Maroso F., Vera M., Ferreiro J., Mayol M., Riba M., Ramil-Rego P., Martínez P. & Bouza C. 2021. Genetic diversity and structure of *Taxus baccata* from the Cantabrian-Atlantic area in northern Spain: A guide for conservation and management actions. **Forest Ecology and Management**, 482: 118844. (doi: 10.1016/j.foreco.2020.118844)

Genetic diversity and structure of *Taxus baccata* from the Cantabrian-Atlantic Area in Northern Spain: a guide for conservation and management actions

Maroso F.[1,2], Vera M.[1], Ferreiro J.[3], Mayol M.[4], Riba M.[4, 5], Ramil-Rego P.[3], Martínez P.[1], Bouza C.[1*]

[1] Departamento de Zooloxía, Xenética e Antropoloxía Física. Facultade de Veterinaria. Universidade de Santiago de Compostela. Campus de Lugo. 27002 Lugo, Spain

[2] Department of Life Science and Biotecnology, University of Ferrara, via Borsari 46, 44121 Ferrara, Italy

[3] IBADER (Instituto de Biodiversidade Agraria e Desenvolvemento Rural), Universidade de Santiago de Compostela. Campus de Lugo, 27002 Lugo, Spain

[4] CREAF, Cerdanyola del Vallès 08193, Spain

[5] Universitat Autònoma Barcelona, Cerdanyola del Vallès 08193, Spain

* Corresponding author

Carmen Bouza. Departamento de Zooloxía, Xenética e Antropoloxía Física. Facultade de Veterinaria. Universidade de Santiago de Compostela. Campus de Lugo. 27002 Lugo, Spain

TLF +34 982 82 24 28

Fax +34 982 82 20 01

e-mail: mcarmen.bouza@usc.es

Highlights

Iberian yew showed high genetic diversity in the Cantabrian-Atlantic region (CR)

Biogeographic genetic differentiation was observed among CR groups.

Local genetic structure and relatedness emerged at a fine spatial scale in CR.

N-to-S genetic diversity decreasing and E-W differentiation was observed in Iberia

Guidelines are provided to assist conservation of yew genetic resources

22 Keywords: conservation, genetic diversity, genetic differentiation, microsatellite markers,
23 palaeartic yew, *Taxus baccata*
24

Abstract

The maintenance of biological diversity at ecosystem, species and intraspecific levels is essential to ensure the survival of forests. The palaeartic yew *Taxus baccata* is part of a number of threatened forest types with prioritized conservation value in Europe. We investigated the spatial distribution of microsatellite genetic diversity in forty-six *T. baccata* populations (1,054 trees) spanning the Cantabrian-Atlantic Region (CR) in Northern Spain, framed within a conservation and restoration plan of the species in that region. Different layers of genetic structure were detected, with low structure at a global scale, suggesting historical connectivity, and a complex structure at smaller spatial scales. A low but significant regional genetic variation was also identified associated with biogeographical groups within CR, of potential interest to assist conservation and restoration programs. These genetic differences were reflected on a heterogeneous contribution to the total heterozygosity and allelic richness by the different regions. Data were contextualized within the Iberian Peninsula using previous data in this species (totaling 2,731 trees from 128 populations) after genotype standardization for a common set of seven microsatellites, confirming higher genetic diversity and more homogeneous structure in CR than in central and southeastern Iberian regions. Evidence of geographical structure between eastern (Betic Range, Catalonian Ranges, Pyrenees) and western (Cantabrian Range, Central System Range, Iberian System Range) populations was detected. This study deepens into the spatial distribution of genetic diversity in *T. baccata* through an intensive survey in CR as a basis for different *in situ* and *ex situ* conservation actions in the region aimed to conserve the genetic resources of this species and improve protected yew-associated natural habitats.

1. Introduction

Changes in biological diversity have been driven by historical climate fluctuations and more recently also by the effects of human activities (Ferreiro da Costa and Ramil-Rego, 2018; Birks, 2019). Concerns on biodiversity loss in forest ecosystems are currently increasing, since there is mounting evidence that the provision of ecosystem services is related to several aspects of biodiversity (Brockerhoff et al., 2017 and references therein). Ecosystems and plant communities dominated by highly competitive species with low colonization rates, many of them found in temperate forests, might be particularly sensitive to biodiversity loss following habitat destruction (Tilman et al., 1994). On the other hand, global analyses of the interacting effects of habitat loss, current climate and climate fluctuations on terrestrial populations during the last decades, suggest strong synergistic effects, the negative effects of habitat loss and fragmentation on biodiversity being greater in areas with high maximum temperatures, especially those that are susceptible to precipitation change (Mantyka-Pringle et al., 2012). This suggests that efforts on management strategies for conserving biodiversity should focus towards those vegetation or habitat types located in the warmest areas within its climatic range.

The main human-related causes behind plant diversity reduction in temperate forests include forestry, agriculture and farming practices, along with introduction of alien species, pollution and climate change, that have led to overexploitation and habitat fragmentation or loss (Millennium Ecosystem Assessment, 2005). Ensuring the conservation and sustainable management of forests is a priority measure (EU, 2011; CBD, 2014). Biodiversity monitoring, from ecosystems and species to intraspecific variation is essential to support conservation actions (Allendorf et al., 2013), such as mitigating threatening factors, improving networks of protected areas and increasing restoration of ecosystems and vulnerable species. In this context, the genetic analyses within species at global and local scales are the basis both to

74 develop and assess strategies for sustainable preservation of forest tree resources (de Vries et
75 al,. 2015; Aravanopoulos, 2016; Fady et al., 2016; Fussi et al., 2016).

76 European yew (*Taxus baccata*) is a long-lived dioecious species, whose distribution ranges
77 from North Africa to Scandinavia and from the Iberian Peninsula to the Caspian Sea (Thomas
78 and Polwart, 2003), and it is currently found in a variety of mixed hardwood forests in
79 temperate European regions. The ecological importance of this species and its association
80 with prioritized habitats has been recognized for wildlife conservation within Natura 2000
81 network (EEA, 2015; Thomas and García-Martí 2015). Despite growing under a wide range
82 of environmental conditions, from oceanic to continental and Mediterranean climate (Thomas
83 and Polwart, 2003) suitable areas for the maintenance of *Taxus baccata* are rapidly decreasing
84 in Europe. Thus, European yew is considered locally rare or endangered in many European
85 regions, including several habitat types in its current southernmost European range (Thomas
86 and Polwart, 2003; Linares, 2013). Besides the progressive climatic constraints, the species
87 decline in Europe seems to be mediated by long-term human pressures including excessive
88 felling, fires, and low rates of natural regeneration due to grazing (Linares, 2013; Thomas and
89 García-Martí, 2015; Litkowiec et al., 2018). In addition, competition for light under excessive
90 canopy closure strongly limits the ability for growth and reproduction (Svenning and Magård
91 1999; Iszkuło, 2010; Iszkuło et al., 2012).

92 The low current census and fragmentation of most yew populations increase the risk of
93 genetic drift and inbreeding. In fact, many studies dealing with genetic diversity and structure
94 of yew in Europe have revealed strong genetic structuring (e.g. Lewandowski et al., 1995;
95 Cao et al., 2004; Myking et al., 2009; Chybicki et al., 2012; Litkowiec et al., 2018),
96 particularly high in southern European areas (Dubreuil et al., 2010; González-Martínez et al.,
97 2010; Trober and Ballian, 2011; Mayol et al., 2015). In the Iberian Peninsula, a cline of
98 increasing genetic variation from southeast to northwest and higher structuring in south-

99 eastern areas has been reported to be driven by glacial climate cycles and human influence
100 (González-Martínez et al., 2010; Mayol et al., 2015). In particular, the Cantabrian Range,
101 which runs parallel to the Cantabrian Sea in the north of the Iberian Peninsula, was found to
102 be one of the most diverse and least structured areas (González-Martínez et al., 2010).
103 However, the limited number of populations analyzed precludes extracting general
104 conclusions, in particular concerning management actions at a regional and local scale in this
105 area. Studies on vegetation dynamics based on biogeographical and paleoecological
106 information (Ramil-Rego et al. 1998a, b; Muñoz-Sobrino et al. 2018) suggest strong spatial
107 and temporal heterogeneity within this this area, which could have influenced the genetic
108 structure of yew. The conservation value of *T. baccata* associated-forests in that region has
109 promoted recovery programs to prevent further decline of natural populations of the species
110 (http://www.life-baccata.eu/es). It is thus pertinent to expand previous genetic studies to a
111 higher number of populations in order to obtain key information to support conservation and
112 management strategies, as reported in other European regions (Litkowiec et al. 2018;
113 Gargiulo et al. 2019).

114 In this study, microsatellite markers were used to investigate the genetic diversity and
115 structure of *T. baccata* populations in the Cantabrian-Atlantic Region (CR) of northern Iberia
116 at a much more detailed geographical scale than previously reported. Genetic data across
117 biogeographical and paleoecological regions within CR were used to delineate areas for
118 conservation and evaluate conservation units for maximizing the maintenance of existing
119 genetic variation. In addition, we integrated the new genetic data with the previous results
120 reported by González-Martínez et al. (2010) and Mayol et al. (2015) using a common set of
121 microsatellites in order to re-evaluate the global patterns of genetic diversity of European yew
122 in the Iberian Peninsula. The study was carried out within the framework of the project LIFE-

123 BACCATA (LIFE15 NAT/ES/000790), aimed at fostering conservation strategies based on
124 genetic data of the most important yew populations in this area.

125 **2. Materials and methods**

126 2.1. Sampling and microsatellite analysis

127 A total of 1,054 new samples were collected from 46 populations in the Cantabrian-Atlantic
128 Area (N = 3 - 50, average 22.9 per population; Figure 1, Table 1). These populations were part
129 of 17 Special Areas for Conservation (SAC) of biodiversity and habitats targeted under the
130 Habitats Directive from the Natura 2000 Network (https://natura2000.eea.europa.eu/; Figure
131 1; Table 1). EUFORGEN's (European Forest Genetic Resources Programme) Genetic
132 Conservation Units (GCU) previously delineated for *Taxus baccata* in the CR and the Iberian
133 Peninsula (http://www.euforgen.org/species/taxus-baccata/) were also integrated with the
134 sampling sites in this study (Figure 1). Our sampling survey encompassed 300 km, from the
135 western limit of the Pyrenees to the northwest Atlantic edge along the coast of the Cantabrian
136 Sea (Figure 1), and included different biogeographical territories, i.e. the Cantabrian Range,
137 the Galician Inner Depression, Navian-Ancares Mountains, Galaico-Duriense Massif and the
138 Cantabrian-Basque Mountains (see below for details on regional groupings). For simplicity
139 we shall refer the area under study to the Cantabrian-Atlantic Region (CR) throughout the text
140 in accordance with biogeographical data in the Iberian Peninsula (Rivas-Martínez et al.,
141 2014). Forest and land planning policies in CR are run by different Spanish Autonomous
142 Communities (Asturias, Cantabria, Galicia, Castilla y León and País Vasco, the last three split
143 into different provinces; Figure 1; Table 1), which might be important in terms of intra- and
144 inter-regional management coordination (Montiel and Galiana, 2005). Population codes
145 reflect administrative information about the presence of different management areas across
146 biogeographic ranges of yew distribution within CR including Autonomous communities and
147 provinces; See Table 1 and Figure 1). Sites for sampling were selected using field surveys and

148 administrative information on the contemporary distribution of yew populations in the area,
149 but also considering environmental, biogeographic and historical scenarios, all of them
150 potentially contributing to shape the current patterns of genetic variation, as reported for yew
151 and other temperate tree species (Mayol et al., 2015, 2020; Muñoz-Sobrino et al., 2018). Field
152 surveys were designed to collect a representative sample size for each population, capturing
153 the geographic position of each individual tree by global positioning system (GPS). Following
154 previous sampling schemes (González-Martínez et al., 2010), a separation > 10 m was taken
155 to minimize the risk of clonality between adjacent tree samples (González-Martínez et al.,
156 2010). Genomic DNA was isolated from 50 mg of fresh leaf samples of each tree, using
157 E.Z.N.A Plant DNA DS Mini Kit (Omega). Nine previously published microsatellites were
158 analysed: Tax23, Tax26, Tax31, Tax36, Tax60, Tax86, Tax92, TS09 (Dubreuil et al., 2008)
159 and ABRII-TB1 (Mahmoodi et al., 2010). Two multiplex PCR reactions were tuned up to
160 amplify seven microsatellites: PCR 1 (Tax23, Tax36, Tax92) and PCR 2 (Tax26, TS09, Tax86,
161 ABRII-TB1) in a final volume of 10 µl including 30 ng of DNA, 1X Master Mix kit Go Taq
162 G2 Hot Start Colorless Master kit (Promega) and 1 µM of each primer, except 0.75 µM for
163 Tax92. For the two remaining microsatellites (Tax31 and Tax60), singleplex PCRs were
164 carried out in 15 µl reaction with 30 ng of template DNA, 1× Gold Buffer, 1.5 mM MgCl2,
165 0.1 mM of each dNTP, 0.3 µM of each primer, 0.5 U AmpliTaq Gold® DNA Polymerase
166 (Applied Biosystems). PCR conditions included an initial denaturation at 95ºC for 2 min,
167 followed by 30 cycles (28 for Tax60 and Tax31) of 94ºC for 30s, annealing at 56ºC (except
168 Tax31 at 61ºC) for 90s and 72ºC for 1 min, and a final extension at 72ºC for 5 min. PCR
169 products were multiplexed (PCR1 with Tax60; PCR2 with Tax92) to be analysed in an ABI
170 PRISM® 3730xl automatic sequencer (Applied Biosystems) at University of Santiago de
171 Compostela (USC). Allele scoring was performed using GeneMapper 4.0 (Applied
172 Biosystems).

173 2.2. Genetic diversity and structure in the Cantabrian-Atlantic Region (CR)

174 Genetic diversity in each population was estimated by expected (He) and observed (Ho)
175 heterozygosity using GenAlEx 6.502 (Peakall and Smouse, 2006) and by allelic richness (AR;
176 minimum sample size N=15) using Contrib 1.02 (Petit et al., 1998). Hardy-Weinberg
177 equilibrium was checked using exact tests for each locus and population, and a global Fisher
178 test was applied over loci within each population using Genepop 4.0 (Raymond and Rousset,
179 1995). This program was used to estimate the intra-population fixation index (F_{IS}) and its
180 significance. F_{IS} was also computed using INEST 2.2 (Chybicki and Burczyk, 2008; $F_{IS\text{-}INEST}$)
181 to correct for the presence of null alleles, as previously reported for these microsatellite loci
182 (Dubreuil et al., 2008; González-Martínez et al., 2010; Chybicki et al., 2011). Null allele
183 frequency was estimated with Genepop. Linkage disequilibrium (LD) between pairs of loci
184 within populations was calculated with Plink 2.0 (Purcell et al., 2007) using an exact test over
185 genotypic frequencies and the global estimation across all populations using Fisher tests.

186 Population samples with less than 15 trees (eight populations totalling 47 individuals; Table 1)
187 were removed for pairwise F_{ST} estimation. Pairwise F_{ST} and AMOVA were performed using
188 Arlequin 3.5 (Excoffier and Lischer, 2010) and the significance calculated with 5,000
189 iterations. Isolation by distance (IBD) of populations was tested by a correlation between F_{ST}
190 and geographical distance using Mantel test with R's package ade4 (http://www.R-
191 project.org/). Clustering analysis was performed under admixture model with STRUCTURE
192 2.3.4 (Pritchard et al., 2000). The whole CR dataset ("global CR" analysis) was analysed for k
193 ranging from 1 to 25 ancestral groups with a priori information of populations to help
194 inferring a more refined population structure, with each k repeated 10 times (100,000 burn-in
195 cycles (BI) and 500,000 iterations). All STRUCTURE results in this study were parsed using
196 STRUCTURE HARVESTER (Earl and vonHoldt, 2012), as an exploratory tool across
197 different k values, to detect the most likely number of groups according to Evanno's Δk

198 values (Evanno et al., 2005). CLUMPAK 1.1 (Kopelman et al., 2015) was used to summarize
199 and evaluate the precision of clustering at different k, and the results plotted using *pophelper*
200 online tool.

201 Genetic data integrated on environmental zoning is essential to evaluate forest genetic
202 resources (Lefèvre et al. 2020), and to identify operational conservation units. This might be
203 particularly relevant for guiding conservation policies when different regional administrative
204 agencies are responsible for conservation and forest management. We therefore further
205 analysed the CR dataset considering different groupings based on biogeographical and
206 paleoecological factors, which might have influenced regional patterns of genetic structure of
207 yew (e.g., Mayol et al., 2015, 2020). We considered evidence from previous reports regarding
208 climate heterogeneity, bedrock lithology and mountain relief, as well as the differential
209 regional impact of glaciarism during the last ice age driving environmental conditions and
210 vegetation ecosystems (Ramil-Rego et al., 1998a, 1998b; López de Heredia et al., 2007;
211 Muñoz-Sobrino et al., 2018). Three regional groupings, using different criteria, were explored
212 (I-III; Table S1): I) paleoecological-based spatial differentiation (Galician Inner Depression,
213 Galaico-Duriense Range, Navian-Ancarense Mountains, Cantabrian Range, Cantabrian-
214 Atlantic Sublitoral Valleys, Basque-Cantabrian Mountains) (Muñoz-Sobrino et al. 2009,
215 2018); II) biogeographical territories associated with the main mountain ranges in the region,
216 including a spatial subdivision within the Cantabrian Range (West, North and South
217 areas)(Muñoz-Sobrino et al. 2018; Rivas-Martínez et al., 2014); III) the same as II, but
218 excluding the most differentiated and isolated populations within less represented
219 biogeographical regions (i.e., populations OU001, ZA001 and LU010; see Results). The
220 relative importance of these regional groupings was evaluated using AMOVA (Arlequin). ,.

221 In addition, STRUCTURE 2.3.4 was run separately within each regional group of populations
222 in CR ("regional CR" analysis), under admixture model and without *a priori* population

information to explore individual assignments irrespective of the sampling location of origin, with *k* ranging from 1 to 5, each k repeated 10 times with 50,000 BI cycles and 100,000 repetitions. Results were parsed using STRUCTURE HARVESTER, to obtain likelihood values for the k values explored, according to Evanno's Δk method; and CLUMPAK 1.1, to summarize clustering resultsacross different runs. Finally, the online tool *pophelper* (Francis, 2017) was used to draw STRUCTURE plots and attach population and group labels.

Contrib 1.02 was applied to assess the contribution of each population to the total genetic diversity, excluding the smallest samples (N < 15). Genetic diversity estimators (He, AR) were partitioned into the intra- and inter-population contributions of each population to the total diversity. The contribution analysis of each regional cluster ("conservation unit") identified in CR was also performed using Contrib 1.02, using the whole population dataset. Average and standard deviation of He, AR and F_{ST} were calculated for each biogeographical group within CR and plotted using R in order to compare genetic diversity between groups (Wilcoxon tests). Effective population size (Ne) was calculated with the LD method implemented in NeEstimator 2.1 (Do et al., 2014). Upper and lower confidence intervals were obtained along with the Ne estimated value per population, since the range was usually very wide and could lead to meaningless biological conclusions. The correlation between Ne and AR at population level was performed using Spearman's Rho test in R. The presence of private alleles and locally common alleles were explored. Following Maguire et al. (2002), locally common alleles were defined as those occurring at >20% frequency at specific locations and < 10% in the other populations. Finally, pairwise relatedness (*r*) was calculated between all pairs of individuals within each population (except N < 15) using the Wang (2001) estimator implemented in Spagedi 1.5 (Hardy and Vekemans, 2002), as the mean over loci after Jackknife re-sampling. Correlation tests and plots between genetic estimators (F_{ST} and *r*) and geographical distances were carried out using Mantel test with R's package ade4.

248 2.3. Genetic diversity in the Iberian Peninsula: genotype standardization with previous studies

249 To integrate genetic data obtained in this study (CR) with those previously reported for the
250 Iberian Peninsula (González-Martínez et al., 2010; Mayol et al., 2015), we standardized
251 genotypes of both datasets. Twenty-four previously genotyped DNA samples by Mayol et al.
252 (2015) were re-genotyped at the USC facilities with the seven microsatellite loci (Tax23,
253 Tax26, Tax31, Tax36, Tax60, Tax86 and Tax92) used by González-Martínez et al. (2010) and
254 Mayol et al. (2015). For each locus, alleles covering the full range were compared and their
255 correspondence established. All data were then merged in a unique dataset for the Iberian
256 Peninsula consisting of a total of 2,791 samples from 126 populations (WIS: Whole Iberian
257 Peninsula Sample; Table S1). To explore the genetic structure of the WIS dataset, pairwise F_{ST}
258 values were estimated with Arlequin 3.5 and their significance calculated with 5,000
259 iterations. STRUCTURE for the complete WIS dataset ("global Iberian" analysis) was run
260 under admixture model and without *a priori* information of populations, for *k* ranging initially
261 from 1 to 15, then every 5 k up to 50 and then every 20 k up to 150, a range that covered the
262 total number of populations studied (126). The analysis was performed using 100,000 BI
263 cycles and 500,000 repetitions. Then the total dataset was split into six groups according to
264 the biogeographical regions considered by González-Martínez et al. (2010) in the Iberian
265 Peninsula: Betic, Catalonian, Central System, Iberian System, Cantabrian and Pyrenees
266 Ranges (Table S1). Average and standard deviation of Ho, He, AR and F_{ST} were calculated for
267 each biogeographical group and plotted using R in order to compare genetic diversity between
268 groups (Wilcoxon tests). A STRUCTURE analysis under the admixture model was also
269 performed for each biogeographic group in the whole Iberian dataset ("regional Iberian"
270 analysis), without *a priori* information, with k ranging from 1 to 10 with 10 replicates each
271 (burn in of 80,000; 100,000 repetitions). The STRUCTURE results obtained ("global" and
272 "regional Iberian" analyses) were parsed separately using STRUCTURE HARVESTER to

analyze the k values explored, according to Evanno's Δk method, summarized using CLUMPAK and plotted using *pophelper* online tool.

3. Results

3.1. Genetic diversity and structure in the Cantabrian Region (CR)

Genetic diversity for 46 populations (1054 individuals) sampled across CR are shown in Table 2 and Table S1. Linkage disequilibrium was not significant for any pair of loci within each population and across all populations (results not shown). Average null allele frequency across loci and populations was 0.055, Tax23 and ABRII-TB1 showing the highest (0.130) and lowest (0.010) values, respectively. After correction for null alleles, population average was 0.664 for He (range: 0.413–0.770), 0.599 for Ho (0.453– 0.759), 4.231 for AR (1.722 - 6.133) and 0.036 for $F_{IS\text{-}INEST}$ (0.009 - 0.121) (Table 2; Table S1). Eight populations showed significant $F_{IS\text{-}INEST}$: LU07, LE15, AS06, CA03, LE045, AR001, GI001, GI002 (minimum sample size of 15). $F_{IS\text{-}INEST}$ were lower than classic F_{IS} values (average: 0.092; range: -0.200 - 0.274; Table S1), suggesting that part of the heterozygote deficit detected was related to null alleles. Estimates of effective population size (Ne) were in general low (very small for LU010, LU013 and LE014; Table S1). Ne ranged from 5.4 (LU013) to 256.8 (LE292), and narrow confidence intervals (CI) were usually observed around the estimated value in most samples. Only in six populations with more than 15 individuals (OU1, LE011, LE017, LE1000, LE292 and LE023) the upper limit could not be estimated ("Infinite" values). Significant correlation between Ne and AR at population level was detected ($r = 0.414$, $P = 0.011$).

Pairwise F_{ST} values ranged between 0.011 (LE023 *vs*. GI002) to 0.480 (LU010 *vs*. OU001) with an average of 0.112 (Table S2A). Average relatedness (*r*) among populations was low (-0.055; range: -0.183 – 0.086), while, as expected, intra-population average was much higher

297 (0.131; range -0.082 – 0.481) (Table S2B). The presence of clones among the sampled
298 individuals was supported by pairwise $r = 1$ in LU013 (27%), LE003 (3.3%), LE014 (10.5%),
299 LE1000 (5.6%), LE046 (6.5%) and CA001 (9%) (Table S1). All these cases involved short
300 distances between trees, close to the 10 m-limit established for sampling, suggesting that it
301 could be increased to avoid clonality for small populations in future studies. Although they
302 represent a very small fraction in the whole dataset (1.2%), these individuals were removed
303 for genetic diversity analyses. The pattern of differentiation due to isolation by distance (IBD)
304 was significant ($P < 0.001$), but mainly explained by two highly differentiated populations in
305 the westernmost part of the distribution (LU010 and OU001, Figure S1). When these
306 populations were removed from the analysis, non-significant correlation between F_{ST} and
307 geographic distance was observed ($P = 0.141$). Nevertheless, the average pairwise relatedness
308 between individuals of different populations showed a negative correlation with distance and
309 it remained significant even after removing OU001 and LU010 ($P < 0.001$).

310 The contribution of each population to the total diversity estimated by heterozygosity and
311 allelic richness was in general low but quite heterogeneous (Table 2; Figure S2). The highest
312 contribution (1.55%) to AR differentiation was estimated for LU010, followed by PA39
313 (1.47%) and OU01 (1.36%). However, the maximum individual contribution to the total AR
314 diversity (inter- plus within-population contributions) was in PA039 (1.1%, Figure S2).

315 In the "global CR" STRUCTURE analysis, and according to Evanno's Delta k values (Evanno
316 et al., 2005), the two most likely number of clusters in CR ($k = 2$ and 4; Figure S3) were
317 explored. The best *k* value revealed two clusters distributed across CR, showing diverse
318 ancestry proportions between many of the western and central/eastern populations ("green"
319 and "red" component predominance, respectively; Figure 1). The results for $k = 4$, as the
320 second highest Δk peak, provided some evidence of regional clustering of populations
321 according to the different biogeographical regions (Figure 1). Differential cluster composition

was observed in the eastern (Basque-Cantabrian Mountains, BCM) and western (Navian-Ancares Mountains, NAM, and Western Cantabrian Range, WCR) areas of CR ("red" and "green" clusters, respectively, in Figure 1). Most populations showed some degree of admixture, and only a few western populations were made up of a single group (e.g., LE001). Some admixture populations with an "orange" component followed a SW to NE gradient (from LE001 to LE005). In central areas of CR, some differences between southern (SCR) and northern (NCR) populations were observed, with a predominant "blue" component in the former (except LE050) and higher admixture in the latter ("blue", "green" and "red" components). Some exceptions were detected, like the Galaico-Duriense Range (GDR), where OU001 with a predominant "orange" cluster component differed from the close population ZA001, admixed with prevalence of the "red" and "blue" components, also observed in LU10 located in the Galician Interior Depression, GID (Figure 1).

As the number of k increased, different minor Δk peaks were detected (k = 16, 6, 21, 10 and 12, in decreasing Δk order), with the main plateau in the probability plot reached at k = 16 (Figure S3). The exploration across the higher k values showed a similar pattern with many populations or small groups of nearby populations characterized by consistent specific genetic patterns , but with varyingl proportion assigned to each population/group (Figure S3), suggesting that a layer of local spatial structure may also exist. Increasing admixture was observed in some populations from the Cantabrian Range, whereas the most homogeneous clusters matched with the least variable and/or highly differentiated populations across CR (e.g., LU010; OU001, PA039, and BU035) (Figure 1; Table 2), which could reflect the tendency of clustering methods to assign pure ancestry components to populations with a strong impact of genetic drift.

We further analyzed different regional groupings in CR (Table S1) using AMOVA (Tables S3). The genetic diversity among-groups was always below 2%, ranging between 1.12% for

Grouping III (when excluding the more isolated populations in the westernmost regions) to 1.93% for Grouping II (according to biogeographical territories), but significant in all cases ($P<0.001$). Grouping II provided evidence of differentiation among seven biogeographical areas sampled across CR (GID, GDR, NAM, NCR, WCR, SCR, BCM; Figure 1). Genetic diversity analyses within and among these biogeographical groups (see Figure 1) showed no significant differences in heterozygosity, and that the NCR populations presented on average the highest allelic richness and lowest pairwise F_{ST} within group (Figure 2). In the "regional CR" STRUCTURE analyses within each of the biogeographical regions, the best number of clusters for NCR was k = 4 with high levels of admixture (Figure S4). In comparison, the other groups showed lower *k* values, except NAM, and none or little admixture. Moreover, BCM showed the highest number of private alleles (29), followed by NCR (15) and BAN (10), which can be partly related to greater sampling size in these areas (Table 4). Private alleles showed in general low frequencies and were abundant in most groups excluding less represented GID and GDR. A few locally common alleles were found, five in GID (LU10 population) and two in GDR (Table 4). The regional contribution to genetic diversity of these two groups was not considered due to their low sampling representation (Figure 2) but it was calculated at population level (Table 2; Figure S2). In the regional analysis BCM was the most contributing group to the global genetic diversity, both to He and AR (Figure 2). All regions contributed to AR differentiation, BCM being the highest differentiated from all others as well as WCR, although this region was less diverse.

3.2. Integration of new samples at the whole Iberian Peninsula scale

The full allelic ranges per locus were compared with reference samples from previous studies, allowing genotyping standardization for seven microsatellite loci. Only Tax60 showed larger fragment length differences between this and previous studies (about 5-10 bp), but population analyses were not affected by the inclusion of this marker (data not shown). As for the

372 Cantabrian Range, a wide variation of within-population genetic diversity was observed for
373 most of the regions analysed in the Iberian Peninsula (Figure 3; Table S1). Expected
374 heterozygosity (He) averaged 0.613 across all (126) locations, and ranged from 0.214
375 (Navarra and Salamanca) to 0.770 (CA001), a similar range to that observed in CR. After
376 excluding populations with small sample size, AR averaged 3.770 across 85 locations,
377 ranging from 1.427 in LU10 to 5.847 in GI002, two CR populations. Wide ranges were also
378 observed in the other regions (Table S1): Betic Range (He: 0.393 - 0.694; AR: 2.347 - 3.994);
379 Catalonian Ranges (He: 0.429 - 0.728; AR: 1.835 - 4.591); Central System Range (He: 0.214 -
380 0.717; AR: 2.236 - 4.703); Iberian System Range (He: 0.360 - 0.628; AR: 3.100- 3.349), and
381 Pyrenees (He: 0.411 - 0.746; AR: 3.024 - 5.044). According to previous results, average
382 genetic diversity was higher in northern (mean He/AR: Cantabrian Range, 0.637/3.999;
383 Pyrenees, 0.629/4.030) than in central, southeastern and Mediterranean regions (Iberian
384 System range, 0.516/3.175; Betic range, 0.595/3.258; Central System range, 0.601/3.594;
385 Catalonian ranges, 0.617/3.413). Analyses of correlations (Spearman's Rho) between genetic
386 variability, latitude and longitude, also showed some significant trends: all three genetic
387 diversity parameters increased with latitude (AR: $r = 0.51$, $P < 0.001$; He: $r = 0.40$, $P < 0.001$;
388 Ho: $r = 0.53$, $P < 0.001$) while decreased with longitude (AR: $r = -0.20$, $P = 0.06$; He : $r = -$
389 0.19, $P < 0.05$; Ho: $r = -0.31$, $P < 0.001$).

390 Pairwise F_{ST} values between all populations averaged 0.129 and ranged from low (LE02 vs.
391 GI002: 0.011) and not significant, to high and significant (OU001 vs. LU010: 0.464),
392 especially for pairwise comparisons involving some particularly differentiated populations
393 (Font Negra, Vidalbar, Hurdes, OU001, GI002, among others; Table S4). Average pairwise
394 F_{ST} values within biogeographic regions showed lower values in CR and the Pyrenees than in
395 the other regions (Figure 3). According to Evanno's method (2005) in the "global Iberian"
396 STRUCTURE analysis, the most likely number of clusters in the Iberian Peninsula was k = 2,

which suggested a pattern separating eastern (Betic Range, Catalonian Ranges, Pyrenees) and western (Cantabrian Range, Central System Range, Iberian System Range) samples (Figure 3).

The "regional Iberian" STRUCTURE analysis suggested different patterns of structure within each regional group of the Iberian Peninsula (Figure S6). The most likely number of genetic groups for the Central System and Betic Ranges (k = 3 in both cases) was higher than for the Catalonian, the Iberian System, the Pyrenees and the Cantabrian Ranges (k = 2), the latter being concordant with the "global CR analysis" (Figure 1). Some local substructure was detected when higher k values were explored in each of the regions studied (Figure S6). The Central and Betic groups, together with the Catalonian Ranges, showed higher average pairwise F_{ST} values than Cantanbrian Range and the Pyrenees (Figure 3; Table S4).

4. Discussion

In this study, microsatellite markers were used to assess genetic diversity in European yew (*T. baccata*) from the Cantabrian-Atlantic Region (CR), previously suggested as one of the main diverse areas within the Iberian Peninsula (González-Martínez et al., 2010). By integrating the current sampling with previously published data (González-Martínez et al., 2010; Mayol et al., 2015), we were able to verify the importance of this area as a source of genetic variation for the species in the Iberian Peninsula. Our results revealed higher genetic structuring than previously reported, likely reflecting the combined effect of historical changes and more recent impact of human pressure, as discussed below. The genetic data here reported also provide key information to advise both management and conservation policies for this locally threatened species.

Genetic diversity in the Cantabrian-Atlantic Region (CR):

420 The yew collection in our study (1,054 trees from 46 CR locations, only 7 with sample size
421 (N) < 10) provided much broader population data than previously reported in the Cantabrian
422 Range (González-Martínez et al., 2010: 81 trees, 5 populations, 1 with N < 10). The average
423 pairwise F_{ST} for the whole region of 11.2% is substantially higher than that reported by
424 González-Martínez et al. (2010; mean F_{ST} 3.6%), and more similar to the genetic
425 differentiation reported for populations within the Pyrenees Range (10.8%), also in the north
426 of Iberian Peninsula. These results are explained by the presence of some highly divergent
427 populations not sampled before and, in fact, some of these samples were responsible for the
428 increased allelic richness in CR, contributing to the total genetic diversity, including a high
429 number of private alleles suggestive of local differentiation (close to 50). Despite the relative
430 contribution of each population to the total genetic diversity was rather low, according to the
431 high number of populations studied, notable differences were observed among them, which
432 reinforced the importance of broad sampling surveys to refine population genetic data and
433 thus provide useful information for conservation purposes (Petit et al., 1998; Ribeiro et al.,
434 2017). In general, much higher population contribution values were observed from AR than
435 He estimates. Interestingly, allelic diversity is acknowledged as a valuable estimator in the
436 context of conservation genetics (Caballero et al., 2010), better reflecting the adaptive
437 potential of populations in the long term than gene-frequency estimators (Caballero and
438 García-Dorado, 2013; Vilas et al., 2015; López-Cortegano et al., 2019). We also found a high
439 intra-population relatedness coefficient (pairwise *r*) and a low Ne for many populations,
440 suggesting that isolation and genetic drift might be among the evolutionary forces currently
441 shaping genetic diversity in CR. Accordingly, IBD was not significant after discarding outlier
442 populations, as reported in different geographical ranges of the species (Dubreuil et al., 2010;
443 González-Martínez et al., 2010; Chybicki et al., 2012). However, some geographic regions
444 within CR, such as NCR, have maintained higher historical gene flow reflected by higher

445 genetic variability and lower pairwise differentiation (mean F_{ST} 5.2%), similar to that reported
446 for low structured populations in Great Britain (Gargiulo et al., 2019).

447 Bayesian clustering revealed different layers of genetic structuring, using a flexible
448 interpretation of k values, as recommended (Novembre 2016; Lawson et al., 2018). The most
449 likely level (k = 2) suggested a certain separation between western and central-easternmost
450 samples of the CR. The genetic differentiation of previously unstudied populations (e.g.,
451 BCM) was also supported by the higher number of private alleles. This k = 2 structure could
452 be also reflecting historical connectivity among the western populations, as suggested for the
453 Cantabrian Range by González-Martínez et al. (2010). The Bayesian analysis also suggested
454 substructure for k = 4, mostly compatible with a biogeographical subdivision in seven
455 regional groups across CR (GID, GDR, NAM, WCR, NCR, SCR and BCM), supported by a
456 small but significant genetic differentiation (AMOVA inter-group: ~ 2%; $P < 0.001$). The
457 regional structure has been considered useful to identify and capture the genetic diversity
458 among seed sourcing areas for British yew populations (Gargiulo et al., 2019; average F_{ST}
459 5.6%; regional variation 3%) and to define yew protection areas in Poland (Litkowiec et al.,
460 2018; average F_{ST} 15.5%; regional variation 5.3%). In addition, when a range of increasing
461 numbers of groups in the STRUCTURE results were explored, a local spatial structure pattern
462 closer to the number of populations was observed, suggesting limited gene flow despite the
463 high potential for dispersal of this species. Some small and fragmented yew populations could
464 be also affected by different factors, such as closure of tree canopies, unequal sex ratios and
465 limited distribution of frugivorous species that might have reduced pollination and seed
466 dispersal (Hilfiker et al., 2004; Dubreil et al., 2010; Iszkuło, 2010; Chybicki et al., 2011;
467 Lavabre and García, 2015; Chybicki and Oleksa, 2018). The average pairwise F_{ST} between
468 yew populations in CR and most regional groups (NAM 10.5%, WCR 10.6%, SCR 10.1%,
469 BCM 9.8%), nearly doubled that found in Great Britain (Gargiulo et al., 2019). Microsatellite-

based Ne estimates in CR revealed quite similar results (19% Ne > 50; mostly from NAM, WCR, NCR and BCM regions) to those found in Poland, where large Ne represented a low fraction of the populations studied (16% Ne > 50; Litkowiec et al., 2018).

Biogeographical and historical climatic factors, important for understanding yew evolution (González-Martínez et al., 2010; Mayol et al., 2015), were considered for interpreting the regional structure detected within CR. Although admixture seems to have been high in the past, specific genetic components within each regional group could be partially related to signatures from historical events. Isolation in different glacial refugia and later re-colonisation and admixture throughout altitudinal gradients has been hypothesized for other forest species (Wójkiewicz et al., 2016; Ribeiro et al., 2017). Molecular and paleo-ecological evidence has been reported for different glacial refugia during the last ice age for tree species in the Cantabrian coast and adjacent mountains, influencing current patterns of genetic structure (Ramil-Rego et al., 1998a,1998b; López de Heredia et al., 2007; Muñoz-Sobrino et al., 2018). Genetic differentiation between eastern (BCM) and western (NAM, WCR) populations of CR could be associated with distinct glacial refugia in western coastal areas and the Basque mountains, as reported for other tree species (López de Heredia et al., 2007; Gómez-Orellana et al., 2013). Other factors cannot be discarded to explain the genetic diversity patterns observed; higher admixture in some regions could suggest larger and more connected populations (e.g., NCR, WCR), whereas more homogeneous clustering could indicate higher isolation and/or reduction in effective population size. Indeed, each regional group showed signals of subclustering pattern, sometimes with high inter-population differentiation and local familiar structure related to increased genetic drift and restricted gene flow in some small and fragmented populations. Such strong subdivision could be partly explained by human impact due to fires and management practices (e.g., Dubreuil et al, 2010).

The Cantabrian Region within the Iberian Peninsula

Genotyping standardization across different datasets (González-Martínez et al., 2010; Mayol et al., 2015) enabled a confident comparison of European yew populations based on a common subset of seven microsatellite loci. The whole Iberian Peninsula dataset (WIS) allowed revisiting the patterns of genetic diversity after increasing the representation of the CR in this study. In agreement with previous data, higher genetic diversity and lower structure characterized populations of CR and Pyrenees compared to central and south-eastern ranges (González-Martínez et al., 2010).

The clustering results on the extended sampling in the Iberian Peninsula suggested a pattern separating eastern (Betic Range, Catalonian Ranges, Pyrenees) and central-western (Cantabrian Range, Central System Range, Iberian System Range) populations, which resembled the differentiation pattern between the Atlantic and Mediterranean drainages observed in other Iberian tree species, and related to genetic signatures of past geological events, consistent with the refugia-within-refugia model (Rodríguez-Sánchez et al., 2010; Macaya-Sanz et al., 2012; Santiso et al., 2016). However, clustering algorithms may sometimes fail to detect much more complex demographic scenarios (Lawson et al. 2018 and references therein). In fact, estimators of genetic diversity showed significant latitudinal and longitudinal trends (See 3.2 Results section). These data support previously reported clines in the Iberian Peninsula of increasing genetic variation from southeast to northwest, which have been attributed to higher climatic stability and lower impact of anthropogenic disturbances (González-Martínez et al., 2010).

On the other hand, the best number of Bayesian clusters (i.e. k values) observed within-region in the "global" WIS scenario was lower than previously reported in some regions (González-Martínez et al., 2010), particularly in the Catalonian Ranges, the Iberian System Range and the Pyrenees. This could be explained by the wider sampling scenario in this study, although other issues related to genotyping standardization and study-specific null alleles cannot be

discarded. Nevertheless, subregional substructure patterns were detected when higher *k* values were explored in the Betic range, the Central System, the Catalonian ranges, the Iberian System and Pyrenees, in accordance with previous data (González-Martínez et al., 2010): admixture was higher in most samples from the western-central Pyrenees (e.g., Bujaruelo1 and Bujaruelo2) than in southern pre-Pyrenean populations (e.g., PasOca, PasEmilio).

Conservation remarks

Conservation of biodiversity should consider issues such as limited spatial, economic or logistic resources (operational conservation or management units; Allendorf et al., 2013). Requirements for defining genetic conservation units in forestry have focused on prioritizing the dynamics of evolutionary processes within natural tree populations (Koskela et al., 2013; Aravanopoulous, 2016). Therefore, maximizing the retention of genetic diversity and adaptive evolutionary potential of populations for conservation and management plans is crucial. Given the extent of *T. baccata* distribution and the genetic structure at different global, regional and local spatial scales in the CR, both "*in-situ*" and "*ex-situ*" strategies of conservation should be considered, which would require cooperation among different local governments in charge of forest management (Montiel and Galiana, 2005).

The small but significant regional structure detected within CR, mostly compatible with its biogeographical subdivision, should be the basis to define regional conservation units for the management of European yew in this region. However, the singularity of some populations/areas should also be considered. For example, the highly differentiated and genetically impoverished populations at the western CR edge, probably as a result of strong geneticdrift, might also contribute important adaptive variation for conservation (see Mayol et al., 2020). Other valuable populations contributing to both genetic diversity and differentiation, such as PA039 and LE14, among others, or retaining a large amount of the total genetic diversity (both He and AR), like GI002, are also of concern, providing added

conservation value at a fine spatial scale.should represent more refined structure to be conserved. Such populations should deserve *in-situ* protection and be also considered for *ex-situ* conservation, e.g. germplasm collection. Moreover, given that the eastern part of CR (BCM) along with WCR and NCR preserve a large proportion of the total genetic diversity, they should be taken into account to be represented in *ex situ* collections, although other less variable but differentiated regions (NAM, SCR) should be also valuable to preserve potential adaptive diversity.

Four out of the six genetic conservation units (GCUs) previously delineated within CR by the EUFORGEN program to preserve adaptive variation in marginal and often quite small populations (http://www.euforgen.org/species/taxus-baccata/) are supported in this study (LU10, LE14, LE1000, BU31). Other GCUs might be complemented with selected populations from less represented regional groups (e.g., NAM, GDR, SCR) based on their contribution to the total diversity and differentiation in CR. Maximizing the maintenance of the total genetic diversity and adaptive potential of the yew across different biogeographic territories in CR is particularly relevant in the light of evidence for local adaption of yew to the wide temperature range in the Iberian Peninsula (Mayol et al., 2020).

This study provides genetic criteria within the context of a conservation programme for improving the habitats of European community interest associated with yew (like 9580*: "Mediterranean Forests of *Taxus baccata*"; Natura 2000 Network). The conservation of yew in their natural habitats, particularly for those populations and regions (e.g., BCM, NAM, WCR) with high allelic diversity associated with larger effective population sizes and higher gene flow in more stable humid conditions, would be benefited by landscape management aimed at maintaining connectivity. On the other hand, smaller and isolated populations showing stronger relatedness (probably related to more recent bottlenecks and limited gene flow associated with stressful environmental conditions and/or human disturbances) would

570 require some reinforcement actions with individuals from the same or nearby populations.
571 Among the actions within the conservation programme, increasing the area of occupation and
572 recruitment should consider using forest plants from the same or genetically related
573 populations at global scales and/or regional biogeographic groupings. Complementary
574 silvicultural actions, like canopy opening to regulate space competition by beech trees,
575 controlling herbivorous fauna or carrying out preventive protection management of the
576 ecological system (e.g., fires, pest control) have been planned to improve the conservation of
577 these protected habitats associated with the persistence of yew populations. The
578 representativeness of the genetic diversity captured in the wild for population reinforcement
579 and/or *ex situ* conservation in germplasm (seeds) and tree collections, including monumental
580 yews, should be also considered, along with their potential contribution to gene flow and risks
581 of hybridization with natural populations (Arnet et al., 2015; Chen and Sun, 2018; Gargiulo et
582 al., 2019). These additional issues would require further research and monitoring in the future.
583 The yew DNA collections gathered in this and other studies in the Iberian Peninsula are
584 essential resources to further research into local and/or regional adaptation to enrich the main
585 conclusions of this study. More powerful population genomic methods can be applied for
586 searching for adaptive variation and selective breedingtowards a sustainable conservation of
587 forest genetic resources (Koskela et al., 2013; Grattapaglia, 2014; Aravanopoulos, 2016;
588 Plomion et al., 2016). Nevertheless, the neutral microsatellite markers still provide useful
589 monitoring of genetic diversity in forest species, including the yew (Mayol et al., 2015;
590 Aravanopoulos, 2016; Litkowiec et al., 2018; Gargiulo et al., 2019), like documented in this
591 study. High-throughput SNP (single nucleotide polymorphisms) datasets (Olsson et al., 2018;
592 Mayol et al., 2020) have confirmed microsatellite patterns of historical demography of
593 populations, but additionally provided new evidence of local adaptation to current and past
594 climatic factors (Mayol et al., 2020).

5. Conclusions

This study improves the genetic characterization of the European yew (*T. baccata)* covering its natural distribution in the Cantabrian-Atlantic Region of the Iberian Peninsula (CR) as a basis to assist conservation and management of protected yew-associated habitats (LIFE-BACCATA; LIFE15 NAT/ES/000790). The results confirmed high genetic diversity of yew resources in this region within the context of the Iberian Peninsula. Compared with other regions in Southern Europe and the Iberian Peninsula, CR shows lower genetic structure at a global scale, probably associated with historical gene flow. Small significant regional clustering linked to ancient biogeographical processes was also suggested, along with some genetic structure and familiar clusters at a finer spatial scale. Consistent with biogeographical subdivisions, we were able to identify seven regional conservation units across the Cantabrian-Atlantic Region, characterized by low within-group differentiation and that maximized inter-group divergence. Besides the contribution to the total genetic diversity of each group, global, regional and local components of genetic variation should be considered for preserving the genetic resources of yew within CR. As a part of the LIFE-BACCATA project, this information will be the basis for prioritizing active preservation of genetic diversity and evolutionary potential of yew populations in the Cantabrian Region through *in situ* and *ex situ* protection and restoration actions.

Acknowledgments

This study was supported by the Life Programme of the European Commission (Project reference: LIFE15 NAT/ES/000790). We thank all the personnel involved in fieldwork actions and sampling coordination from LIFE BACCATA partners (IBADER-USC Junta de Castilla y León, Fundación CESEFOR Fundación HAZI), as well as from regional and local public

bodies that have provided the necessary permits and authorizations: Xunta de Galicia, Principado de Asturias, Gipuzkoa Foru Aldundia and Bizkaiko Foru Aldundia.The collaboration of Asier Díez and Nere Amaia Laskurain, from the University of the Basque Country, has been essential to collect samples in the Basque Country. The authors are also grateful to Ana Bella Díez Gutiérrez (Forest Service of General Directorate for Biodiversity, Environment and Climate Change, Gobierno de Cantabria) for kindly providing samples from Cantabria region. Data previously reported for the Iberian Peninsula was supported by projects TAXUS (CGL2007-63107/BOS) and ADAPCON (CGL2011-30182-C02-01/02) from the Spanish Ministry of Economy and Competitiveness. We acknowledge the support of Adrián Millán and Raquel Fernández-Cebrián on microsatellite genotyping. Thanks to María Portela and Mónica Otero for technical laboratory and administrative support.

References

Allendorf, F.W., Luikart, G., Aitken, S.N., 2013. Conservation and the Genetics of Populations, 2nd ed. Wiley-Blackwell Publishing Ltd.

Aravanopoulos, F.A., 2016. Conservation and monitoring of tree genetic resources in temperate forests. Curr. Forestry Rep. 2, 119–129. https://doi.org/10.1007/s40725-016-0038-8

Arnet, M., Santos, B., Brockerhoff, E.G., Pelser, P.B., Ecroyd, C., Clemens, J., 2015. Importance of arboreta for ex situ conservation of threatened trees. Biodivers. Conserv. 24, 3601–3620. https://doi.org/10.1007/s10531-015-1024-3

Birks, H.J.B., 2019. Contributions of Quaternary botany to modern ecology and biogeography. Plant Ecol. Divers. 12, 189-385. https://doi.org/10.1080/17550874.2019.1646831

Brockerhoff, E.G., Barbaro, L., Castagneyrol, B., Forrester, D.I., Gardiner, B., González-

643 Olabarria, J.R., O'B. Lyver, P., Meurisse, N., Oxbrough A., Taki H., Thompson I.D.,
644 van der Plas F., Jactel, H. 2017. Forest biodiversity, ecosystem functioning and the
645 provision of ecosystem services. Biodivers. Conserv. 26. 3005-3035.
646 https://doi.org/10.1007/s10531-017-1453-2

647 Caballero, A., Rodríguez-Ramilo, S.T., Avila, V., Fernández, J., 2010. Management of genetic
648 diversity of subdivided populations in conservation programmes. Conserv. Genet.11,
649 409–419. https://doi.org/10.1007/s10592-009-0020-0

650 Caballero, A., García-Dorado, A., 2013. Allelic diversity and its implications for the rate of
651 adaptation. Genetics 195, 1373–1384. https://doi.org/10.1534/genetics.113.158410

652 Cao, C.P., Leinemann, L., Ziehe, M., Finkeldey, R., 2004. Study of genetic variation and
653 differentiation of yew (*Taxus baccata* L.) stands using isozyme and DNA marker. Allg.
654 Forst. Jagdztg. 175, 21-28.

655 CBD, 2014. Global Biodiversity Outlook 4. A mid-term assessment of progress towards the
656 implementation of the Strategic Plan for Biodiversity 2011-2020. Montreal: Secretariat
657 of the Convention on Biological Diversity (CBD).

658 Chen, G., Sun, W., 2018. The role of botanical gardens in scientific research, conservation,
659 and citizen science. Plant Divers. 40, 181–188.
660 https://doi.org/10.1016/j.pld.2018.07.006

661 Chybicki, I.J., Burczyk, J., 2008. Simultaneous estimation of null alleles and inbreeding
662 coefficients. J. Hered. 100, 106–113. https://doi.org/10.1093/jhered/esn088

663 Chybicki, I.J., Oleksa, A., Burczyk, J., 2011. Increased inbreeding and strong kinship
664 structure in *Taxus baccata* estimated from both AFLP and SSR data. Heredity 107,
665 589–600. https://doi.org/10.1038/hdy.2011.51

666 Chybicki, I.J., Oleksa, A., Kowalkowska, D., 2012. Variable rates of random genetic drift in

667 protected populations of English yew: implications for gene pool conservation.
668 Conserv. Genet. 13, 899–911. https://doi.org/10.1007/s10592-012-0339-9

669 Chybicki, I.J., Oleksa, A., 2018. Seed and pollen gene dispersal in *Taxus baccata*, a dioecious
670 conifer in the face of strong population fragmentation. Ann. Bot. 122, 409-421.
671 https://doi.org/10.1093/aob/mcy081

672 De Vries, S.M.G., Alan, M., Bozzano, M., Burianek, V., Collin, E., Cottrell, J., Ivankovic, M.,
673 Kelleher, C.T., Koskela, J., Rotach, P., Vietto, L., Yrjänä, L., 2015. Pan-European
674 strategy for genetic conservation of forest trees and establishment of a core network of
675 dynamic conservation units. European Forest Genetic Resources Programme
676 (EUFORGEN), Bioversity International, Rome, Italy. xii, 40p.
677 http://www.euforgen.org/fileadmin/templates/euforgen.org/upload/Publications/Thema
678 tic_publications/EUFORGEN_FGR_conservation_strategy.pdf

679 Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator
680 v2: re-implementation of software for the estimation of contemporary effective
681 population size (Ne) from genetic data. Mol. Ecol. Resour. 14, 209–214.
682 https://doi.org/10.1111/1755-0998.12157

683 Dubreuil, M., Sebastiani, F., Mayol, M., González-Martínez, S.C., Riba, M., Vendramin,
684 G.G., 2008. Isolation and characterization of polymorphic nuclear microsatellite loci
685 in *Taxus baccata* L. Conserv. Genet. 9, 1665–1668. https://doi.org/10.1007/s10592-
686 008-9515-3

687 Dubreuil, M., Riba, M., González-Martínez, S.C., Vendramin, G.G., Sebastiani, F., Mayol,
688 M., 2010. Genetic effects of chronic habitat fragmentation revisited: strong genetic
689 structure in a temperate tree, *Taxus baccata* (Taxaceae), with great dispersal capability.
690 Am. J. Bot. 97, 303-310. https://doi.org/10.3732/ajb.0900148

691 Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for
692 visualizing STRUCTURE output and implementing the Evanno method. Conserv.
693 Genet. Resour. 4, 359–361. https://doi.org/10.1007/s12686-011-9548-7

694 EEA, 2015. Linking in situ vegetation data to the EUNIS habitat classification: results for
695 forest habitats. European Environment Agency (EEA), Technical report. No 18/2015.
696 Luxembourg: Publications Office of the European Union.

697 European Union, 2011. The EU Biodiversity Strategy to 2020. Luxembourg: Publications
698 Office of the European Union. doi: 10.2779/39229

699 Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to
700 perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour.
701 10, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

702 Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals
703 using the software STRUCTURE: a simulation study. Mol Ecol 14, 2611-2620.
704 https://doi.org/10.1111/j.1365-294X.2005.02553.x

705 Fady, B., Aravanopoulos, F.A., Alizoti, P., Mátyás, C., von Wühlisch, G., Westergren, M.,
706 Belletti, P., Cvjetkovic, B., Ducci, F., Huber, G., et al., 2016. Evolution-based
707 approach needed for the conservation and silviculture of peripheral forest tree
708 populations. Forest Ecol. Manag. 375, 66–75.
709 https://doi.org/10.1016/j.foreco.2016.05.015

710 Ferreiro da Costa, J., Ramil-Rego, P., 2018. Biological Conservation and Nature Protection
711 Strategies in Spanish Atlantic Region, in: Bülent Sen (Ed.) Selected Studies in
712 Biodiversity. InTech. Rijeka. Croacia. pp. 25-44.
713 http://dx.doi.org/10.5772/intechopen.72515

714 Francis, R.M., 2017. pophelper: an R package and web app to analyse and visualize

715 population structure. Mol. Ecol. Resour. 17, 27–32. https://doi.org/10.1111/1755-
716 0998.12509

717 Fussi, B., Westergren, M., Aravanopoulos, F., Baier, R., Kavaliauskas, D., Finzgar, D., Alizoti,
718 P., Bozic, G., Avramidou, E., Konnert, M., Kraigher, H., 2016. Forest genetic
719 monitoring: an overview of concepts and definitions. Environ. Monit. Assess. 188,
720 493. https://doi.org/10.1007/s10661-016-5489-7

721 Gargiulo, R., Saubin, M., Rizzuto, G., West, B., Fay, M.F., Kallow, S., Trivedi, C., 2019.
722 Genetic diversity in British populations of *Taxus baccata* L.: Is the seedbank
723 collection representative of the genetic variation in the wild? Biol. Conserv. 233, 289–
724 297. https://doi.org/10.1016/j.biocon.2019.01.014

725 Gómez-Orellana, L., Ramil-Rego, P., Muñoz Sobrino, C., 2013. The response of vegetation at
726 the end of the last glacial period (MIS 3 and MIS 2) in littoral areas of NW Iberia.
727 Boreas 42, 729–744. https://doi.org/10.1111/j.1502-3885.2012.00310.x

728 González-Martínez, S.C., Dubreuil, M., Riba, M., Vendramin, G.G., Sebastiani, F., Mayol,
729 M., 2010. Spatial genetic structure of *Taxus baccata* L. in the western Mediterranean
730 Basin: past and present limits to gene movement over a broad geographic scale. Mol.
731 Phylogenet. Evol. 55, 805–815. https://doi.org/10.1016/j.ympev.2010.03.001

732 Grattapaglia, D., 2014. Breeding forest trees by genomic selection: current progress and the
733 way forward, in: Tuberosa, R., Graner, A., Frison, E., (eds,) Genomics of Plant
734 Genetic Resources. Springer, Dordrecht. pp, 651-682. https://doi.org/10.1007/978-94-
735 007-7572-5_26

736 Hardy, O.J., Vekemans, X., 2002. SPAGeDi: a versatile computer program to analyse spatial
737 genetic structure at the individual or population levels. Mol. Ecol. Notes 2, 618–620.
738 https://doi.org/10.1046/j.1471-8286.2002.00305.x

739 Hilfiker, K, Gugerli, F, Schütz, J.-P., Rotach, P, Holderegger, R., 2004. Low RAPD variation
740 and female-biased sex ratio indicate genetic drift in small populations of the dioecious
741 conifer *Taxus baccata* in Switzerland. Conserv. Genet. 5, 357-365.
742 https://doi.org/10.1023/B:COGE.0000031144.95293.1b

743 Iszkuło, G., 2010. Success and failure of endangered tree species: low temperatures and low
744 light availability affect survival and growth of European yew (*Taxus baccata* L.)
745 seedlings. Pol. J. Ecol. 58, 259-271.

746 Iszkuło, G., Didukh, Y., Giertych, M.J., Jasińska, A.K., Sobierajska, K., Szmyt, J., 2012. Weak
747 competitive ability may explain decline of *Taxus baccata*. Ann. Forest Sci. 69, 705-
748 712. https://doi.org/10.1007/s13595-012-0193-4

749 Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: a
750 program for identifying clustering modes and packaging population structure
751 inferences across *K*. Mol. Ecol. Resour. 15, 1179–1191. https://doi.org/10.1111/1755-
752 0998.12387

753 Koskela, J., Lefèvre, F., Schueler, S., Kraigher, H., Olrik, D.C., Hubert, J., Longauer, R.,
754 Bozzano, M., Yrjänä, L., Alizoti, P., Rotach, P., Vietto, L., Bordács, S., Myking, T.,
755 Eysteinsson, T., Souvannavong, O., Fady, B., De Cuyper, B., Heinze, B., von
756 Wühlisch, G., Ducousso, A., Ditlevsen, B., 2013. Translating conservation genetics
757 into management: Pan-European minimum requirements for dynamic conservation
758 units of forest tree genetic diversity. Biol. Conserv. 157, 39–49.
759 https://doi.org/10.1016/j.biocon.2012.07.023

760 Lavabre, J.E., García, D., 2015. Geographic consistency in the seed dispersal patterns of
761 *Taxus baccata* L. in the Iberian Peninsula. Forest Systems, 24, e040.
762 http://dx.doi.org/10.5424/fs/2015243-07462

763 Lawson, D.J., van Dorp, L., Falush, D., 2018. A tutorial on how not to over-interpret
764 STRUCTURE and ADMIXTURE bar plots. Nat. Commun. 9, 3258.
765 https://doi.org/10.1038/s41467-018-05257-7

766 Lefèvre, F., Alia, R., Bakkebø Fjellstad, K., Graudal, L., Oggioni, S.D., Rusanen, M.,
767 Vendramin, G.G., Bozzano, M., 2020. Dynamic conservation and utilization of forest
768 tree genetic resources: indicators for *in situ* and *ex situ* genetic conservation and forest
769 reproductive material. European Forest Genetic Resources Programme
770 (EUFORGEN), European Forest Institute.
771 http://www.euforgen.org/publications/publication/dynamic-conservation-and-
772 utilization-of-forest-tree-genetic-resources-indicators-for-in-situ/

773 Lewandowski, A., Burczyk, J., Mejnartowicz, L., 1995. Genetic structure of English yew
774 (*Taxus baccata* L.) in the Wierzchlas Reserve: implications for genetic conservation.
775 Forest Ecol. Manag. 73, 221-227. https://doi.org/10.1016/0378-1127(94)03477-E.

776 Linares, J.C., 2013. Shifting limiting factors for population dynamics and conservation status
777 of the endangered English yew (*Taxus baccata* L., Taxaceae). Forest Ecol. Manag.
778 291, 119–127. https://doi.org/10.1016/j.foreco.2012.11.009

779 Litkowiec, M., Lewandowski, A. Wachowiak, W., 2018. Genetic variation in *Taxus baccata*
780 L.: a case study supporting Poland's protection and restoration program. Forest Ecol.
781 Manag. 409, 148–160. https://doi.org/10.1016/j.foreco.2017.11.026

782 López-Cortegano, E., Pouso, R., Labrador, A., Pérez-Figueroa, A., Fernández, J., Caballero,
783 A., 2019. Optimal management of genetic diversity in subdivided populations. Front.
784 Genet. 10, 843. https://doi.org/10.3389/fgene.2019.00843

785 López de Heredia, U., Carrión, J.S., Jiménez, P., Collada, C., Gil, L., 2007. Molecular and
786 palaeoecological evidence for multiple glacial refugia for evergreen oaks on the

787 Iberian Peninsula. J. Biogeogr. 34, 1505–1517. https://doi.org/10.1111/j.1365-
788 2699.2007.01715.x

789 Macaya-Sanz, D., Heuertz, M., López de Heredia, U., De Lucas, A.I., Hidalgo, E., Maestro,
790 C., Prada, A., Alía, R., González-Martínez, S.C., 2012. The Atlantic-Mediterranean
791 watershed, river basins and glacial history shape the genetic structure of Iberian
792 poplars. Mol. Ecol. 21, 3593-3609. https://doi.org/10.1111/j.1365-294X.2012.05619.x

793 Maguire, T.L., Peakall, R., Saenger., P., 2002. Comparative analysis of genetic diversity in the
794 mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by
795 AFLPs and SSRs. Theor. Appl. Genet. 104, 388-398.
796 https://doi.org/10.1007/s001220100724

797 Mahmoodi, P., Khayam Nekouei, S.S., Mardi, M., Pirseyedi, S.S., Ghaffari, M.R., Ramroodi,
798 M., Siahsar, B.A., 2010. Isolation and characterization of new microsatellite marker in
799 *Taxus baccata* L. Conserv. Genet. Resour. 2, 195–199. https://doi.org/10.1007/s12686-
800 009-9133-5

801 Mantyka-Pringle, C.S., Martin, T.G., Rhodes, J.T. 2012. Interactions between climate change
802 and habitat loss effects on biodiversity: a systematic review and meta-analysis. Global
803 Change Biol. 18, 1239-1252. https://doi.org/10.1111/j.1365-2486.2011.02593.x

804 Mayol, M., Riba M., González-Martínez, S.C., Bagnoli, F., de Beaulieu, J.L., Berganzo, E.,
805 Burgarella, C., Dubreuil, M., Krajmerová, D., Paule, D., Romsakova, I., Vettori, C.,
806 Vincenot, L., Vendramin, G.G., 2015. Adapting through glacial cycles: insights from a
807 long-lived tree (*Taxus baccata*). New Phytol. 208, 973–986.
808 https://doi.org/10.1111/nph.13496

809 Mayol, M., Riba, M., Cavers, S., Grivet, D., Vincenot, L., Cattonaro, F., Vendramin, G.G.,
810 González-Martínez, S.C., 2020. A multiscale approach to detect selection in nonmodel

811 tree species: Widespread adaptation despite population decline in *Taxus baccata* L.
812 Evol. Appl. 13, 143-160. https://doi.org/10.1111/eva.12838

813 Millennium Ecosystem Assessment, 2005. Ecosystems and human well-being: Synthesis. Vol.
814 5. Island press Washington, DC

815 Montiel, C., Galiana, L., 2005. Forest policy and land planning policy in Spain: a regional
816 approach. Forest Policy Econ. 7, 131–142. https://doi.org/10.1016/S1389-
817 9341(03)00026-1

818 Muñoz Sobrino, C., Ramil-Rego, P., Gómez-Orellana, L., Ferreiro da Costa, J., Díaz Varela,
819 R.A., 2009. Climatic and human effects on the post-glacial dynamics of *Fagus*
820 *sylvatica* L. in NW Iberia. Plant Ecol. 203, 317–340. https://doi.org/10.1007/s11258-
821 008-9552-5

822 Muñoz Sobrino, C., García-Moreiras, I., Gómez-Orellana, L., Iriarte Chiapusso, M.J., Heiri,
823 O., Lotter, A.F., Ramil-Rego, P., 2018. The last hornbeam forests in SW Europe: new
824 evidence on the demise of *Carpinus betulus* in NW Iberia. Veget. Hist. Archaeobot.
825 27, 551–576. https://doi.org/10.1007/s00334-017-0654-7

826 Myking, T., Vakkari, P., Skrøppa T., 2009. Genetic variation in northern marginal *Taxus*
827 *baccata* L. populations. Implications for conservation. Forestry 82, 529–539.
828 https://doi.org/10.1093/forestry/cpp022

829 Novembre, J., 2016. Pritchard, Stephens, and Donnelly on Population Structure. Genetics,
830 204, 391-393. https://doi.org/10.1534/genetics.116.195164

831 Olsson, S., Pinosio, S., González-Martínez, S.C., Abascal, F., Mayol, M., Grivet, D.,
832 Vendramin, G.G., 2018. De novo assembly of English yew (*Taxus baccata*)
833 transcriptome and its applications for intra-and inter-specific analyses. Plant Mol.
834 Biol. 97, 337–345. https://doi.org/10.1007/s11103-018-0742-9

835 Peakall, R.O.D., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population
836 genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.
837 https://doi.org/10.1111/j.1471-8286.2005.01155.x

838 Petit, R.J., El Mousadik, A., Pons, O., 1998. Identifying populations for conservation on the
839 basis of genetic markers. Conserv. Biol. 12, 844–855. https://doi.org/10.1111/j.1523-
840 1739.1998.96489.x

841 Plomion, C., Bastien, C., Bogeat-Triboulot, M.B., Bouffier, L., Déjardin, A., Duplessis, S.,
842 Fady, B., Heuertz, M., Le Gac, A.L., Le Provost, G., Legué, V., Lelu-Walter, M.A.,
843 Leplé, J.C., Maury, S., Morel, A., Oddou-Muratorio, S., Pilate, G., Sanchez, L., Scotti,
844 I., Scotti-Saintagne, C., Segura, V., Trontin, J.F., Vacher, C., 2016. Forest tree
845 genomics: 10 achievements from the past 10 years and future prospects. Ann. Forest
846 Sci. 73, 77–103. https://doi.org/10.1007/s13595-015-0488-3

847 Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using
848 multilocus genotype data. Genetics 155, 945–959.

849 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J.,
850 Sklar, P., De Bakker, P.I.W., Daly, M.J., 2007. PLINK: a tool set for whole-genome
851 association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.
852 https://doi.org/10.1086/519795

853 Ramil-Rego, P., Muñoz-Sobrino, C., Rodríguez-Guitián, M., Gómez-Orellana, L., 1998a.
854 Differences in the vegetation of the North Iberian Peninsula during the last 16 000
855 years, Plant Ecol. 138, 41–62. https://doi.org/10.1023/A:1009736432739

856 Ramil-Rego, P., Rodríguez-Guitián, M., Muñoz-Sobrino, C., 1998b. Sclerophyllous
857 vegetation dynamics in the north of the Iberian Peninsula during the last 16,000 years.
858 Glob. Ecol. Biogeogr. Let. 7, 335–351. https://doi.org/10.1046/j.1466-

822x.1998.00300.x

Raymond, M., Rousset, F., 1995. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. J. Hered. 86, 248–249.

Ribeiro, M.M., Piotti, A., Ricardo, A., Gaspar, D., Costa, R., Parducci, L., Vendramin, G.G., 2017. Genetic diversity and divergence at the *Arbutus unedo* L. (Ericaceae) westernmost distribution limit. PloS One 12, e0175239. https://doi.org/10.1371/journal.pone.0175239

Rivas-Martínez, S., Penas, A., Díaz-González, T.E., del Río, S., Cantó, P., Herrero, L., Pinto Gomes C., Costa, J.C., 2014. Biogeography of Spain and Portugal. Preliminary typological synopsis. Int. J. Geobot. Res. 4, 1-64.

Rodríguez-Sánchez, F., Hampe, A., Jordano, P., Arroyo, J., 2010. Past tree range dynamics in the Iberian Peninsula inferred through phylogeography and palaeodistribution modelling: A review. Rev. Palaeobot. Palynol. 162, 507-521. https://doi.org/10.1016/j.revpalbo.2010.03.008

Santiso, X., Lopez, L., Retuerto, R., Barreiro, R., 2016. Phylogeography of a widespread species: pre-glacial vicariance, refugia, occasional blocking straits and long-distance migrations. AoB Plants, 8: plw003. https://doi.org/10.1093/aobpla/plw003

Svenning, J.C., Magård, E., 1999. Population ecology and conservation status of the last natural population of English yew *Taxus baccata* in Denmark. Biol. Conserv. 88, 173–182. https://doi.org/10.1016/S0006-3207(98)00106-2

Thomas, P.A., Polwart, A., 2003. *Taxus baccata* L. J. Ecol. 91, 489–524. https://doi.org/10.1046/j.1365-2745.2003.00783.x

Thomas, P.A., García-Marti, X., 2015. Response of European yews to climate change: a review. Forest Syst. 24, eR01. https://doi.org/10.5424/fs/2015243-07465

Tilman, D., May, R.M., Lehman, C.L., Nowak, M.A., 1994. Habitat destruction and the extinction debt. Nature 371, 65-66. https://doi.org/10.1038/371065a0

Tröber, U., Ballian, D., 2011. Genetic characterization of English yew (*Taxus baccata* L.) populations in Bosnia and Herzegovina. Eur J Forest Res. 130, 479–489. https://doi.org/10.1007/s10342-010-0436-6

Vilas, A., Pérez-Figueroa, A., Quesada, H., Caballero, A., 2015. Allelic diversity for neutral markers retains a higher adaptive potential for quantitative traits than expected heterozygosity. Mol Ecol. 24, 4419-4432. https://doi.org/10.1111/mec.13334

Wang, J., 2001. A pseudo-likelihood method for estimating effective population size from temporally spaced samples. Genet. Res. 78, 243–257. https://doi.org/10.1017/S0016672301005286

Wójkiewicz, B., Litkowiec, M., Wachowiak, W., 2016. Contrasting patterns of genetic variation in core and peripheral populations of highly outcrossing and wind pollinated forest tree species. AoB Plants 8, plw054. https://doi.org/10.1093/aobpla/plw054

Figure Captions

Figure 1. Geographical location and genetic analysis of *Taxus baccata* from 46 populations in the Cantabrian-Atlantic Region (CR) of the Iberian Peninsula. A) Genetic diversity estimates (He: expected heterozygosity; AR: allelic richness for minimum sampling size of 15 trees) in each population; B) Global Bayesian clustering for the two most likely k values (2, 4) according to Evanno's method (Figure S3); C) Map distribution and Bayesian clustering at regional level (proportion of individual membership is shown for the inferred regional groupings at k = 4, supported by AMOVA analyses (Table 3)). Population codes (See Table 1) reflect the regional political administrations in charge of forest management where the yew populations studied were sampled (Autonomous communities (provinces)): Galicia (LU-Lugo, OU-Ourense); Castilla y León (LE-León, ZA-Zamora, PA-Palencia and BU-Burgos); AS-Asturias, CA-Cantabria; País Vasco (BZ-Bizkaia; AR-Araba and GI-Gipuzkoa). The map represented biogeographical and paleoecological regions, Natura 2000 SACs (Special Areas of Conservation; https://natura2000.eea.europa.eu/) and EUFORGEN's GCUs (Genetic conservation units; http://www.euforgen.org/species/taxus-baccata/) within the CR.

Figure 2. (a-c) Regional genetic diversity and differentiation of *Taxus baccata* in the Cantabrian-Atlantic Region (CR) of the Iberian Peninsula. Black lines represent average values, white boxes 50% intervals, vertical lines 90% intervals and black points values outside these intervals; * $P < 0.05$; ** $P < 0.01$ (Wilcoxon test). (d-e) Contribution of each regional group to the total diversity (Heterozygosity) and allelic richness (AR) subdivided into within-region diversity (blue) and differentiation (orange) components. Regional biogeographical groupings were described according to STRUCTURE and AMOVA analyses (Figure 1; Table 3).

Figure 3. Genetic diversity and structure of *Taxus baccata* in the Iberian Peninsula (WIS). A) Regional distribution of the global Bayesian clustering analysis in the Iberian Peninsula. The

925 proportion of individual membership was shown for the best number of genetic clusters
926 detected (k = 2). B) Regional comparisons of genetic diversity ((allelic richness and expected
927 heterozygosity (He))) and interpopulation coefficient of differentiation (F_{ST}). The names of
928 Iberian System and Central System mountain ranges in A) were shortened in B) (Iberian and
929 Central ranges, respectively). Black lines represent average values, white boxes 50%
930 intervals, vertical lines 90% intervals and black points values outside these intervals; * $P <$
931 0.05; ** $P < 0.01$ (Wilcoxon test).

932

933 **Supplementary Tables and Figures**

934 **Table S1:** A) Sampling information and genetic diversity of yew (*Taxus baccata*) for: A1) 46
935 yew populations across the Cantabrian-Atlantic Region (CR) of the Iberian Peninsula, and
936 A2) 80 additional populations of the Iberian Peninsula from previous studies (WIS). B)
937 Population groupings tested for AMOVA across CR. C and D) CR and WIS microsatellite
938 genotyping data, respectively.

939 **Table S2**: Pairwise A) F_{ST} and B) relatedness (*r*) estimates between yew (*Taxus baccata*)
940 populations from the Cantabrian-Atlantic Region (CR) of the Iberian Peninsula.

941 **Table S3**. AMOVA results for regional groupings of yew (*Taxus baccata*) in the Cantabrian-
942 Atlantic Region (CR) of the Iberian Peninsula.

943 **Table S4:** Pairwise F_{ST} between yew (*Taxus baccata*) populations analysed in the Iberian
944 Peninsula (WIS).

945 **Figure S1:** Plots of geographical distance between yew (*Taxus baccata*) populations in the
946 Cantabrian-Atlantic Region *vs*. A) F_{ST}; B) relatedness (*r*); C and D) F_{ST} and *r*, respectively,
947 but without populations OU001 and LU010.

948 **Figure S2.** Contribution (%) of each yew (*Taxus baccata*) population in the Cantabrian
949 Region (CR) to the total genetic variability (Heterozygosity, H) and Allelic Richness (AR).
950 The contribution is split in both cases into components of within-population diversity (blue
951 bars) and interpopulation differentiation (orange bars). Populations are arranged from west to
952 east along CR (left to right; See Figure 1).

953 **Figure S3**: Global STRUCTURE clustering of yew (*Taxus baccata*) from 46 populations in
954 the Cantabrian-Atlantic Region (CR) of the Iberian Peninsula ("Global CR" analysis). The
955 plots of Evanno's Δk values and suggestive local substructure at higher k values are shown.
956 Populations are arranged from west to east along CR (left to right; See Figure 1).

957 **Figure S4:** Regional STRUCTURE clustering of yew (*Taxus baccata*) within each
958 biogeographical region studied in the Cantabrian Region (CR) of the Iberian Peninsula
959 ("regional CR" analysis). The best number of *k* clusters is shown for each regional group.

960 **Figure S5**: Structure plots according to the Evanno's method in the "global Iberian"
961 STRUCTURE analysis of yew (*Taxus baccata*) from 128 populations in the whole dataset
962 from the Iberian Peninsula (WIS).

963 **Figure S6:** "Regional Iberian" STRUCTURE analysis of yew (*Taxus baccata*) within each
964 region considered in the whole dataset from the Iberian Peninsula (WIS).

965 **Table 1.** Sampling of the 46 *Taxus baccata* populations studied in the Cantabrian-Atlantic Region.

966 Map localization and additional information are shown on Figure 1 and Table S1, respectively.

Pop[a]	Samples	Longitude	Latitude	Grouping[b]	Autonomous Community[a]	SAC[c]/GCU[d]
LU010	20	-7.661306	43.161045	GID	Galicia	ESP0279[c]
ZA001	20	-6.793085	42.021946	GDR	Castilla y Leon	ES4190060
OU001	30	-6.780882	42.273249	GDR	Galicia	ES1130007
LU007	30	-7.115222	42.606625	NAM	Galicia	ES1120001
LU013	22	-7.058750	42.629333	NAM	Galicia	ES1120001
LE001	50	-6.879080	42.745603	NAM	Castilla y Leon	ES4130010
LE006	33	-6.832860	42.811317	NAM	Castilla y Leon	ES4130011
LE003	30	-6.745807	42.868853	NAM	Castilla y Leon	ES4130012
LE007	30	-6.700682	42.875140	NAM	Castilla y Leon	ES4130013
LE005	33	-6.688912	42.836513	NAM	Castilla y Leon	ES4130014
LE011	17	-6.362202	42.873355	WCR	Castilla y Leon	ES0000210
LE015	20	-6.286706	42.810948	WCR	Castilla y Leon	ES4130149
LE014	19	-6.281246	42.902915	WCR	Castilla y Leon	ES000021/ ESP0258[c]
LE016	5	-6.130278	42.880056	WCR	Castilla y Leon	-
LE030	3	-6.088832	42.913676	WCR	Castilla y Leon	ES4130035
LE1018	27	-5.736233	42.858954	NCR	Castilla y Leon	ES4130050
LE017	31	-5.551478	42.902348	NCR	Castilla y Leon	ES4130037
LE057	31	-5.550022	42.969282	NCR	Castilla y Leon	ES4130050
LE1000	18	-5.257607	43.066121	NCR	Castilla y Leon	ES4130003/ ESP0255 [c]
AS006	30	-5.246857	43.434120	NCR	Asturias	ES1200043

LE292	20	-5.082650	43.121100	NCR	Castilla y Leon	ES0000003
LE293	19	-4.972927	43.109128	NCR	Castilla y Leon	ES0000004
CA003	20	-4.540004	43.202000	NCR	Cantabria	ES1300001
CA001	22	-4.385000	43.296000	NCR	Cantabria	-
CA004	20	-4.195000	43.083000	NCR	Cantabria	ES1300021
LE050	30	-5.251352	42.928687	SCR	Castilla y Leon	ES4130003
LE023	20	-5.088313	42.993693	SCR	Castilla y Leon	ES4130004
LE045	31	-5.064052	42.818959	SCR	Castilla y Leon	-
LE046	31	-4.897152	42.870968	SCR	Castilla y Leon	ES4130004
PA040	13	-4.552738	42.972854	SCR	Castilla y Leon	ES4140011
PA039	24	-4.551331	42.841502	SCR	Castilla y Leon	ES4140012
BU1042	35	-3.954133	42.916152	BCM	Castilla y Leon	ES4120090
CA002	21	-3.488000	43.198110	BCM	Cantabria	-
BU031	31	-3.433368	42.790977	BCM	Castilla y Leon	ES4120030/ ESP0234[c]
BU035	32	-3.375363	43.113485	BCM	Castilla y Leon	ES4120049
BZ002	29	-3.304359	43.182748	BCM	Pais Vasco	ES2130002
BU061	32	-3.240002	43.091814	BCM	Castilla y Leon	ES4120049
BU064	20	-3.140008	42.750710	BCM	Castilla y Leon	ES4120030
BZ001	6	-2.763419	43.044047	BCM	Pais Vasco	ES2110009
AR002	5	-2.746641	42.625116	BCM	Pais Vasco	ES2110018
AR001	18	-2.605471	42.775877	BCM	Pais Vasco	ES2110015
GI005	3	-2.390794	43.273848	BCM	Pais Vasco	ES2120001
GI003	6	-2.355710	42.987780	BCM	Pais Vasco	ES2120002
GI001	30	-2.202313	43.211841	BCM	Pais Vasco	ES2120006

GI002	31	-2.108469	42.988244	BCM	Pais Vasco	ES2120011
GI004	6	-1.864141	43.260734	BCM	Pais Vasco	ES2120016

967

[a] Population codes show the regional political administrations in charge of forest management (Autonomous communities (provinces)): Galicia (LU-Lugo, OU-Ourense); Castilla y León (LE-León, ZA-Zamora, PA-Palencia and BU-Burgos); AS-Asturias, CA-Cantabria; País Vasco (BZ-Bizkaia; AR-Araba and GI-Gipuzkoa) (See Figure 1).

[b] Conservation groupings (Figure 1): GID. Galician interior depression; GDR. Galaico-Durienses Range; NAM. Navian-Ancares Mountains; WCR/ NCR/ SCR, West/North/South Cantabrian Range; BCM. Basque-Cantabrian Mountains

[c] Special areas of conservation (Natura 2000; https://natura2000.eea.europa.eu/): ES4190060 (Tejedelo); ES1130007 (Pena Trevinca); ES1120001 (Ancares-Courel); ES4130010 (Sierra Ancares); ES0000210 (Alto Sil); ES4130149 (Omaña); ES4130035 (Valle San Emiliano); ES4130050 (Montaña Central de León); ES4130037 (Hoces de Vegacervera); ES4130003 (Picos de Europa en Castilla y León); ES1200043 (Sierra Sueve); ES0000003 (Picos de Europa); ES1300001 (Liébana); ES1300021 (Valles altos del Nansa y Saja y Alto Campo); ES4140011 (Montaña Palentina); ES4120090 (Embalse Ebro, Monte Hijedo); ES4120030 (Montes Obarenes); ES4120049 – (Valle de Mena); ES2130002 (Ordunte); ES2110009 (Gorbeia); ES2110018 (Arabako hegoaldeko mendilerroak); ES2110015 (Gasteizko mendi garaiak); ES2120001 (Arno); ES2120002 (Aizkorri-Aratz); ES2120006 (Pagoeta); ES2120011 (Aralar); ES2120016 (Aiako harria)

[d] Genetic conservation Units (EUFORGEN; http://www.euforgen.org/species/taxus-baccata/)

968 **Table 2.** Genetic diversity estimates for 46 *Taxus baccata* populations in the Cantabrian-Atlantic

969 Region.

Pop	N	He	Ho	$F_{IS\text{-}INEST}$	A_R	r	C-H (Div)	C-H (Diff)	C-A_R (Div)	C-A_R (Diff)	A_P
LU010	20	0.413	0.462	0.012	1.722	0.481	-0.432	0.596	-0.870	1.549	1
ZA001	20	0.661	0.585	0.033	4.235	0.150	0.066	-0.079	0.226	0.036	1
OU001	30	0.511	0.453	0.014	2.179	0.399	-0.330	0.850	-0.803	1.356	0
LU007	30	0.706	0.603	**0.050**	4.568	0.056	0.178	-0.095	0.212	0.137	3
LU013	22	0.580	0.593	0.013	3.017	0.242	-0.102	-0.076	-0.434	0.263	0
LE001	50	0.690	0.610	0.016	4.487	0.123	0.067	0.023	0.258	0.712	2
LE006	33	0.706	0.627	0.016	4.394	0.101	0.192	-0.026	0.133	0.349	4
LE003	30	0.628	0.587	0.015	3.885	0.171	-0.117	0.085	-0.103	0.480	0
LE007	30	0.674	0.593	0.020	4.404	0.114	0.089	-0.057	0.242	0.399	0
LE005	33	0.657	0.605	0.018	3.955	0.161	0.055	-0.092	0.007	0.440	1
LE011	17	0.717	0.639	0.047	4.377	0.080	0.225	-0.100	0.090	0.292	0
LE015	20	0.606	0.522	**0.061**	3.572	0.177	-0.139	0.217	-0.252	0.358	0
LE014	19	0.721	0.692	0.013	4.424	0.146	0.294	-0.055	0.324	0.341	1
LE016	5	0.620	0.578	0.065	-	0.170	-	-	-	-	1
LE030	3	0.463	0.556	0.032	-	0.323	-	-	-	-	-
LE1018	27	0.681	0.568	0.042	4.519	0.061	0.043	-0.194	0.152	-0.033	0
LE017	31	0.769	0.703	0.026	5.356	0.029	0.347	-0.198	0.562	-0.123	2
LE057	31	0.652	0.592	0.032	4.123	0.129	-0.002	-0.023	0.177	0.615	0
LE1000	18	0.691	0.675	0.023	4.889	0.073	0.122	-0.219	0.392	-0.264	1
AS006	30	0.686	0.575	**0.021**	4.619	0.067	0.135	-0.076	0.236	0.362	3

LE292	20	0.691	0.570	0.071	4.951	0.036	0.135	-0.197	0.395	-0.028	2
LE293	19	0.715	0.759	0.009	4.605	0.118	0.194	-0.116	0.241	0.076	1
CA003	20	0.709	0.653	**0.034**	4.815	0.062	0.164	-0.244	0.361	-0.288	0
CA001	22	0.770	0.714	0.010	5.653	0.038	0.378	-0.203	0.689	-0.172	1
CA004	20	0.654	0.565	0.035	4.470	0.072	0.087	-0.098	0.262	0.267	2
LE050	30	0.677	0.619	0.016	3.923	0.144	0.066	0.027	-0.094	0.427	1
LE023	20	0.733	0.655	0.018	5.171	0.033	0.245	-0.293	0.424	-0.205	1
LE045	31	0.677	0.541	**0.075**	4.406	0.073	0.051	-0.106	0.104	0.184	0
LE046	31	0.577	0.558	0.029	3.523	0.210	-0.202	0.023	-0.248	0.311	0
PA040	13	0.634	0.692	0.015	-	0.292	-	-	-	-	2
PA039	24	0.651	0.614	0.018	3.140	0.237	0.005	0.198	-0.332	1.473	4
BU1042	35	0.684	0.634	0.028	4.504	0.141	0.074	-0.107	0.219	0.233	0
CA002	21	0.603	0.567	0.010	3.450	0.270	-0.140	0.027	-0.281	0.413	1
BU031	31	0.660	0.547	0.016	3.245	0.193	0.018	0.121	-0.385	0.677	1
BU035	32	0.648	0.561	0.028	3.332	0.169	0.077	0.120	-0.288	0.958	0
BZ002	29	0.687	0.545	0.060	4.494	0.100	0.107	-0.048	0.199	0.368	2
BU061	32	0.596	0.552	0.011	3.231	0.250	-0.010	0.232	-0.257	0.133	1
BU064	20	0.750	0.693	0.032	4.662	0.043	0.327	0.096	0.241	0.383	1
BZ001	6	0.696	0.548	**0.109**	-	0.121	-	-	-	-	0
AR002	5	0.733	0.622	0.044	-	-0.082	-	-	-	-	0
AR001	18	0.701	0.509	**0.075**	4.417	0.067	0.207	-0.164	0.135	0.449	1
GI005	3	0.634	0.648	0.034	-	0.173	-	-	-	-	0
GI003	6	0.665	0.552	**0.080**	-	0.022	-	-	-	-	1
GI001	30	0.762	0.607	**0.064**	5.917	-0.044	0.328	-0.205	0.772	0.197	1

GI002	31	0.765	0.651	**0.034**	6.133	-0.013	0.316	-0.253	0.863	-0.095	5
GI004	6	0.625	0.556	**0.121**	-	-0.071	-	-	-	-	0

970

Null allele corrected estimates of expected (He) and observed (Ho) heterozygosity, and inbreeding coefficient ($F_{IS\text{-}INEST}$; in bold P<0.05); A_R (allelic richness; minimum N of 15); r (population average relatedness); C-H and C-A_R: Population contribution for heterozygosis (H) and A_R estimates, respectively, to diversity (Div) and differentiation (Diff) in the Cantabrian-Atlantic Region. (-) No data computed for small sample size (N<15). A_P: Private alleles for each population.

971

972

973 **Table 3**. Analysis of molecular variance (AMOVA) of *Taxus baccata* populations in the
974 Cantabrian-Atlantic Region.

Grouping II Structure-Conservation	***F*-statistic**	**Variance component**	**% Variation**
Among locations (F_{ST})	0.12106***	0.37935	12.10
Among groups (F_{CT})	0.01934***	0.06523	1.93
Among locations within groups (F_{SC})	0.10372***	0.31412	10.17
Within locations		2.96415	87.89

975
976
977
978
979 **Table 4**. Regional distribution of private (A_P) and locally common (LCA) alleles in
980 *Taxus baccata* across Cantabrian-Atlantic Region. N: Total number of individuals.
981

Regional Group	**N**	**LCA**	**A_P**
Galician Inner depression GID	20	5	1
Navian-Asturian Mountains NAM	228	0	10
Galician-Duriense Range GDR	50	2	1
West Cantabrian Range WCR	64	0	2
North Cantabrian Range NCR	238	0	15
South Cantabrian Range SCR	149	0	8
Basque-Cantabrian Mountains BCM	305	0	29

982
983

Figure 1

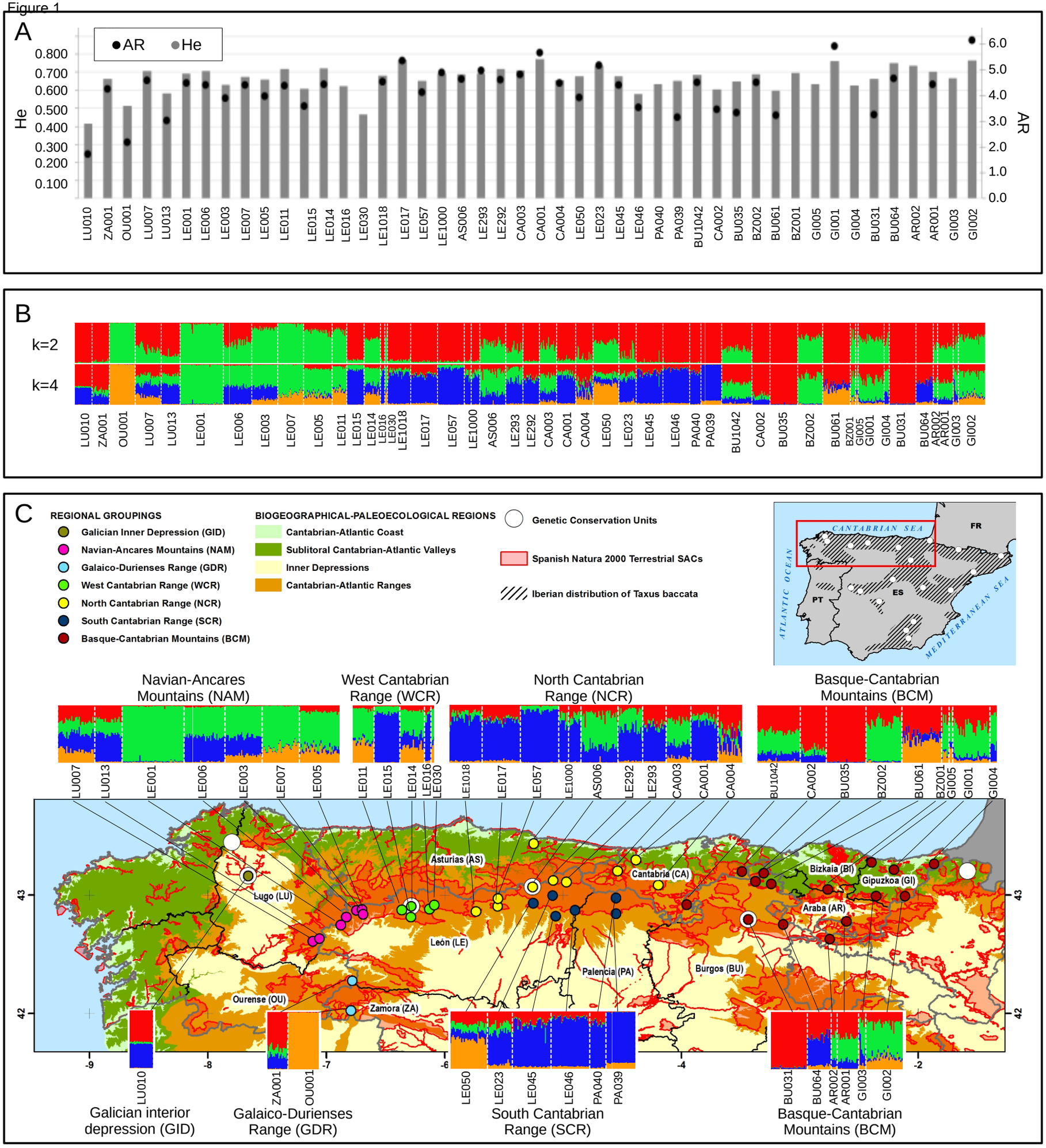
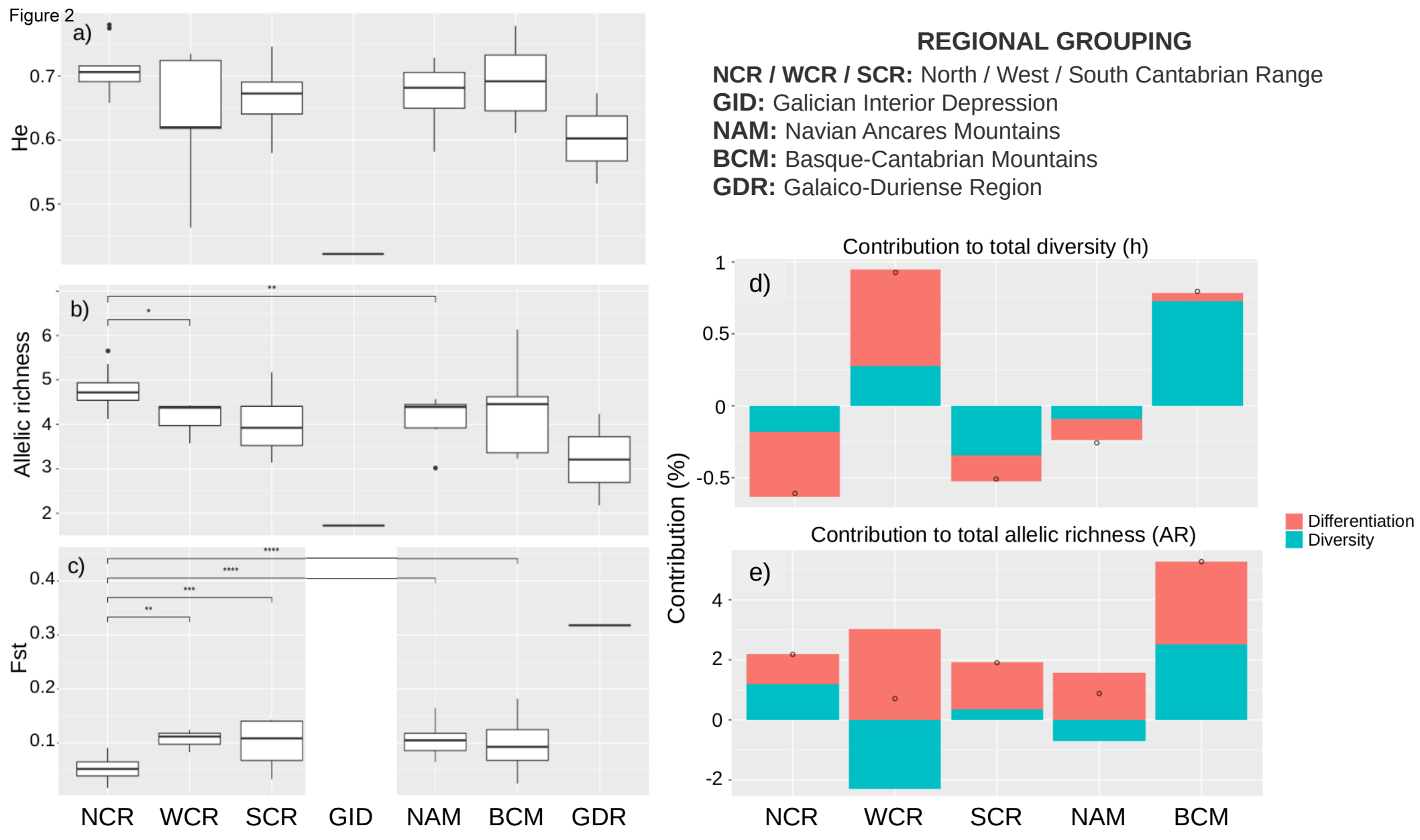

A
AR
He
He
AR
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
6.0
5.0
4.0
3.0
2.0
1.0
0.0
LU010 ZA001 OU001 LU007 LU013 LE001 LE006 LE003 LE007 LE005 LE011 LE015 LE014 LE016 LE030 LE1018 LE017 LE057 LE1000 AS006 LE293 LE292 CA003 CA001 CA004 LE050 LE023 LE045 LE046 PA040 PA039 BU1042 CA002 BU035 BZ002 BU061 BZ001 GI005 GI001 GI004 BU031 BU064 AR002 AR001 GI003 GI002
B
k=2
k=4
LU010 ZA001 OU001 LU007 LU013 LE001 LE006 LE003 LE007 LE005 LE011 LE015 LE014 LE016 LE030 LE1018 LE017 LE057 LE1000 AS006 LE293 LE292 CA003 CA001 CA004 LE050 LE023 LE045 LE046 PA040 PA039 BU1042 CA002 BU035 BZ002 BU061 BZ001 GI005 GI001 GI004 BU031 BU064 AR002 AR001 GI003 GI002
C
REGIONAL GROUPINGS
Galician Inner Depression (GID)
Navian-Ancares Mountains (NAM)
Galaico-Durienses Range (GDR)
West Cantabrian Range (WCR)
North Cantabrian Range (NCR)
South Cantabrian Range (SCR)
Basque-Cantabrian Mountains (BCM)
BIOGEOGRAPHICAL-PALEOECOLOGICAL REGIONS
Cantabrian-Atlantic Coast
Sublitoral Cantabrian-Atlantic Valleys
Inner Depressions
Cantabrian-Atlantic Ranges
Genetic Conservation Units
Spanish Natura 2000 Terrestrial SACs
Iberian distribution of Taxus baccata
CANTABRIAN SEA
FR
ATLANTIC OCEAN
PT
ES
MEDITERRANEAN SEA
Navian-Ancares Mountains (NAM)
West Cantabrian Range (WCR)
North Cantabrian Range (NCR)
Basque-Cantabrian Mountains (BCM)
LU007 LU013 LE001 LE006 LE003 LE007 LE005
LE011 LE015 LE014 LE016 LE030
LE1018 LE017 LE057 LE1000 AS006 LE292 LE293 CA003 CA001 CA004
BU1042 CA002 BU035 BZ002 BU061 BZ001 GI005 GI001 GI004
Asturias (AS)
Cantabria (CA)
Bizkaia (BI)
Gipuzkoa (GI)
Lugo (LU)
Araba (AR)
León (LE)
Palencia (PA)
Burgos (BU)
Ourense (OU)
Zamora (ZA)
43
42
-9
-8
-7
-6
-4
-2
LU010
ZA001 OU001
LE050 LE023 LE045 LE046 PA040 PA039
BU031 BU064 AR002 AR001 GI003 GI002
Galician interior depression (GID)
Galaico-Durienses Range (GDR)
South Cantabrian Range (SCR)
Basque-Cantabrian Mountains (BCM)

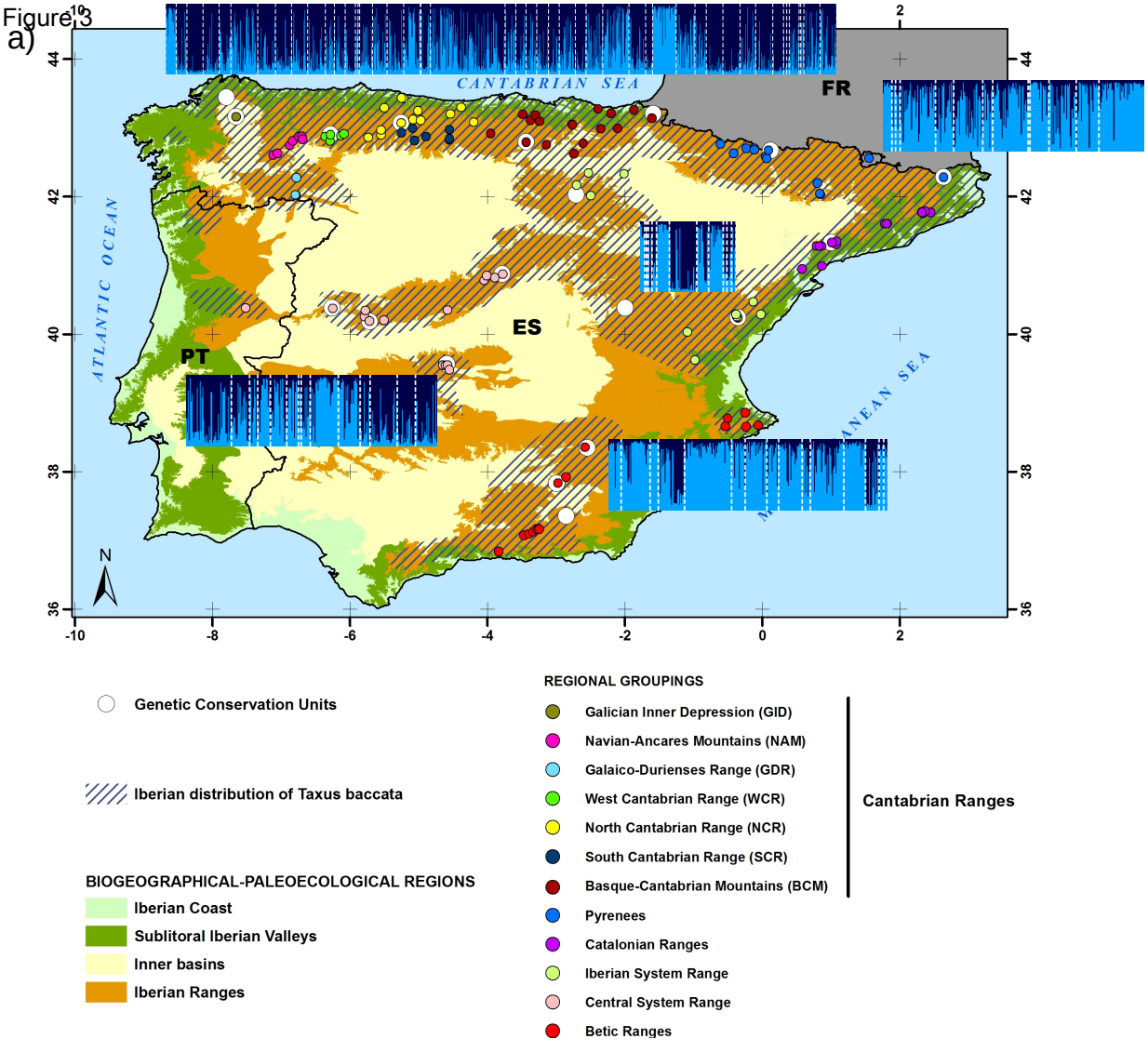
Figure 2

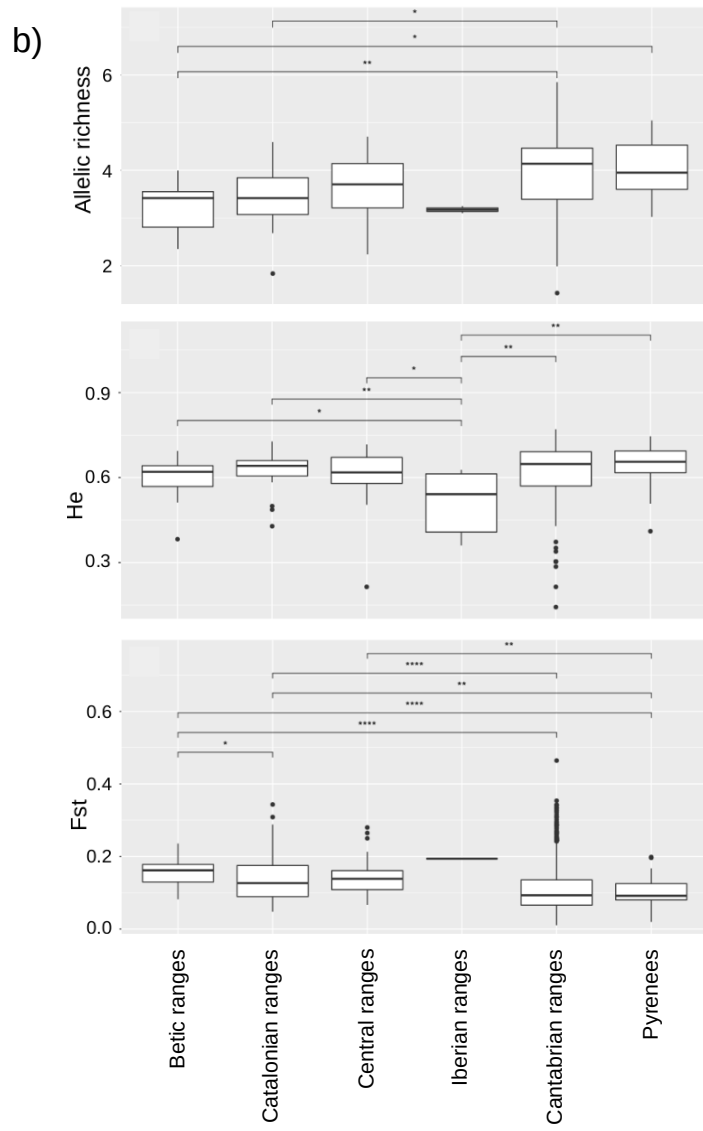


REGIONAL GROUPING

NCR / WCR / SCR: North / West / South Cantabrian Range
GID: Galician Interior Depression
NAM: Navian Ancares Mountains
BCM: Basque-Cantabrian Mountains
GDR: Galaico-Duriense Region

Figure 3


a)
CANTABRIAN SEA
FR
ES
PT
ATLANTIC OCEAN
MEDITERRANEAN SEA
N
Genetic Conservation Units
Iberian distribution of Taxus baccata
BIOGEOGRAPHICAL-PALEOECOLOGICAL REGIONS
Iberian Coast
Sublitoral Iberian Valleys
Inner basins
Iberian Ranges
REGIONAL GROUPINGS
Galician Inner Depression (GID)
Navian-Ancares Mountains (NAM)
Galaico-Durienses Range (GDR)
West Cantabrian Range (WCR)
North Cantabrian Range (NCR)
South Cantabrian Range (SCR)
Basque-Cantabrian Mountains (BCM)
Pyrenees
Catalonian Ranges
Iberian System Range
Central System Range
Betic Ranges
Cantabrian Ranges


b)
Allelic richness
He
Fst
Betic ranges
Catalonian ranges
Central ranges
Iberian ranges
Cantabrian ranges
Pyrenees

No conflict of interest

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Francesco Maroso: Formal analysis, Investigation, Writing-original draft, review and editing; **Manuel Vera**: Formal analysis, Writing-review and editing; **Javier Ferreiro**: Investigation, Resources, Writing-review and editing; **Miquel Riba**: Formal analysis, Resources, Writing-review and editing; **María Mayol**: Formal analysis, Resources, Writing-review and editing; **Pablo Ramil**: Funding acquisition, Resources, Writing-review and editing; **Paulino Martínez**: Conceptualization, Formal analysis, Writing-review and editing, Supervision; **Carmen Bouza**: Conceptualization, Formal analysis; Investigation, Writing-original draft, review and editing, Supervision.

Figure S1


Click here to access/download
Supplementary Material
FigureS1.pdf

Figure S2

Figure S3

Figure S4

Figure S5

Figure S6

Table S1

Click here to access/download

Supplementary Material

Table_S1-Rev2.xlsx

Table S2

Table S3

Table S4

Click here to access/download
Supplementary Material
Table_S4-Rev2.xlsx

MS_tracked changes