas.
563 As a service to our authors and readers, this journal provides supporting information
564 supplied by the authors. Such materials are peer-reviewed and may be re-organized for
565 online delivery, but are not copy-edited or typeset. Technical support issues arising
566 from supporting information (other than missing files) should be addressed to the
567 authors.

568
569 **BIOSKETCH**

570 **Pelayo Acevedo** is a researcher at the CIBIO, Centro de Investigação em
571 Biodiversidade e Recursos Genéticos, InBio Laboratório Associado. His interests
572 include the study of factors affecting the distribution and abundance of pathogens, and
573 their hosts and vectors, through fragmented habitats.
574

575 **Author contributions:** P.A. and P.C.A. conceived the ideas; P.A. analysed the data;
576 P.A., J.M.-F., R.R. and P.C.A. participated in the discussion of the results and wrote the
577 manuscript.

578
579 Editor: Miles Silman
580

Table 1. Predictors used in the different models to study the *Lepus castroviejoi* and *L. corsicanus* distributions. A quarter is a period of three months (1/4 of the year).

Factor	Codes	Description (units)
Topoclimate	BIO0	Altitude (masl)
	BIO2	Mean diurnal range (mean of monthly [max T - min T]) (°C*10)
	BIO3	Isothermality (BIO2/Temperature annual range) (*100)
	BIO4	Temperature seasonality (standard deviation*100)
	BIO9	Mean temperature of driest quarter (°C*10)
	BIO15	Precipitation seasonality (coefficient of variation)
	BIO18	Precipitation of warmest quarter (mm)
	BIO19	Precipitation of coldest quarter (mm)
Land uses	T11	Post-flooding or irrigated croplands (or aquatic) (%)
	T14	Rainfed croplands (%)
	T20	Mosaic cropland / vegetation (grassland/shrubland/forest) (%)
	T30	Mosaic vegetation / cropland (%)
	T70	Closed needleleaved evergreen forest (%)
	T100	Closed to open (>15%) mixed forest (%)
	T110	Mosaic forest-shrubland / grassland (%)
	T120	Mosaic grassland / forest or shrubland (%)
	T130	Closed to open shrubland (%)
	T150	Sparse (<15%) vegetation (%)
	T210	Water bodies (%)

Table 2. Results of the models developed on the current distribution of *Lepus castroviejo* and *L. corsicanus* distributions. Predictors were listed following the order of entrance (the first one in the top) in the stepwise procedure. B parameter coefficient and its standard error (SE), z-value test statistics, Sig. significance (*** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$ and ns $P > 0.05$). Predictors coded as in Table 1.

Model	Codes	B(SE)	z-value
Overall species distribution	BIO15	0.1708(0.0126)	13.510***
	BIO0	0.0066(0.0005)	14.011***
	BIO19	-0.0209(0.0018)	-11.822***
	BIO4	-0.0022(0.0002)	-10.383***
	BIO9	0.0231(0.0059)	3.943***
	T11	-0.3337(0.1682)	-1.984*
	BIO2	0.0182(0.0061)	2.964**
	T30	-0.0299(0.0142)	-2.104*
	T120	0.0959(0.0455)	2.106*
	T130	-0.0350(0.0166)	-2.111*
	T150	0.0251(0.0148)	-2.111ns
	Intercept	-1.1033(0.5063)	-2.179*
<i>L. castroviejo</i>	BIO9	-0.1622(0.0217)	-7.475***
	T120	-3.4344(1.7825)	-1.927ns
	T14	-0.0942(0.0377)	-2.501*
	Intercept	25.6815(3.4690)	7.403***
<i>L. corsicanus</i>	BIO15	0.1633(0.0140)	11.666***
	BIO19	-0.0251(0.0023)	-10.863***
	BIO0	0.0063(0.0023)	9.054***
	BIO4	-0.0020(0.0003)	-6.240***
	BIO9	0.0253(0.0076)	3.341***
	BIO2	0.0272(0.0114)	2.397*
	T30	-0.0443(0.0176)	-2.518*
	T120	0.1180(0.0493)	2.397*
	T11	-5.7220(20.680)	-0.028ns
	Intercept	-1.4300(0.5510)	-2.595**
<i>L. corsicanus</i> in mainland Italy	T70	0.0666(0.0228)	2.916**
	BIO0	0.0030(0.0005)	5.510***
	BIO15	0.0834(0.0156)	5.359***
	BIO2	0.0444(0.0157)	2.828**
	T210	-0.1804(0.1131)	-1.595ns
	Intercept	-11.2500(1.8910)	-5.952***

FIGURE LEGENDS

Figure 1. Location of the study areas. Current distributions of *Lepus castroviejo* and *L. corsicanus* (excluding range in Corsica) represented in UTM 10×10 km (marks).

Species distribution data were obtained from Palomo *et al.* (2007) for Iberian Peninsula and adapted from Boitani *et al.* (2002) for Italy. Trend surface analyses (see text for details) were used to delimit the arena for modelling purposes (areas in grey).

Figure 2. Probability of species occurrence in the study areas (from dark blue=1 to yellow=0) obtained from the different models: overall species distribution (a), model trained on *L. castroviejo* and then projected onto Italy (b), model trained on *L. corsicanus* and projected onto Iberia (c), and finally, model trained on *L. corsicanus*, but excluding its range in Sicily, and projected onto both Sicily and Iberia (d). Arrows denote the projection of a model from the training area onto the territory of the sister species.

Figure 3. Calibration plots showing the relationship between the predicted probability of occurrence for a model and the observed proportion of evaluation localities occupied by the species: overall species distribution (a), model trained on *L. castroviejo* (b), model trained on *L. corsicanus* (c), and finally, model trained on *L. corsicanus* but excluding Sicily (d). Arrows denote the projection of a model from the training area to the territory of the sister species. Full/open symbols indicate bins with ≥ 15 or < 15 localities, respectively. In bins with < 15 localities the observed frequency should be cautiously considered (Jovani & Tella, 2006). Significant codes: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$ and ns $P > 0.05$, refer to Hosmer-Lemeshow goodness-of-fit statistic (Lemeshow & Hosmer, 1982).

Figure 1

618 **Figure 2**

619

620

621 Figure 3

622

Supplementary material – Journal of Biogeography

Evidence for niche similarities in the allopatric sister species *Lepus castroviejoi* and *L. corsicanus* - - P. Acevedo, J. Melo-Ferreira, R. Real, P.C. Alves

Appendix S1. Controlling dimensionality effects in niche conservatism interpretations.

We carried out parallel analyses to be sure that the number of parameters in the models – model's dimensionality (see Peterson, 2011) – is not affecting our interpretations on niche conservatism. We run models following three different approaches: i) considering all available climate and land uses predictors (n=19; approach 1), ii) only climatic predictors (n=8) were considered, as they usually yield higher predictive capacity than land uses (approach 2), and iii) models were fitted with the orthogonal factors (n=6) resulting from a principal component analysis (approach 3). For each approach, final models were selected following a forwards-backwards stepwise procedure based on Akaike Information Criterion (AIC; Akaike, 1974). Results are summarized in Table S1 and Figure S1.

Table S1. Summary of the models carried out under approaches differing in the set of predictors. Akaike Information Criterion is provided as a measure of model parsimony. The number of predictors included in each final model is reported in brackets.

Approach	Model		
	<i>L. castroviejoi</i>	<i>L. corsicanus</i>	<i>L. corsicanus</i> in mainland Italy
<i>Approach 1</i>	157.16 (3)	626.94 (9)	410.97 (5)
<i>Approach 2</i>	161.69 (2)	630.29 (7)	413.56 (4)
<i>Approach 3</i>	161.79 (6)	862.35 (6)	426.73 (5)

Independently of the approach, similar interpretations could be obtained when models predictions were mapped (Table S1, Figure S1): only models trained on data for *L.*

corsicanus predicted potentiality in areas occupied by both species. Thus, no effects derived of the number of predictors are expected to affect our interpretations, and only results from approach 1 are shown in the main text since it consistently yielded the most parsimonious models (Table S1).

Figure S1. Probability of species occurrence in the study areas (from dark-blue=1 to yellow=0) obtained from the different approaches (see above for details) and models: model trained on *L. castroviejo* and then projected to Italy [A], model trained on *L. corsicanus* and transferred to Iberia [B], and finally, model trained on *L. corsicanus*, but excluding its range in Sicily, and projected to both Sicily and Iberia [C]).

Appendix S2. Checking for models transferability between study areas.

Three pivotal requirements should be checked when models are transferred outside their training area. First, for each model, Mahalanobis distances were computed to represent the ecological dissimilarity of the transferring area to the training area, i.e. the degree of extrapolation. Second, multicollinearity among predictors in a model can bias the predictions when this model is transferred outside the range where it was trained. The collinearity was quantified using the variance inflation factor (VIF). Given that collinearity can be a problem with VIFs higher than 3 (see Zuur *et al.*, 2010), VIFs were calculated for each predictor and study area as the inverse of the coefficient of non-determination of the regression of each predictor against all others. Finally, Mantel tests were run to assess the maintenance of the correlation structure of the predictors between the training and the transferring areas (see Jiménez-Valverde *et al.*, 2011). Pearson's correlation coefficient between the elements of the matrices was used as the statistical test, and its significance was assessed by permuting the row labels of one matrix relative to the other 9999 times.

Statistical analyses were carried out in R 2.15.2 (R Core Team 2012). The 'HH' package was used for the variance inflation factor analyses (Heiberger, 2012) and 'ade4' for the Mantel tests (Chessel *et al.*, 2004).

Results suggest that only in very few localities were the models extrapolated; thus, a certain degree of uncertainty in the predictions can be expected in these localities (Figure S2). According to the obtained VIF values, no effects of multicollinearity are expected when models were spatially transferred (mean VIF value / range: 1.72 / 1.37-2.82; 1.81 / 1.02-2.99; 1.80 / 1.25-2.91 for the areas of Iberia, Italy and mainland Italy (excluding Sicily), respectively). Similarly, the correlation structures of the predictors between training and transferring areas were maintained in all study cases, precisely

58 between Italian and Iberian areas ($r=0.27$, $p<0.05$), mainland Italy and Iberia ($r=0.28$,
59 $p<0.05$), and mainland Italy and Sicily ($r=0.51$, $p<0.01$). That is, the correlation
60 structures of the predictors between training and transferring areas were maintained in
61 all study cases.
62

63
64 **Figure S2.** Maps showing the degree of environmental dissimilarity – in terms of the predictors
65 retained in the models (see Table 2) – between the transferring area and the training ones.
66 Localities with values in the predictors that were outside their range in the training areas are
67 depicted in blue.

68
69 **References for the supplementary material**

70 Akaike, H. (1974) New look at statistical-model identification. *Transactions on*
71 *Automatic Control*, **19**, 716-723.
72 Chessel, D., Dufour, A.B. & Thioulouse, J. (2004) The ade4 package-I- One-table
73 methods. *R News* **4**, 5-10.

- 1
2
3 74 Heiberger, R.M. (2012) HH: Statistical Analysis and Data Display: Heiberger and
4
5 75 Holland. R package version 2.3-27. <http://CRAN.R-project.org/package=HH>
6
7 76 Jiménez-Valverde, A., Decae, A.E. & Arnedo, M.A. (2011) Environmental suitability of
8
9 77 new reported localities of the funnelweb spider *Macrothele calpeiana*: an
10
11 78 assessment using potential distribution modelling with presence-only
12
13 79 techniques. *Journal of Biogeography*, **38**, 1213-1223.
14
15
16 80 Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of
17
18 81 evidence. *Journal of Biogeography*, **38**, 817-827.
19
20
21 82 R Core Team (2012) R: A language and environment for statistical computing. R
22
23 83 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0,
24
25 84 URL <http://www.R-project.org/>.
26
27 85 Zuur, A.F., Ieno, E.N., Elphick, C.S. (2010) A protocol for data exploration to avoid
28
29 86 common statistical problems. *Methods in Ecology and Evolution*, **1**, 3-14.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60