

1
2
3 1 **Original Article**
4

5 2
6

7 3 **Range dynamics driven by Quaternary climate oscillations explain the distribution**
8

9 4 **of introgressed mtDNA of *Lepus timidus* origin in hares on the Iberian Peninsula**
10

11 5 Pelayo Acevedo^{1,2*}, José Melo-Ferreira², Liliana Farelo², Beatriz Beltran-Beck²,
12

13 6 Raimundo Real³, Rita Campos², Paulo C. Alves²
14

15 7 ¹ SaBio, Instituto de Investigación en Recursos Cinegéticos (IREC), CSIC-UCLM-
16

17 8 JCCM, Ciudad Real, Spain.
18

19 9 ² CIBIO (Centro de Investigação em Biodiversidade e Recursos Genéticos), InBIO
20

21 10 Laboratório Associado. Universidade do Porto, Portugal.
22

23 11 ³ Grupo de Biogeografía, Diversidad y Conservación, Departamento de Biología
24

25 12 Animal, Universidad de Málaga, Spain.
26

27 13 * **Correspondence:** Pelayo Acevedo. SaBio, Instituto de Investigación en Recursos
28

29 14 Cinegéticos (IREC), CSIC-UCLM-JCCM, Ciudad Real, Spain. E-mail:
30

31 15 pelayo.acevedo@gmail.com
32

33 16 **Running title:** Quaternary climate oscillations and range of introgressed mtDNA
34

35 17 **Word count:** 6482 (2 pages for tables and figures)
36

37 18
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 19 **ABSTRACT**
4

5 20 **Aim** A striking case of mitochondrial DNA (mtDNA) introgression occurs in the
6
7 21 Iberian Peninsula: the mtDNA of *Lepus timidus*, a species that inhabited this territory
8
9 22 during the Last Glacial Maximum (LGM), is almost fixed in the Iberian range of *Lepus*
10
11 23 *europaeus*, is very frequent in the northern half of Iberia within *Lepus granatensis* and
12
13 24 is fixed in *Lepus castroviejoii*. Here, we have used a transdisciplinary approach to
14
15 25 understand the ecological drivers of the distribution of this introgressed mtDNA in the
16
17 26 Iberian hare species.
18
19

20
21 27 **Methods** First, we have modelled the climatic niche of *L. timidus* to predict the
22
23 28 favourability for the species during the LGM. Second, we have gathered data regarding
24
25 29 *L. timidus* mtDNA introgression for 1137 individuals (139 populations) from the Iberian
26
27 30 Peninsula. We have tested whether the climatic favourability for *L. timidus* during the
28
29 31 LGM reflects the frequency of the introgressed mtDNA in the carrier hare populations.
30
31

32 32 **Results** The climatic favourability for the LGM is positively related to the presence of
33
34 33 individuals with *L. timidus* mtDNA (z -value=3.513, $P<0.001$), and discriminates better
35
36 34 than by chance between individuals with and without this lineage (AUC=0.712). At the
37
38 35 population level, the favourability values relate to the observed frequencies of mtDNA
39
40 36 introgression ($r=0.454$, $P<0.001$) and discriminate between populations with and
41
42 37 without *L. timidus* mtDNA (AUC=0.835).
43
44

45 38 **Main conclusions** The results suggest that the geographic distribution of *L. timidus*
46
47 39 mtDNA introgression is linked to *L. timidus* distribution in Iberia during the LGM. Our
48
49 40 study helps to understand the role of climate in the reticulated evolutionary history of
50
51 41 hares, and is a first step toward understanding the ecological impact of mtDNA
52
53 42 introgression on Iberian hares.
54
55
56
57
58
59
60

1
2
3 43 **Keywords:** climatic niche, Iberian hare, *Lepus granatensis*, mountain hare, reticulate
4
5 44 evolution, species distribution modelling.
6
7
8
9
10 46
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

For Peer Review

1
2
3 47 **INTRODUCTION**
4

5 48 Quaternary climatic changes have governed the geographical range dynamics of species
6
7 49 worldwide, affecting their spatial distribution and genetic diversity (Hewitt, 2000).
8
9 50 Patterns of genetic variation, when interpreted within the context of geography (s.s.
10
11 51 phylogeography), can provide important insights into the history of species (Avice,
12
13 52 2000). However, research in phylogeography has placed more emphasis on the phylo-
14
15 53 component than on geography, despite the inherent information that the latter contains
16
17 54 about the evolutionary envelope and ecology of the species (Kidd *et al.*, 2006).
18
19

20 55 The geography of species ranges can be understood by using species distribution
21
22 56 models (SDMs), which are descriptive and/or explanatory approaches for the
23
24 57 assessment of range drivers and the prediction of species occurrence in space and/or
25
26 58 time (see Guisan & Zimmermann, 2000). These approaches have been widely used to
27
28 59 address basic questions in ecology (Soberón, 2007) and conservation biogeography
29
30 60 (Lomolino, 2004). Moreover, SDMs can provide the information needed to generate
31
32 61 and test hypotheses concerning the past structure of genetic diversity (Richards *et al.*,
33
34 62 2007; Nogués-Bravo, 2009). When SDMs are analysed in an evolutionary context, they
35
36 63 can supply complementary information to that obtained from phylogeography (Diniz-
37
38 64 Filho *et al.*, 2013), as both SDMs and phylogeography are focused on the understanding
39
40 65 of macroevolutionary patterns, and the microevolutionary processes related to them
41
42 66 (Bermingham & Moritz, 1998; Ruegg *et al.*, 2006). Phylogeographical evidence can
43
44 67 thus be complemented by SDM studies, as the latter provide quantitative hypotheses to
45
46 68 test regarding species distribution in past scenarios and a statistical inference of the
47
48 69 main determinants of the distribution ranges (Svenning *et al.*, 2011). For instance, the
49
50 70 applications of SDMs in evolutionary studies allow researchers to propose and validate
51
52 71 hypotheses about the existence and location of glacial refugia (e.g. Carnaval & Moritz,
53
54
55
56
57
58
59
60

1
2
3 72 2008; Acevedo *et al.*, 2012a; de Lima *et al.*, 2014), to explore ecological relationships
4
5 73 within and between phylogenetically related taxa (Peterson *et al.*, 1999; Wiens &
6
7 74 Graham, 2005; Acevedo *et al.*, 2014) or to reconstruct the conditions that allowed,
8
9 75 among other things, historical hybridisation and genetic introgression to occur (Wielstra
10
11 & Arntzen, 2012).
12
13

14 77 Hares (*Lepus* spp.) are an important example in the study of hybridisation and
15
16 78 genetic introgression (Arnold, 2008). Past and ongoing events of hybridisation and/or
17
18 79 mitochondrial DNA (mtDNA) introgression among species of hares are being
19
20 80 increasingly reported in literature (e.g. Melo-Ferreira *et al.*, 2012). Most of the cases
21
22 81 reported involve the mountain hare, *Lepus timidus*, an arctic/boreal species that is
23
24 82 currently widespread in northern Eurasia, from the British Isles to the Russian Far East,
25
26 83 with some isolated populations in Scotland, Poland, the Alps and Japan (Smith &
27
28 84 Johnston, 2008). The range fluctuations of this species throughout the glacial cycles,
29
30 85 attested by fossil records that show the presence of the species during the Upper
31
32 86 Pleistocene in areas from where it later disappeared, such as Southern France (Lopez-
33
34 87 Martinez, 1980) or the Iberian Peninsula (Altuna, 1970), created multiple contacts with
35
36 88 other temperate hare species, thus facilitating hybridisation and introgression, most
37
38 89 notably of mtDNA (Alves *et al.*, 2008a). One of the most extreme cases of mtDNA
39
40 90 introgression among hares occurs in the Iberian Peninsula (e.g. Melo-Ferreira *et al.*,
41
42 91 2011) where mtDNA of *L. timidus* origin is almost fixed in the Iberian range of *Lepus*
43
44 92 *europaeus*, is very frequent in the northern half of Iberia for *Lepus granatensis* and is
45
46 93 fixed in *Lepus castroviejoi* (Alves *et al.*, 2008b; Melo-Ferreira *et al.*, 2012). On the
47
48 94 contrary, introgression at nuclear loci seems generally limited (Melo-Ferreira *et al.*,
49
50 95 2009, 2012; but see Melo-Ferreira *et al.*, 2011 for a case of massive X-linked
51
52 96 introgression into *L. granatensis*). Molecular data suggest two major time-frames for
53
54
55
56
57
58
59
60

1
2
3 97 these mtDNA introgression events, one during Middle Pleistocene, affecting the
4
5 98 common ancestor of *L. castroviejoi* and Italian *Lepus corsicanus*, and another at the
6
7 99 LGM, affecting all three Iberian species (Melo-Ferreira *et al.*, 2007, 2012). Whether
8
9
10 100 mtDNA introgression from *L. timidus* resulted from neutral demography, adaptive
11
12 101 processes or both is a matter of debate (Alves *et al.*, 2008b; Melo-Ferreira *et al.*, 2005,
13
14 102 2009, 2011). While adaptive introgression predicts that the prevalence of the
15
16 103 introgressed variants depends on the adaptive landscape of the haplotypes, introgression
17
18 104 driven by purely demographic processes strongly relies on the biogeographic and
19
20 105 demographic history of the involved species. Theoretical simulations show that
21
22 106 situations of competitive replacement with hybridization may lead to massive
23
24 107 introgression of genetic variants from the local into the invading species, due to drift in
25
26 108 the invasion front and/or recurrent hybridization, and that the pattern should be stronger
27
28 109 in markers linked to the least dispersing sex, as mtDNA in species with female
29
30 110 philopatry (Currat *et al.*, 2008). This may explain mtDNA introgression from *L. timidus*
31
32 111 into *L. granatensis* if the latter competitively replaced the former in northern Iberia by
33
34 112 the end of the last glacial period (Melo-Ferreira *et al.*, 2007). A recent study carried out
35
36 113 in the contact zone between *L. granatensis* and *L. europaeus* in northern Iberia
37
38 114 suggested that recurrent introgression along the invasion front may better explain the
39
40 115 patterns of mtDNA differentiation and that the current distribution of mtDNA of *L.*
41
42 116 *timidus* origin in the Iberian hare species could thus be linked to the historical
43
44 117 distribution of the species in the region (Melo-Ferreira *et al.*, 2014a).

45
46
47 118 To date the case of *L. timidus* mtDNA introgression in Iberia has only been
48
49 119 addressed by means of molecular approaches. In this work, using SDMs, we test
50
51 120 whether the distribution of this mtDNA lineage in other hare species reflects the past
52
53 121 distribution of *L. timidus* in Iberia (Melo-Ferreira *et al.*, 2014a). This analysis helps
54
55
56
57
58
59
60

1
2
3 122 understanding the role of climate-driven range shifts in the reticulated evolutionary
4
5 123 history of hares.
6

7 124

9 125 **MATERIAL AND METHODS**

11 126 **Species data and geographical extent for modelling**

12 127 The current distribution of *L. timidus* in Europe was extracted from The Atlas of
13
14
15
16 128 European Mammals (Mitchell-Jones *et al.*, 1999), with UTM 50 × 50 km squares as
17
18 129 territorial units (Fig. 1). Two procedures were applied in order to select an appropriate
19
20
21 130 geographical background in which to parameterise the species niche model:

22
23 131 1) Given that a geographically-biased sampling effort was acknowledged by the
24
25 132 authors of the Atlas, we calculated the number of mammal species in each square in
26
27 133 order to consider only for modelling those squares in which at least one species was
28
29 134 reported. This filter excluded territories from Eastern Europe which coincided with the
30
31 135 most incomplete sampling areas identified by the Atlas' authors (A.J. Mitchell-Jones,
32
33 136 personal communication), and resulted in the selection of 2557 squares.

34
35
36 137 2) Since the extent of the geographical background (i.e. extent of the study area)
37
38 138 has substantial effects on the outputs of species distribution modelling (see Barve *et al.*,
39
40 139 2011; Acevedo *et al.*, 2012b), we delimited an adequate territory in order to study *L.*
41
42 140 *timidus* distribution within the 2557 squares selected in the previous step. This was done
43
44
45 141 by modelling the species distribution in which the third-degree polynomial of the spatial
46
47 142 coordinates were considered as predictors (TSA; Trend Surface Analysis). The
48
49 143 geographical background for niche modelling was represented by the localities which,
50
51 144 after carrying out the TSA, had a predicted favourability that was higher than the
52
53
54 145 minimum value assigned to a presence (for further details see Acevedo *et al.*, 2012b; see
55
56 146 Fig. 1). This procedure selected 1582 squares, of which *L. timidus* was reported in 760.
57
58
59
60

147 **Determining the species climatic niche**

148 The climatic drivers of *L. timidus* distribution – i.e. its “climatic niche” – were
149 determined on the basis of the locations in which the species currently occurs. The niche
150 model was parameterized using an 80% random sample of the species distribution data
151 (training dataset) and evaluated against the remaining 20% of the data (validation
152 dataset). Nineteen bioclimatic variables (spatial resolution of ~1km) were obtained from
153 the WorldClim project database for the present (see Hijmans *et al.*, 2005): annual mean
154 temperature (BIO1), mean diurnal range (mean of monthly (BIO2=max temp-min
155 temp), isothermality (BIO3=(BIO2/BIO7) * 100), temperature seasonality
156 (BIO4=standard deviation *100); max temperature of warmest month (BIO5), min
157 temperature of coldest month (BIO6), temperature annual range (BIO7=BIO5-BIO6);
158 mean temperature of wettest quarter (BIO8), mean temperature of driest quarter (BIO9),
159 mean temperature of warmest quarter (BIO10), mean temperature of coldest quarter
160 (BIO11), annual precipitation (BIO12), precipitation of wettest month (BIO13),
161 precipitation of driest month (BIO14), precipitation seasonality (BIO15; coefficient of
162 variation), precipitation of wettest quarter (BIO16), precipitation of driest quarter
163 (BIO17), precipitation of warmest quarter (BIO18), and precipitation of coldest quarter
164 (BIO19). Multicollinearity among predictor variables in a model can bias the
165 predictions when the model is transferred outside the range/timeframe in which it was
166 trained, and in which correlation among variables may be different, as in this study. In
167 order to avoid multicollinearity-derived problems, we quantified the variance inflation
168 factor (VIF) in the training dataset in order to exclude those predictors with VIF>10
169 from the analyses prior to modelling (Montgomery & Peck, 1992). VIFs were
170 calculated for each predictor as the inverse of the coefficient of non-determination of
171 the regression of each predictor against all others by using the 'HH' R package

1
2
3 172 (Heiberger, 2012). The predictors selected after controlling the VIF were considered in
4
5 173 a multiple logistic regression (Hosmer & Lemeshow, 1989), in which the *L. timidus*
6
7 174 distribution was used as a response variable and the bioclimatic variables selected were
8
9 175 used as predictors. The final model was obtained by using a forwards-backwards
10
11 176 stepwise procedure based on the Akaike Information Criterion (AIC).

12
13
14 177 In order to test that our predictions are not biased for the selection of an unique
15
16 178 modelling technique, we assessed the consistency of the predictions obtained from
17
18 179 logistic regression when they are compared with those derived from seven different
19
20 180 techniques widely used in species distribution modelling (namely generalized additive
21
22 181 model, boosted regression trees, classification tree analysis, artificial neural network,
23
24 182 BIOCLIM and flexible discriminant analysis) and, in addition, an ensemble of their
25
26 183 forecasts (see Appendix 1 in the Supporting Information). The results showed that the
27
28 184 logistic regression is as robust as obtained by other methods and the predicted pattern
29
30 185 for LGM from logistic regression is consistent with the obtained for most explored
31
32 186 methods (Fig. S1, Tables S1 and S2). As model's complexity (e.g. Merow *et al.*, 2014)
33
34 187 limit the capability of models to be projected, in our study we opted to show results
35
36 188 from logistic regression as a simpler, well-known technique able to produce robust
37
38 189 inference.

39
40
41
42 190 Two components of the model's predictive performance were assessed in the
43
44 191 evaluation dataset, namely discrimination and reliability (Jiménez-Valverde *et al.*,
45
46 192 2013). The area under the curve of the receiver operating characteristic plot (AUC) was
47
48 193 computed by using the 'ROCR' R package (Sing *et al.*, 2012) to assess the
49
50 194 discriminatory power of the model in the evaluation data. The reliability of the predicted
51
52 195 probabilities was estimated by exploring the calibration plots (Pearce & Ferrier, 2000).
53
54 196 Graphs were constructed using the 'ggplot2' R package (Wickham, 2009) by plotting the
55
56
57
58
59
60

1
2
3 197 proportion of occupied evaluation sites against the predicted probability of presence (for
4
5 198 the ten equally sized probability intervals); points located along the 45° line indicate
6
7 199 perfect calibration. The Hosmer–Lemeshow goodness-of-fit statistic (H-L; Lemeshow
8
9 200 & Hosmer, 1982) was used to statistically assess whether the observed values matched
10
11 201 those expected in the calibration plots.

12
13
14 202 The objective of the model calibrated for the present was to identify the climatic
15
16 203 favourability for the species. The parameterised model was therefore used to identify
17
18 204 the location of the climatically favourable areas for the species during the LGM, when
19
20 205 *L. timidus* was present in the Iberian Peninsula. This was done by projecting the model
21
22 206 on bioclimatic variables for the LGM (Braconnot *et al.*, 2007), which are available at a
23
24 207 spatial resolution of ~5km. Previously, as the models are not able to accurately predict
25
26 208 beyond the range of values of the predictors used for training (Campbell, 2004), a
27
28 209 multivariate environmental similarity surface (MESS) analysis was developed (Elith *et*
29
30 210 *al.*, 2010). The MESS was used to estimate how similar each square in the LGM was to
31
32 211 the training dataset (the present), in relation to the predictors retained in the final model.
33
34 212 Negative values were obtained for those squares in which at least one predictor had a
35
36 213 value that was outside the range of environments represented in the training dataset, and
37
38 214 these are therefore climatic conditions that differ from those that the species is currently
39
40 215 experiencing, where the uncertainty of the predictions is high. Therefore, the model was
41
42 216 not projected to the past in those squares outside the climatic range represented in the
43
44 217 training dataset.

45 218 **Determining *L. timidus* mtDNA introgression in hares from the Iberian Peninsula**

46 219 We used both data concerning *L. timidus* introgression reported in previous studies
47
48 220 carried out by our research group (Alves *et al.*, 2008b; Melo-Ferreira *et al.*, 2011) and
49
50 221 new data obtained to improve the representativeness of the populations sampled. A total
51
52
53
54
55
56
57
58
59
60

1
2
3 222 of 1137 individuals (139 populations) were analysed, including samples of *L. europaeus*
4
5 223 ($n=179$ individuals) and *L. granatensis* ($n=958$ individuals) – see Fig. 1 and Appendix
6
7 224 S2.

8
9 225 Total genomic DNA was extracted from frozen liver or ear tissue using the
10
11 226 EasySpin Genomic DNA Tissue Kit (Citomed) or the JETQUICK Tissue DNA
12
13 227 Purification Kit (Genomed). *Lepus timidus* mtDNA lineage was detected in each of the
14
15
16 228 animals sampled, using the polymerase chain reaction-restriction fragment length
17
18 229 polymorphism (PCR-RFLP) approach implemented by Melo-Ferreira *et al.* (2005) or by
19
20 230 the amplification and sequencing of a mitochondrial control region fragment, as
21
22 231 described by Melo-Ferreira *et al.* (2007).

232 **Relating model predictions for the LGM to *L. timidus* mtDNA in the Iberian**

233 **Peninsula**

234 Climatic favourability for *L. timidus* occurrence during the LGM was related to the *L.*
235 *timidus* mtDNA detection in hares from the Iberian Peninsula at both the individual and
236 the population levels. At the individual level we first explored the capability of the
237 predicted probabilities to explain the presence of mtDNA foreign lineage by employing
238 a generalised linear mixed model (GLMM; family "binomial-logit"), in which
239 population was included as a random factor, and which was run with the use of the
240 'lme4' R package (Bates *et al.*, 2012). We additionally estimated the probability that an
241 animal with *L. timidus* mtDNA lineage chosen at random would be sampled in a square
242 with a higher probability for *L. timidus* occurrence during the LGM than an animal
243 without *L. timidus* mtDNA also chosen at random, i.e. the AUC at the individual level.

244 The same rationale used at the individual level was also used to assess the
245 predictive performance of the model at the population level. In order to avoid the
246 inclusion of potential bias owing to an insufficient sampling size in the populations

1
2
3 247 sampled, only those populations that contained more than 10 sampled animals were
4
5 248 considered in these analyses. The discriminatory capacity of the model predictions to
6
7 249 separate populations with and without introgressed lineages, independently of their
8
9 250 frequency, was assessed using AUC, and a Pearson correlation was carried out to
10
11 251 explore the relationship between the predicted probabilities (probabilities of *L. timidus*
12
13 252 occurrence during LGM) and the observed *L. timidus* mtDNA origin frequencies.
14
15
16
17

18 254 **RESULTS**

19
20 255 From the initial set of predictors, only eight (BIO2, BIO3, BIO5, BIO8, BIO9, BIO15,
21
22 256 BIO18 and BIO19) were considered for building up the climatic niche model, as they
23
24 257 had all achieved VIFs<10. In the stepwise procedure (Table 1, see also Fig. 2) all the
25
26 258 predictors related to temperature were selected earlier than those related to precipitation.
27
28 259 The model achieved a good performance in terms of both discrimination (AUC=0.907)
29
30 260 and calibration (H-L: $\chi^2=11.28$, $P=0.186$; see Fig. 2) when it was assessed in the
31
32 261 validation dataset.
33
34
35

36 262 Previous to the model's transference to the LGM, the MESS analysis showed
37
38 263 that only southern areas of Europe had the LGM climatic conditions included in the
39
40 264 gradient represented in the calibration dataset, and the projection of the model to the
41
42 265 past was thus restricted to these areas. The results showed that the climatic favourability
43
44 266 for *L. timidus* is currently more restricted than it was in the LGM, when climatic
45
46 267 conditions in the Mediterranean peninsulas were favourable for it (Fig. 3).
47
48

49 268 The predicted probability for the LGM was significantly related to the presence
50
51 269 of individuals with *L. timidus* mtDNA (GLMM: z -value=3.513, $P<0.001$), and there
52
53 270 was a better than random discrimination between individuals with and without *L.*
54
55 271 *timidus* mtDNA (AUC=0.712). We also obtained consistent results at the population
56
57
58
59
60

1
2
3 272 level: the predicted probabilities related to the observed frequencies of mtDNA
4
5 273 introgression (Pearson's $r=0.454$, $n=42$, $P<0.001$) and the discrimination between
6
7 274 populations hosting the *L. timidus* mtDNA lineage and populations without this mtDNA
8
9 275 were better than by chance (AUC=0.835).
10

11 276

12 277 **DISCUSSION**

13
14
15
16 278 Genetic and paleontological data suggest that climate oscillations played an important
17
18 279 role in modulating the geographic distribution, evolutionary history and interspecific
19
20 280 interactions of *L. timidus* in Western Europe (Altuna, 1970; Lopez-Martinez, 1980;
21
22 281 Melo-Ferreira *et al.*, 2007). Ancient hybridisation in areas in which *L. timidus* is
23
24 282 currently locally extinct, such as in the Iberian Peninsula, left genetic traces in the native
25
26 283 species, which is particularly visible in their mtDNA (Melo-Ferreira *et al.*, 2005; Alves
27
28 284 *et al.*, 2008a). A recent genetic study has suggested that the current geographic
29
30 285 distribution of mtDNA of *L. timidus* origin in the hare species from the Iberian
31
32 286 Peninsula could be a testament to *L. timidus* distribution in the region during the LGM
33
34 287 (Melo-Ferreira *et al.*, 2014a). The results reported here confirm this hypothesis from a
35
36 288 macroecological perspective.
37
38

39
40 289 According to our model (Table 1), *L. timidus* is distributed in cold areas with
41
42 290 strong diurnal contrasts in temperature and a high seasonality in the precipitation
43
44 291 regime, with very dry summers and very wet winters (e.g. Thulin, 2003; Acevedo *et al.*,
45
46 292 2012c). The four variables first selected in the stepwise procedure, and therefore acting
47
48 293 on a larger spatial scale (see Muñoz *et al.*, 2005), were related to temperature. The
49
50 294 strong dependence of *L. timidus* on temperature, and in particular on the maximum
51
52 295 temperature, reinforces the idea of high sensitivity to ongoing climate change, which
53
54 296 had previously been suggested for both this (Acevedo *et al.*, 2012c) and other species
55
56
57
58
59
60

1
2
3 297 inhabiting the Boreal-Alpine region (Thuiller *et al.*, 2005). Working on a large
4
5 298 ecophysiological dataset including terrestrial ectotherms, endotherms and plants, Araújo
6
7 299 *et al.* (2013) found that, in contrast to the lower thermal limit of the species climatic
8
9
10 300 niche, the upper thermal limit is highly conserved in nature. As a consequence the
11
12 301 species' capacity to adapt to climatic oscillations is principally constrained when the
13
14 302 climatic niche of the species is close to its upper limit. The fact that *L. timidus*
15
16 303 distribution is strongly limited by maximum temperatures suggests that its realized
17
18 304 climatic niche is close to its upper temperature limit. In that case, *L. timidus* would be
19
20 305 expected to have been strongly affected by post-glacial climate warming and to be also
21
22 306 highly affected by ongoing climate change. This prediction is in agreement with the
23
24 307 results obtained in previous works (Acevedo *et al.*, 2012c). This study thus provides
25
26 308 new support for the potential role of climate in the decreasing demographic trend of this
27
28 309 species in some European countries (e.g. Thulin, 2003), and we strongly suggest that
29
30 310 climate must be considered among the potential factors that limit the viability of
31
32 311 populations of *L. timidus* in Western Europe (Smith & Johnston, 2008).

33
34
35
36 312 Our model of *L. timidus* climatic favourability at the LGM produced a plausible
37
38 313 scenario for the past distribution of this species in Iberia, with relevant insights into the
39
40 314 biogeographical settings that drove mtDNA introgression into other Iberian hare
41
42 315 species, in particular *L. granatensis*, despite the local extinction of *L. timidus* in the
43
44 316 region. In keeping with theoretical predictions of Currat *et al.* (2008), the northwards
45
46 317 gradient of increasing frequencies of mtDNA introgression of *L. timidus* origin into *L.*
47
48 318 *granatensis* and the signs of expansion of the introgressed lineage, conform to a model
49
50 319 of competitive replacement with hybridization of the resident *L. timidus* by an invading
51
52 320 *L. granatensis* after the LGM, if the first was present in northern Iberia and the latter
53
54 321 expanded from the south, spreading the traces of introgression into the newly colonized
55
56
57
58
59
60

1
2
3 322 territory (Melo-Ferreira *et al.*, 2005, 2007). Analyses of genetic diversity among the
4
5 323 native *L. granatensis* mtDNA haplotypes further refined this model, suggesting that the
6
7 324 species may have expanded from a refugium in central Iberia, and that the post-glacial
8
9 325 expansion of the species may have been centrifugal (Melo-Ferreira *et al.*, 2011). This
10
11 326 hypothesis agrees with SDMs, which suggest that central Iberia (area around Caceres)
12
13 327 was strongly suitable for *L. granatensis* at the LGM (Acevedo *et al.*, 2012a).

14
15
16 328 Interestingly, the genetic isolation by distance of native *L. granatensis*
17
18 329 haplotypes is maximized if this area is considered the origin of a post-glacial expansion.
19
20 330 In addition, Acevedo *et al.* (2012a) suggested that some other areas, in the north and the
21
22 331 south, were also suitable for the species at the LGM. The present work suggests that *L.*
23
24 332 *timidus* may have been widespread in northern Iberia at the LGM, which could have
25
26 333 then impeded the presence of *L. granatensis* in the region. These areas could have
27
28 334 nevertheless acted as points of attraction, defining the post-glacial colonization routes of
29
30 335 an expanding *L. granatensis*. Similar range replacements may also explain massive
31
32 336 introgression into *L. europaeus*, which likely have arrived to Iberia after the LGM
33
34 337 (Stamatis *et al.*, 2009).

35
36
37 338 Another interesting pattern has been inferred from molecular data and deserves
38
39 339 consideration here: the introgressed haplotypes of *L. timidus* origin show strong east-
40
41 340 west structure in northern Iberia (Melo-Ferreira *et al.*, 2007). Considering this model of
42
43 341 northwards range replacement with hybridization of the resident *L. timidus* by an
44
45 342 invading *L. granatensis*, two phenomena may have contributed to this structured pattern
46
47 343 of introgression: i) the fixation of different introgressed alleles by drift in the invasion
48
49 344 front and amplification of these variants in the wave of colonisation; and/or ii) recurrent
50
51 345 hybridisation with a structured *L. timidus* population along the invasion front. A recent
52
53 346 molecular study suggested that recurrent hybridisation was probably involved, and that
54
55
56
57
58
59
60

1
2
3 347 the areas in which the introgressed haplotypes are currently found might thus represent
4
5 348 the ancestral distribution of *L. timidus* in Iberia during the LGM (Melo-Ferreira *et al.*,
6
7 349 2014a). The results presented here agree with this hypothesis, by showing that the
8
9 350 climatic favourability for the presence of *L. timidus* in northern Iberia during the LGM
10
11 351 is correlated with the frequency of introgression (Fig. 3).

12
13
14 352 The biogeographical scenarios put forward to explain mtDNA introgression do
15
16 353 not exclude the possibility that natural selection may have favoured mtDNA
17
18 354 introgression or explain the persistence of the introgressed haplotypes in northern
19
20 355 Iberia. Indeed, the comparison of the demographic signals of the introgressed and native
21
22 356 mtDNA lineages in northern Iberia suggest that the former may have outcompeted the
23
24 357 latter (Melo-Ferreira *et al.*, 2011). Also, a recent study has shown evidence of positive
25
26 358 selection in the evolution of mtDNA in hares, particularly in the arctic lineage, of which
27
28 359 *L. timidus* is part (Melo-Ferreira *et al.*, 2014b). Our results suggest that the current
29
30 360 distribution of mtDNA of *L. timidus* origin in the Iberian Peninsula could mimic a
31
32 361 particular ecological niche closely related to that expressed by *L. timidus* in Western
33
34 362 Europe, which may indicate some potential for local adaptation. Even though our results
35
36 363 do not demonstrate an adaptive advantage of *L. timidus* mtDNA in hares from Iberia (or
37
38 364 of any other genes that may have been affected by ancient hybridisation events), they
39
40 365 provide hypotheses to be tested in future field studies (Richards *et al.*, 2007). Future
41
42 366 studies based on sampling populations across the introgression gradient are needed to
43
44 367 characterise fitness-related parameters in both introgressed and non-introgressed
45
46 368 individuals, and to provide conclusive evidence regarding the ecological role of *L.*
47
48 369 *timidus* mtDNA in hares from Iberia. Also, given the close interaction of mitochondrial
49
50 370 and nuclear encoded peptides in key aspects of energy metabolism, the inspection of
51
52
53
54
55
56
57
58
59
60

1
2
3 371 putative co-evolution and co-introgression of nuclear genes would shed light onto the
4
5 372 adaptive nature of historical introgressive hybridization in this system.
6

7 373

8
9 374 **ACKNOWLEDGMENTS**

10
11 375 Sally Newton kindly reviewed the manuscript for grammar. We thank the Societas
12
13 376 Europaea Mammalogica and Tony Mitchell-Jones for providing the distribution data
14
15 377 used to prepare the Atlas of European mammals. P.A. is currently supported by the
16
17 378 Spanish Ministerio de Economía y Competitividad (MINECO) and Universidad de
18
19 379 Castilla-La Mancha (UCLM) through a ‘Ramón y Cajal’ contract (RYC-2012-11970).
20
21 380 J.M.-F. and R.C. hold FCT Investigator and postdoctoral grants, respectively, from the
22
23 381 Fundação para a Ciência e a Tecnologia (FCT), funded by Programa Operacional
24
25 382 Potencial Humano (POPH) – Quadro de Referencia Estrategico Nacional (QREN) from
26
27 383 the European Social Fund and by the Portuguese Ministerio da Educacao e Ciencia
28
29 384 (IF/00033/2014 and SFRH/BPD/64365/2009, respectively). This work is funded by
30
31 385 FEDER funds through the Operational Programme for Competitiveness Factors -
32
33 386 COMPETE and by National Funds through FCT - Foundation for Science and
34
35 387 Technology under the EXPL/AAG-MAA/1082/2013 and FCOMP-01-0124-FEDER-
36
37 388 041082. Partial found was also obtained from FCT-ANR/BIA-EVF/0250/2012 and by
38
39 389 the Spanish Ministry of Agriculture, Food and Environment, Spanish National Park’s
40
41 390 Network (project 1098/2014).
42
43
44
45

46
47 391 **REFERENCES**

48
49 392 Acevedo, P., Melo-Ferreira, J., Real, R. & Alves, P.C. (2012a) Past, present and future
50
51 393 distributions of an iberian endemic, *Lepus granatensis*: ecological and
52
53 394 evolutionary clues from species distribution models. *PLoS ONE*, **7(12)**, e51529
54
55
56
57
58
59
60

- 1
2
3 395 Acevedo, P., Jiménez-Valverde, A., Lobo, J.M. & Real, R. (2012b) Delimiting the
4
5 396 geographical background in species distribution modelling. *Journal of*
6
7 397 *Biogeography*, **39**, 1383-1390.
- 8
9 398 Acevedo, P., Jiménez-Valverde, A., Melo-Ferreira, J., Real, R. & Alves, P.C. (2012c)
10
11 399 Parapatric species and the implications for climate change studies: a case study
12
13 400 on hares in Europe. *Global Change Biology*, **18**, 1509-1519.
- 14
15 401 Acevedo, P., Melo-Ferreira, J., Real, R. & Alves, P.C. (2014) Evidence for niche
16
17 402 similarities in the allopatric sister species *Lepus castroviejoii* and *L. corsicanus*.
18
19 403 *Journal of Biogeography*, **41**, 977-986.
- 20
21 404 Altuna, J. (1970) Hallazgo de una liebre ártica (*Lepus timidus*) en el yacimiento
22
23 405 prehistorico de Urtiga (Guipuzcoa). *Munibe*, **22**, 165-168.
- 24
25 406 Alves, P.C., Melo-Ferreira, J., Branco, M., Suchentrunk, F., Ferrand, N. & Harris, D.J.
26
27 407 (2008b) Evidence for genetic similarity of two allopatric European hares (*Lepus*
28
29 408 *corsicanus* and *L. castroviejoii*) inferred from nuclear DNA sequences.
30
31 409 *Molecular Phylogenetics and Evolution*, **46**, 1191-1197.
- 32
33 410 Alves, P.C., Melo-Ferreira, J., Freitas, H. & Boursot, P. (2008a) The ubiquitous
34
35 411 mountain hare mitochondria: multiple introgressive hybridization in hares, genus
36
37 412 *Lepus*. *Philosophical Transactions of the Royal Society B*, **363**, 2831-2839.
- 38
39 413 Araújo, M.B., Ferri-Yáñez, F., Bozinovic, F., Marquet, P.A., Valladares, F. & Chown,
40
41 414 S.L. (2013) Heat freezes niche evolution. *Ecology Letters*, **16**, 1206-1219.
- 42
43 415 Arnold, M.L. (2008) *Reticulate evolution and humans*. Oxford University Press, NY.
- 44
45 416 Avise, J.C. (2000) *Phylogeography: the history and formation of species*. Harvard
46
47 417 University Press, Cambridge, MA.
- 48
49 418 Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson,
50
51 419 A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area
52
53
54
55
56
57
58
59
60

- 1
2
3 420 in ecological niche modeling and species distribution modeling. *Ecological*
4
5 421 *Modelling*, **222**, 1810-1819
- 6
7 422 Bates, D., Maechler, M. & Bolker, B. (2012) lme4: Linear mixed-effects models using
8
9 423 S4 classes. R package version 0.999999-0. [http://CRAN.R-](http://CRAN.R-project.org/package=lme4)
10
11 424 [project.org/package=lme4](http://CRAN.R-project.org/package=lme4)
- 12
13 425 Bermingham, E. & Moritz, C. (1998) Comparative phylogeography: concepts and
14
15 426 applications. *Molecular Ecology*, **7**, 367-369.
- 16
17 427 Braconnot, P., Otto-Bliesner, B., Harrison, S. *et al.* (2007) Results of PMIP2 coupled
18
19 428 simulations of the Mid-Holocene and Last Glacial Maximum - Part 1:
20
21 429 experiments and large-scale features. *Climate of the Past*, **3**, 261-277.
- 22
23 430 Campbell, S.K. (2004) *Flaws and fallacies in statistical thinking*. Dover Publications,
24
25 431 Inc., NY.
- 26
27 432 Carnaval, A.C. & Moritz, C. (2008) Historical climate modelling predicts patterns of
28
29 433 current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*,
30
31 434 **35**, 1187-1201.
- 32
33 435 Currat, M., Ruedi, M., Petit, R.J. & Excoffier, L. (2008) The hidden side of invasions:
34
35 436 massive introgression by local genes. *Evolution*, **62**, 1908-1920.
- 36
37 437 de Lima, N.E., Lima-Ribeiro, M.S., Tinoco, C.F., Terribile, L.C. & Collevatti, R.G.
38
39 438 (2014) Phylogeography and ecological niche modelling, coupled with the fossil
40
41 439 pollen record, unravel the demographic history of a Neotropical swamp palm
42
43 440 through the Quaternary. *Journal of Biogeography*, **4**, 673-686.
- 44
45 441 Diniz-Filho, J.A.F., Gouveia, S.F. & Lima-Ribeiro, M.S. (2013) Evolutionary
46
47 442 macroecology. *Frontiers of Biogeography*, **5.3**
- 48
49 443 Elith, J., Kearney, M., Phillips, S. (2010) The art of modelling range-shifting species.
50
51 444 *Methods in Ecology and Evolution*, **1**, 330-342.
- 52
53
54
55
56
57
58
59
60

- 1
2
3 445 Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in
4
5 446 ecology. *Ecological Modelling*, **135**, 147-186.
6
7 447 Heiberger, R.M. (2012) HH: Statistical analysis and data. Display: Heiberger and
8
9 448 Holland. R package version 2.3-27. <http://CRAN.R-project.org/package=HH>
10
11 449 Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907e913
12
13 450 Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high
14
15 451 resolution interpolated climate surfaces for global land areas. *International*
16
17 452 *Journal of Climatology*, **25**, 1965-1978.
18
19 453 Hosmer, D.W. & Lemeshow, S. (1989) *Applied logistic regression*. John Wiley and
20
21 454 Sons, NY.
22
23 455 Jiménez-Valverde, A., Acevedo, P., Barbosa, A.M., Lobo, J.M. & Real, R. (2013)
24
25 456 Discrimination capacity in species distribution models depends on the
26
27 457 representativeness of the environmental domain. *Global Ecology and*
28
29 458 *Biogeography*, **22**, 508-516.
30
31 459 Kidd, D.M. & Ritchie, M.G. (2006) Phylogeographic information systems: putting the
32
33 460 geography into phylogeography. *Journal of Biogeography*, **33**, 1851-1865.
34
35 461 Lemeshow, S. & Hosmer, D.W. (1982) A review of goodness of fit statistics for use in
36
37 462 the development of logistic regression models. *American Journal of*
38
39 463 *Epidemiology*, **115**, 92-106.
40
41 464 Lomolino, M.V. (2004) Conservation biogeography. In: *Frontiers of biogeography:*
42
43 465 *new directions in the geography of nature* (ed. by M.V. Lomolino and L.R.
44
45 466 Heaney), PP. 293–296. Sinauer Associates, Sunderland, MA.
46
47 467 Lopez-Martinez, N. (1980) Les lagomorphs (Mammalia) du pléistocène supérieur de
48
49 468 Jaurens. *Nouvelles Archives du Museum d'Histoire Naturelle de Lyon*, **18**, 5-16.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 469 Melo-Ferreira, J., Alves, P.C., Freitas, H., Ferrand, N. & Boursot, P. (2009) The
4
5 470 genomic legacy from the extinct *Lepus timidus* to the three hare species of
6
7 471 Iberia: contrast between mtDNA, sex chromosomes and autosomes. *Molecular*
8
9 472 *Ecology*, **18**, 2643-2658.
- 10
11 473 Melo-Ferreira, J., Alves, P.C., Rocha, J., Ferrand, N. & Boursot, P. (2011) Interspecific
12
13 474 X-chromosome and mitochondrial DNA introgression in the Iberian hare:
14
15 475 selection or allele surfing? *Evolution*, **65**, 1956-1968.
- 16
17 476 Melo-Ferreira, J., Boursot, P., Carneiro, M., Esteves, P.J., Farelo, L. & Alves, P.C.
18
19 477 (2012) Recurrent introgression of mitochondrial DNA among hares (*Lepus* spp.)
20
21 478 revealed by species-tree inference and coalescent simulations. *Systematic*
22
23 479 *Biology*, **61**, 367-381.
- 24
25 480 Melo-Ferreira, J., Boursot, P., Randi, E., Kryukov, A., Suchentrunk, F., Ferrand, N. &
26
27 481 Alves, P.C. (2007) The rise and fall of the mountain hare (*Lepus timidus*) during
28
29 482 Pleistocene glaciations: expansion and retreat with hybridization in the Iberian
30
31 483 Peninsula. *Molecular Ecology*, **16**, 605-618.
- 32
33 484 Melo-Ferreira, J., Boursot, P., Suchentrunk, F., Ferrand, N. & Alves, P.C. (2005)
34
35 485 Invasion from the cold past: extensive introgression of mountain hare (*Lepus*
36
37 486 *timidus*) mitochondrial DNA into three other hare species in northern Iberia.
38
39 487 *Molecular Ecology*, **14**, 2459-2464.
- 40
41 488 Melo-Ferreira, J., Farelo, L., Freitas, H., Suchentrunk, F., Boursot, P. & Alves, P.C.
42
43 489 (2014a) Home-loving boreal hare mitochondria survived several invasions in
44
45 490 Iberia: the relative roles of recurrent hybridisation and allele surfing. *Heredity*,
46
47 491 **112**, 265-273.
- 48
49 492 Melo-Ferreira, J., Vilela, J., Fonseca, M.M., da Fonseca, R., Boursot, P. & Alves, P.C.
50
51 493 (2014b) The elusive nature of adaptive mitochondrial DNA evolution of an
52
53
54
55
56
57
58
59
60

- 1
2
3 494 arctic lineage prone to frequent introgression. *Genome Biology and Evolution*, **6**,
4
5 495 886-896
6
7 496 Merow, C., Smith, M.J., Edwards, T.C. Jr, Guisan, A., McMahon, S.M., Normand, S.,
8
9 497 Thuiller, W., Wüest, R.O., Zimmermann, N.E. & Elith, J. (2014) What do we
10
11 498 gain from simplicity versus complexity in species distribution models?
12
13 499 *Ecography*, **37**, 1267-1281.
14
15
16 500 Mitchell-Jones, A.J., Amori, G., Bogdanowicz, W. *et al.* (1999) *The atlas of european*
17
18 501 *mammals*. T & AD Poyser Ltd, London.
19
20 502 Montgomery, D.C. & Peck, E.A. (1992) *Introduction to linear regression analysis*.
21
22 503 Wiley, NY.
23
24 504 Muñoz, A.R., Real, R., Barbosa, A.M. & Vargas, J.M. (2005) Modelling the distribution
25
26 505 of Bonelli's eagle in Spain: implications for conservation planning. *Diversity*
27
28 506 *and Distributions*, **11**, 477-486.
29
30
31 507 Nogués-Bravo, D. (2009) Predicting the past distribution of species climatic niches.
32
33 508 *Global Ecology and Biogeography*, **18**, 521-531.
34
35
36 509 Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models
37
38 510 developed using logistic regression. *Ecological Modelling*, **133**, 225-245.
39
40 511 Peterson, A.T., Soberón, J. & Sánchez-Cordero, V. (1999) Conservatism of ecological
41
42 512 niches in evolutionary time. *Science*, **285**, 1265-1267.
43
44
45 513 Richards, C.L., Carstens, B.C. & Knowles, L.L. (2007) Distribution modelling and
46
47 514 statistical phylogeography: an integrative framework for generating and testing
48
49 515 alternative biogeographical hypotheses. *Journal of Biogeography*, **34**, 1833-
50
51 516 1845.
52
53
54
55
56
57
58
59
60

- 1
2
3 517 Ruegg, K.C., Hijmans, R.J. & Moritz, C. (2006) Climate change and the origin of
4
5 518 migratory pathways in the Swainson's thrush, *Catharus ustulatus*. *Journal of*
6
7 519 *Biogeography*, **33**, 1172-1182.
- 8
9
10 520 Sing, T., Sander, O., Beerenwinkel, N. & Lengauer, T. (2009) ROCR: visualizing the
11
12 521 performance of scoring classifiers. R package version 1.0-4. Available at:
13
14 522 <http://CRAN.Rproject.org/package=ROCR> (accessed January 2011).
- 15
16 523 Smith, A.T. & Johnston, C.H. (2008) *Lepus timidus*. In: *IUCN 2013. IUCN Red List of*
17
18 524 *Threatened Species*. Version 2013.2. <www.iucnredlist.org>. Downloaded on 21
19
20 525 January 2014.
- 21
22
23 526 Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of
24
25 527 species. *Ecology Letters*, **10**, 1115-1123.
- 26
27 528 Stamatis, C., Suchentrunk, F., Moutou, K.A., *et al.* (2009) Phylogeography of the brown
28
29 529 hare (*Lepus europaeus*) in Europe: a legacy of southeastern Mediterranean
30
31 530 refugia? *Journal of Biogeography*, **36**, 515-528.
- 32
33
34 531 Svenning, J.C., Fløjgaard, C., Marske, K.A., Nógues-Bravo, D. & Normand, S. (2011)
35
36 532 Applications of species distribution modeling to paleobiology. *Quaternary*
37
38 533 *Science Reviews*, **30**, 2930-2947.
- 39
40 534 Thuiller, W., Lavorel, S. & Araújo, M.B. (2005) Niche properties and geographical
41
42 535 extent as predictors of species sensitivity to climate change. *Global Ecology and*
43
44 536 *Biogeography*, **14**, 347-357.
- 45
46
47 537 Thulin, C.-G. (2003) The distribution of mountain hares *Lepus timidus* in Europe: a
48
49 538 challenge from brown hares *L. europaeus*? *Mammal Review*, **33**, 29-42.
- 50
51 539 Wickham, H. (2009) *ggplot2: elegant graphics for data analysis*. Springer, NY.
52
53
54
55
56
57
58
59
60

- 1
2
3 540 Wielstra, B. & Arntzen, J.W. (2012) Postglacial species displacement in *Triturus* newts
4
5 541 deduced from asymmetrically introgressed mitochondrial DNA and ecological
6
7 542 niche models. *BMC Evolutionary Biology*, **12**, 161.
8
9 543 Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology,
10
11 544 and conservation biology. *Annual Review of Ecology, Evolution, and*
12
13 545 *Systematics*, **36**, 519-539.
14
15
16 546
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

For Peer Review

1
2
3 547 **SUPPORTING INFORMATION**

4
5 548 Additional Supporting Information may be found in the online version of this article:

6
7 549 **Appendix S1** Assessing the reliability of the past model predictions: comparative

8
9 550 analyses of seven – plus an ensemble of forecasts – techniques .

10
11 551 **Appendix S2** Introgression frequencies at the population level used in this manuscript.

12
13 552 As a service to our authors and readers, this journal provides supporting information

14
15 553 supplied by the authors. Such materials are peer-reviewed and may be re-organized for

16
17 554 online delivery, but are not copy-edited or typeset. Technical support issues arising

18
19 555 from supporting information (other than missing files) should be addressed to the

20
21 556 authors.

22
23 557

24
25 558 **BIOSKETCH**

26
27 559 Pelayo Acevedo is a researcher at the Instituto de Investigación en Recursos

28
29 560 Cinegéticos. His interests include the integration of macroecology and genomics to

30
31 561 extract further insights for conservation biology. In addition, he studies the factors

32
33 562 affecting the distribution and abundance of pathogens, and their hosts and vectors,

34
35 563 through fragmented habitats.

36
37 564

38
39 565 **Author contributions:** P.A., J.M.-F. and P.C.A. conceived the ideas; J.M.-F., L.F. and

40
41 566 R.C. compiled introgression data and performed the genetic analyses; P.A. calibrated

42
43 567 the niche models and carried out the statistical analyses; P.A., J.M.-F., L.F., B.B., R.R.,

44
45 568 R.C. and P.C.A. participated in the discussion of the results and wrote the manuscript.

46
47 569

48
49 570 Editor: Miles Silman

50
51 571

1
2
3 572 **Table 1.** Results of the model developed for the current distribution of *Lepus timidus*.
4
5 573 Predictors are listed following the order of entrance (the first at the top) in the stepwise
6
7 574 procedure. *B* parameter coefficient and its standard error (*SE*), *z-value* test statistics. All
8
9
10 575 coefficients are significant at $P < 0.01$.

Predictor	<i>B</i> (<i>SE</i>)	<i>z-value</i>
Max temperature of warmest month	-0.067 (0.005)	-13.121
Isothermality	-0.207 (0.024)	-8.766
Mean diurnal range	0.097 (0.012)	8.171
Mean temperature of driest quarter	0.001 (1E-4)	5.274
Precipitation seasonality	0.038 (0.013)	2.931
Precipitation of warmest quarter	-0.008 (0.002)	-4.314
Precipitation of coldest quarter	0.007 (0.002)	4.232
Intercept	12.162 (1.136)	10.704

576

577

578 **FIGURE LEGENDS**

579 **Figure 1.** (a) Current distribution of *Lepus timidus* in Western Europe; presence data
580 were referred to UTM 50 × 50 km grid cells and were taken from Mitchell-Jones *et al.*
581 (1999). The arena for modelling that was delimited according to trend surface analysis
582 (TSA) using a third-degree polynomial of the spatial coordinates applied to the
583 occurrence localities of *L. timidus* (see Acevedo *et al.*, 2012b. (b) Distribution of the
584 studied populations of *L. granantensis* and *L. europaeus* is shown.

585 **Figure 2.** (a) Probability of *Lepus timidus* occurrence in Western Europe for the present
586 according to the climate niche of the species presented in Table 1. (b) Calibration plots
587 showing the relationship between the predicted probability of occurrence for the model
588 and the observed proportion of evaluation localities currently occupied by *L. timidus*.
589 The number of evaluation localities in each interval of probability is also shown.

590 **Figure 3.** (a) Probability of *Lepus timidus* occurrence in Western Europe during the last
591 glacial maximum (LGM; ~21 ky before the present) according to the climate niche of
592 the species presented in Table 1. Territories with climatic conditions in the LGM
593 beyond those considered when modelling the climate niche of *L. timidus* were delimited
594 by means of a multivariate environmental similarity surface (MESS; see Elith *et al.*,
595 2010) and were not used to represent the probability of species occurrence. (b)
596 Frequencies of the *L. timidus* mtDNA lineages (section in black) in populations with a
597 sample size of ≥ 10 ; only the presence (grey boxes) or absence (white boxes) of *L.*
598 *timidus* mtDNA lineage is shown for the remaining sample populations.

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

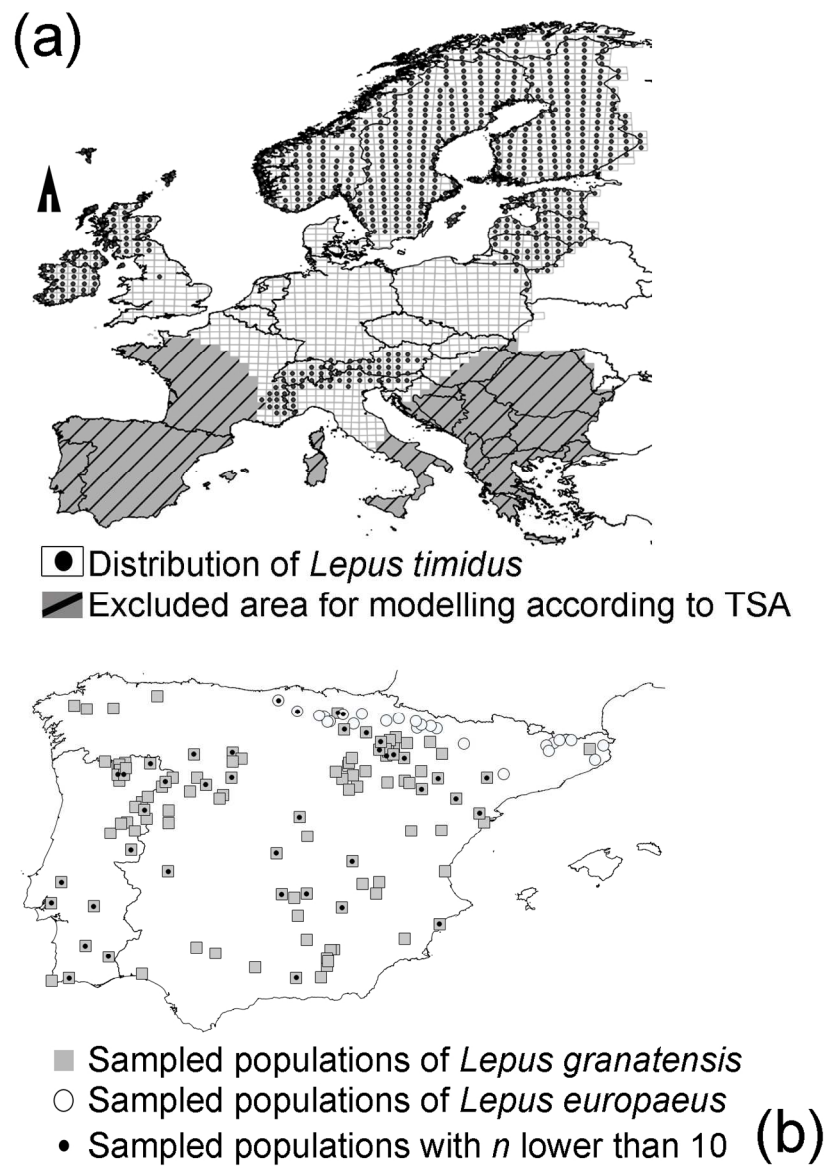


Fig.1

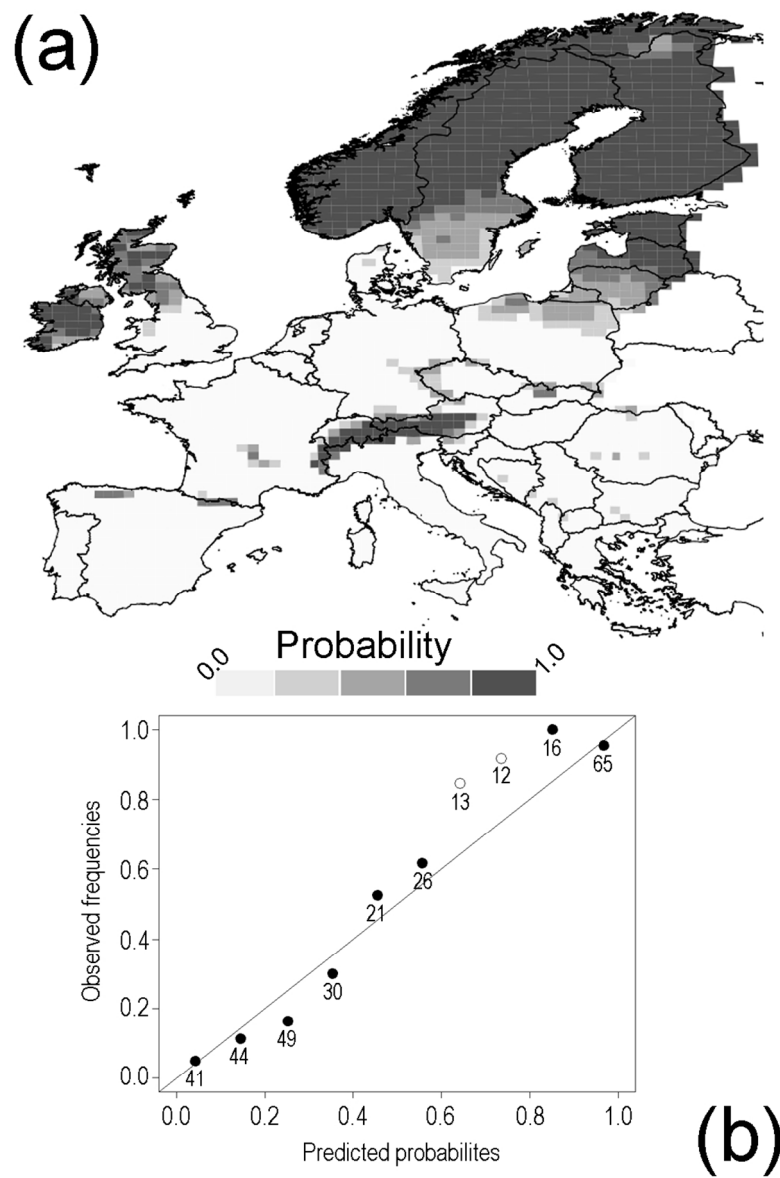


Fig. 2

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

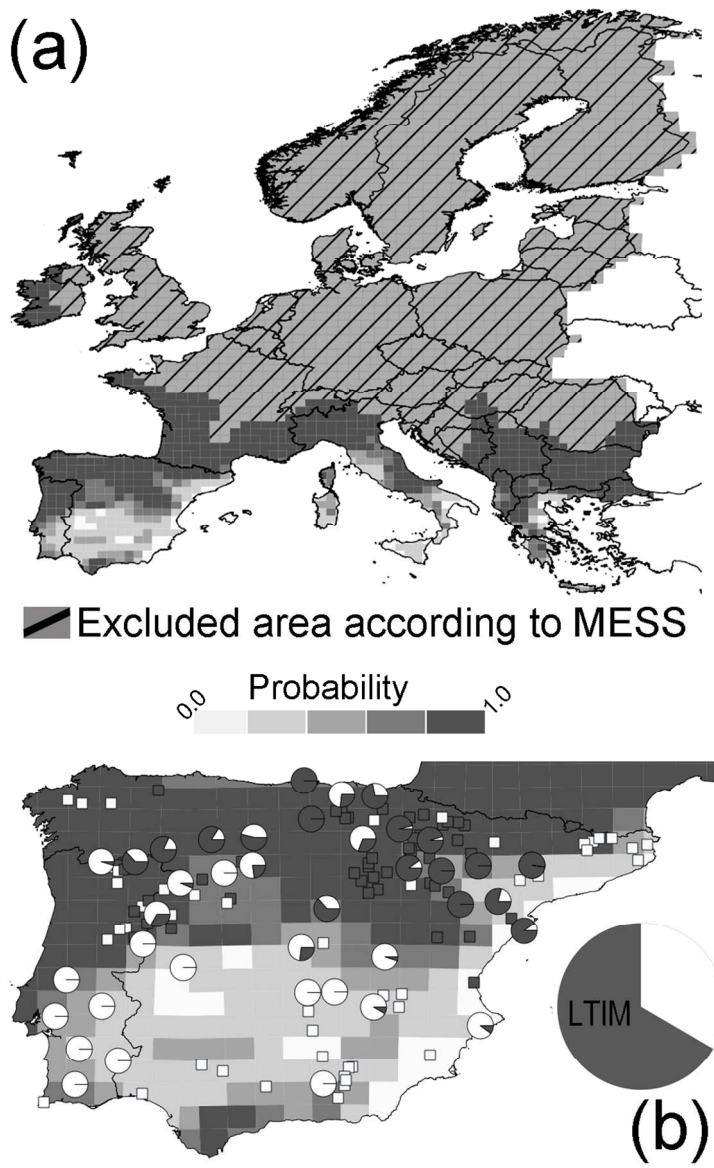


Fig. 3

Supplementary material – Journal of Biogeography

Range dynamics driven by Quaternary climate oscillations explain the distribution of introgressed mtDNA of *Lepus timidus* origin in hares on the Iberian Peninsula -

- P. Acevedo, J. Melo-Ferreira, L. Farelo, B. Beltran-Beck, R. Real, R. Campos, P. C.

Alves

Appendix S1. Assessing the reliability of the past model predictions: comparative analyses of seven – plus an ensemble of forecasts – techniques.

When modelling species distribution discrepancies among predictions obtained from different techniques can be produced. An operational solution to account for this variability is to simultaneously apply more than one technique and/or to fit ensembles of forecasts and analyze the resulting range of uncertainties with consensus methodologies (e.g. Araújo and New, 2007). We here showed the consistency of the predictions obtained from logistic regression when they are compared with those derived from seven different techniques widely used in species distribution modelling (namely Generalized Additive Model [GAM], Boosted Regression Trees [GBM], Classification Tree Analysis [CTA], Artificial Neural Network [ANN], BIOCLIM and Flexible Discriminant Analysis [FDA]) and, in addition, an ensemble of their forecasts. Analyses were performed with biomod2 using default specifications for each technique (for further details see e.g. Thuiller *et al.*, 2009, 2014).

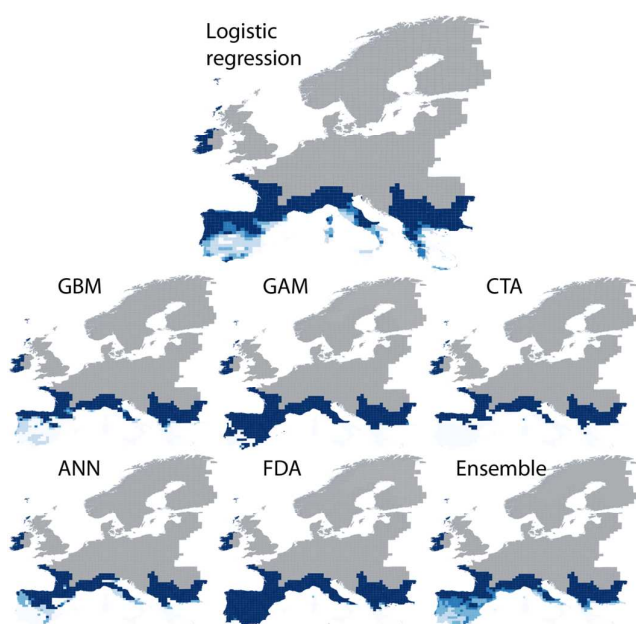
We have followed the same analytical design described in the main text. Briefly, previous to modelling we have delimited the geographical background using a Trend Surface Analyses (Acevedo *et al.*, 2012). The same background-delimited arena was used for parameterizing models from all techniques. Data were randomly split in calibration (80%) and validation (20%). Results suggested that all single techniques,

except BIOCLIM, performed adequately according to discrimination measures, concretely Cohen's Kappa, true skill statistic [TSS] and AUC (see Table S1). Only techniques with TSS>0.6 were considered for ensemble and the predictive performance of ensemble was similar to that achieved by other techniques.

Table S1. Predictive performance measures of each technique when they were assessed on the independent validation dataset.

Technique/Parameter	Kappa	TSS	AUC
Logistic regression	0.855	0.856	0.975
GBM	0.867	0.868	0.984
GAM	0.927	0.928	0.994
CTA	0.809	0.811	0.915
ANN	0.836	0.837	0.956
BIOCLIM	0.400	0.398	0.699
FDA	0.858	0.859	0.974
Ensemble of forecasts	0.885	0.886	0.987

Fig. S1. Climatic suitability for *Lepus timidus* occurrence in Western Europe during the Last Glacial Maximum (LGM). Territories with climatic conditions in the LGM beyond those considered when modelling the climate niche of *L. timidus* were delimited (grey-areas) by MESS (Elith *et al.*, 2010) and excluded of the analyses.



Cartographically, there is also a strong consistency when models were hindcasted to the Last Maximum Glacial (LGM; see Fig. S1). After the exclusion of territories with climatic conditions in the LGM beyond those considered when modelling the climate niche of *Lepus timidus* – defined by means of a multivariate environmental similarity surface (MESS; see Elith *et al.*, 2010) – predictions of species favourability from the different techniques were strongly correlated (Table S2).

Table S2. Pearson's correlation coefficients among predictions for LGM from different techniques on the area do not excluded after MESS analyses ($n=949$; $P<0.001$ in all cases). BIOCLIM did not achieve the threshold of $TSS>0.6$ and therefore it was not included in the ensemble and was excluded of these analyses.

Technique	GBM	GAM	CTA	ANN	FDA	Ensemble of forecasts
Logistic regression (main text)	0.832	0.679	0.763	0.755	0.607	0.863
GBM	-	0.701	0.964	0.807	0.605	0.909
GAM		-	0.622	0.694	0.774	0.886
CTA			-	0.750	0.497	0.845
ANN				-	0.600	0.863
FDA					-	0.828

Overall, the results showed that the logistic regression is as robust as obtained by other methods and the predicted pattern for LGM from logistic regression is consistent with the obtained for most explored methods. As model's dimensionality (Peterson, 2011) and complexity (e.g. Merow *et al.*, 2014) limit the capability of models to be projected, in our study we opted to show results from logistic regression as a simpler, well-known technique able to produce robust inference.

References

Acevedo, P., Jiménez-Valverde, A., Lobo, J.M. & Real, R. (2012b) Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*, 39, 1383-1390.

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22, 42-47.

Elith, J., Kearney, M., Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1, 330-342.

Merow, C., Smith, M.J., Edwards, T.C. Jr, Guisan, A., McMahon, S.M., Normand, S., Thuiller, W., Wüest, R.O., Zimmermann, N.E. & Elith, J. (2014) What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37, 1267-1281.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, 38, 817-827

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-373.

Thuiller, W., Georges, D. & Engler, R. (2014). biomod2: Ensemble platform for species distribution modeling. R package version 3.1-48. <http://CRAN.R-project.org/package=biomod2>.

Appendix S2. Raw data for the number of individuals of *Lepus granatensis* (LGR) and *Lepus europaeus* (LER) with mtDNA of *Lepus granatensis* lineage (GRA), *Lepus timidus* lineage (TIM) and *L. europaeus* lineage (EUR). The number of individuals that were published in previous studies (PUB data) and those which are new for this study (NEW data) are also shown.

Species	Code	XCOORD	YCOORD	Size	mtDNA lineage			Source	
					GRA	TIM	EUR	PUB data	NEW data
LGR	ALA1	-2.698387	42.910000	15	11	4	0	15	0
LGR	ALA1	-2.554819	42.550332	44	27	17	0	44	0
LGR	ALB	-1.858542	38.994350	3	3	0	0	3	0
LGR	ALB	-2.156911	39.207123	4	4	0	0	4	0
LGR	ALB	-1.797291	39.237885	6	6	0	0	6	0
LGR	ALB	-2.603424	38.700745	12	11	1	0	12	0
LGR	ALC	-0.133537	41.051037	24	5	19	0	24	0
LGR	ALI	-0.490686	38.345997	20	18	2	0	20	0
LGR	ALJ	-8.164983	37.876057	25	25	0	0	25	0
LGR	AMD	-6.840460	40.619846	3	1	2	0	3	0
LGR	AMD	-6.886843	40.803907	15	10	5	0	8	7
LGR	ARR	-7.986761	38.723221	21	21	0	0	21	0
LGR	BEN	-5.814288	42.010277	25	4	21	0	25	0
LGR	BRG	-6.756738	41.806114	11	2	9	0	11	0
LGR	CAC	-6.370961	39.471329	20	20	0	0	20	0
LGR	CBR	-7.172305	39.934232	24	24	0	0	24	0
LGR	CHV	-7.473165	41.741779	1	1	0	0	0	1
LGR	CHV	-7.306936	41.828781	3	0	3	0	0	3
LGR	CHV	-7.423473	41.790836	2	0	2	0	0	2
LGR	CRD	-4.485322	37.405640	6	6	0	0	6	0
LGR	CRE	-3.642048	38.910183	2	2	0	0	2	0
LGR	CRE	-3.374207	38.995762	12	12	0	0	12	0
LGR	CRE	-3.562719	38.521481	8	8	0	0	8	0
LGR	CRE	-3.910044	38.980040	11	11	0	0	3	8
LGR	CRN	-8.133835	43.000793	4	4	0	0	4	0
LGR	CRN	-8.408066	43.076374	2	2	0	0	2	0
LGR	CUE	-2.381535	39.690079	20	19	1	0	20	0
LGR	FCR	-6.942384	40.971539	1	1	0	0	1	0
LGR	FCR	-7.087496	40.877239	3	0	3	0	3	0
LGR	GER	2.765149	42.118183	1	1	0	0	0	1
LGR	GRN	-3.598557	37.177338	25	25	0	0	25	0
LGR	GRN	-2.929444	37.428333	1	1	0	0	0	1
LGR	GRN	-3.061281	37.182758	3	3	0	0	0	3
LGR	GRN	-2.780160	37.795662	1	1	0	0	0	1
LGR	GRN	-2.851806	37.790000	1	1	0	0	0	1
LGR	GRN	-2.927338	37.607674	1	1	0	0	0	1
LGR	GRN	-2.908233	37.529099	1	1	0	0	0	1
LGR	HLV	-6.944722	37.261421	2	2	0	0	0	2
LGR	HUE	-0.690008	42.275848	2	0	2	0	2	0
LGR	HUE	-0.433141	42.018589	5	0	5	0	5	0
LGR	JAE	-3.371246	38.011425	1	1	0	0	0	1
LGR	VBP	-8.889770	37.102550	8	8	0	0	8	0
LGR	LRD	0.538652	41.500664	10	0	0	10	0	10

LGR	LUG	-7.556758	43.009739	1	1	0	0	0	1
LGR	MAD	-3.351601	40.232246	3	3	0	0	3	0
LGR	MAD	-3.523562	40.650913	22	8	14	0	22	0
LGR	MER	-7.660666	37.641514	12	12	0	0	12	0
LGR	MNT	-7.732006	41.850042	8	8	0	0	0	8
LGR	MOS	-0.449361	40.362770	1	0	1	0	0	1
LGR	MUR	-1.243637	38.033028	7	7	0	0	7	0
LGR	NAV1	-1.640235	41.974884	10	1	9	0	0	10
LGR	NAV1	-1.482148	42.187229	3	0	3	0	0	3
LGR	NAV1	-1.442261	42.378651	7	0	7	0	0	7
LGR	NAV1	-1.573341	42.319778	3	1	2	0	0	3
LGR	NAV1	-2.082455	42.476921	25	1	24	0	25	0
LGR	NAV1	-1.258404	41.922241	10	0	8	2	0	10
LGR	NAV1	-1.612827	42.064831	1	0	1	0	0	1
LGR	NAV1	-1.653248	42.119507	6	0	6	0	0	6
LGR	NAV1	-1.798200	42.101800	17	11	6	0	0	17
LGR	NAV1	-1.765864	42.277707	28	5	23	0	0	28
LGR	NAV1	-1.490363	41.999619	22	1	20	1	0	22
LGR	PAN	-8.901123	38.793315	29	29	0	0	29	0
LGR	PTM	-8.515796	37.169695	13	13	0	0	13	0
LGR	POL	-6.611051	43.271709	1	0	1	0	0	1
LGR	SAL	-5.184983	41.124939	8	6	2	0	8	0
LGR	SAL	-5.260701	41.047848	2	2	0	0	2	0
LGR	SAL	-6.359504	40.801659	1	1	0	0	1	0
LGR	SAL	-5.897394	41.206612	1	1	0	0	1	0
LGR	SAL	-6.359352	40.522388	1	0	1	0	1	0
LGR	SAN	-8.685994	39.236668	16	16	0	0	16	0
LGR	SEN	-6.271231	41.499542	1	0	1	0	1	0
LGR	SEN	-6.503643	41.307781	1	1	0	0	1	0
LGR	SEN	-6.808799	41.090080	3	1	2	0	3	0
LGR	SEN	-6.431105	41.415914	16	15	1	0	16	0
LGR	SES	-7.625556	40.298332	3	3	0	0	3	0
LGR	SES	-7.266131	40.538349	1	1	0	0	1	0
LGR	SES	-7.089524	40.351662	6	4	2	0	6	0
LGR	SES	-7.392667	40.514511	2	2	0	0	2	0
LGR	SEV	-5.764025	37.840450	1	1	0	0	1	0
LGR	SEV	-5.344389	37.711113	7	7	0	0	7	0
LGR	SOR	-2.463772	41.764431	1	0	1	0	1	0
LGR	SOR	-2.272888	41.308670	2	0	2	0	2	0
LGR	SOR	-2.365938	41.552589	1	0	1	0	1	0
LGR	SOR	-2.599558	41.474736	1	0	1	0	1	0
LGR	SOR	-2.458889	41.830002	1	0	1	0	1	0
LGR	SOR	-2.743580	41.634573	8	4	4	0	8	0
LGR	SOR	-2.457360	41.242658	4	0	4	0	4	0
LGR	SOR	-2.095995	41.648623	6	2	4	0	6	0
LGR	SOR	-2.392854	41.923808	2	0	2	0	2	0
LGR	TAR	0.480001	40.544415	3	0	3	0	0	3
LGR	TAR	0.373882	40.734818	10	1	9	0	0	10
LGR	TCA	-4.976654	42.049622	13	6	7	0	13	0
LGR	TCA	-4.780810	41.915787	1	1	0	0	1	0
LGR	TCA	-4.969508	41.855160	8	2	6	0	8	0

LGR	TER	-1.103675	40.341839	1	0	1	0	1	0
LGR	TOL	-4.027323	39.862831	26	19	7	0	26	0
LGR	TOR	-4.995759	41.504719	22	17	5	0	22	0
LGR	VAL	-0.376288	39.469910	5	4	1	0	5	0
LGR	VLP	-7.329707	41.569266	49	18	31	0	0	49
LGR	VLP	-7.431022	41.439006	1	1	0	0	0	1
LGR	VLP	-7.417325	41.675739	1	1	0	0	0	1
LGR	VLP	-7.467154	41.571795	22	21	1	0	0	22
LGR	VLP	-7.291761	41.706713	8	7	1	0	0	8
LGR	ZAM	-5.745313	41.507324	1	0	1	0	1	0
LGR	ZAM	-5.557500	41.351877	14	14	0	0	14	0
LGR	ZAR	-0.889581	41.648792	1	0	1	0	1	0
LGR	ZAR	-1.244587	41.438297	7	2	5	0	7	0
LGR	ZAR	-0.875759	41.257969	10	0	10	0	10	0
LGR	ZAR	-1.192911	42.249786	1	1	0	0	1	0
LGR	ZAR	-1.616740	41.395985	1	1	0	0	1	0
LGR	ZAR	-1.144898	41.385502	5	0	5	0	5	0
LGR	ZAR	-0.526964	41.488827	14	0	11	3	11	3
LER	ALA2	-2.860516	42.727879	1	0	1	0	1	0
LER	ALA2	-2.352369	42.669483	2	0	0	2	2	0
LER	ALA2	-2.992245	42.846432	1	0	1	0	1	0
LER	ALA2	-2.918374	42.712128	1	0	1	0	1	0
LER	ALA2	-3.098279	42.848080	1	0	1	0	1	0
LER	ALA2	-2.581248	42.882163	61	17	44	0	61	0
LER	ALA2	-2.583586	42.625389	11	11	0	0	11	0
LER	BAR	1.975309	42.244564	4	0	0	4	0	4
LER	BAR	1.818061	42.203552	4	0	0	4	0	4
LER	BAR	1.872023	42.080879	8	0	0	8	0	8
LER	CAN	-3.983120	43.183715	15	0	10	5	15	0
LGR	CAN	-3.921429	43.394523	1	0	1	0	0	1
LER	GER	2.103915	42.316177	1	0	0	1	0	1
LER	GER	2.874128	41.889561	5	0	0	5	0	5
LER	GER	3.056821	42.197102	4	1	0	3	0	4
LER	GER	2.365142	42.313667	4	0	0	4	0	4
LER	HUE	0.035862	42.230381	1	0	0	1	0	1
LER	JAC	-0.547055	42.571716	3	0	3	0	3	0
LER	JAC	-0.983557	42.586388	3	0	3	0	3	0
LER	JAC	-0.677305	42.627350	1	0	1	0	1	0
LER	JAC	-0.889167	42.641388	5	0	4	1	5	0
LER	LRD	0.899247	41.575160	8	0	0	8	0	8
LER	NAV2	-1.370966	42.786083	3	0	3	0	3	0
LER	NAV2	-1.645104	42.733093	1	0	1	0	1	0
LER	NAV2	-2.170814	42.895634	1	0	1	0	1	0
LER	NAV2	-1.002983	42.720951	1	0	0	1	1	0
LER	VLC	-3.572031	42.938454	28	0	28	0	28	0