1 **Molecular evidence of introgression of Asian germplasm into a natural *Castanea sativa***
2 **forest in Spain**
3
4 **Francisco Alcaide[1], Alejandro Solla[1], Beatriz Cuenca[2], M. Ángela Martín[3*]**
5
6 [1]Faculty of Forestry, Institute for Dehesa Research (INDEHESA), University of
7 Extremadura, Avenida Virgen del Puerto 2, 10600 Plasencia, Spain
8 [2]Maceda Nursery, Tragsa-SEPI group, Carretera de Maceda a Baldrei km 2, 32700 Maceda,
9 Ourense, Spain
10 [3]Department of Genetics, ETSIAM, University of Córdoba, Campus de Rabanales, 14071
11 Córdoba, Spain
12 [*]Corresponding author Tel.: +34-957-218505; E-mail: angela.martin@uco.es
13
14 **Running title** Introgression of Asiatic germplasm into *Castanea sativa*
15

Abstract

Evidence has been documented in Europe of introgression, the transfer of genetic material, between Asian chestnut species (*Castanea crenata* and *C. mollissima*) and the native *C. sativa* through spontaneous hybridization and subsequent backcrossing. However, the extent of this introgression has not been monitored in a particular forest stand or in adult and juvenile trees simultaneously. We assessed introgression in a natural *C. sativa* forest in northwest Spain and developed a reliable method to detect the presence of Asian germplasm of *Castanea* spp. A total of 34 adult trees and 42 saplings were genotyped at 13 SSRs in a forest where ink-disease-resistant *C. crenata* and *C. mollissima* seedlings had been planted in the 1940s. The 13 SSR loci selected were highly polymorphic and 115 different alleles were detected for the individuals sampled. Bayesian clustering identified two groups for *C. sativa* and one group each for *C. mollissima* and *C. crenata*. Within the forest, 70.6% of adults and 28.6% of juveniles were classified as pure *C. sativa.* Most juveniles were *C. sativa* × *C. crenata* (> 40%) and ca. 10% were *C. crenata* × *C. mollissima* hybrids. Six new alleles private to *C. crenata* are reported here. The study quantifies, for the first time in Europe, introgression of Asian germplasm into a natural *C. sativa* forest. It also examines the extent of introgression in offspring and provides a method to detect exotic germplasm in *C. sativa* forests. We discuss why the forest studied may benefit from transfer of alleles involved in ink-disease resistance and why introgression will be detrimental to drought tolerance in offspring.

Keywords: interspecific hybridization, genetic differentiation, microsatellite markers, monitoring system, *Phytophthora cinnamomi*

40 **Introduction**

41 Introgression is the transfer of a gene from one species into the gene pool of another species
42 by hybridization and backcrossing. Gene introgression in forests has been widely
43 acknowledged (Brunet *et al*., 2013; Geraldes *et al*., 2014; Buck *et al*., 2020; Crowl *et al*.,
44 2020) and may be referred to as 'adaptive' when the foreign genes incorporated are
45 maintained by natural selection, leading to positive fitness consequences in the recipient
46 species (Suarez-Gonzalez *et al*., 2018; Burgarella *et al*., 2019). Favourable effects of
47 introgression in forests include increased genetic variation, a wider range of habitats where
48 trees can become established, and transfer of alleles involved in disease resistance (Brunet
49 *et al*., 2013; Suarez-Gonzalez *et al*., 2018). Introgression usually facilitates rapid adaptation
50 of tree populations to new environments (Hamilton *et al*., 2015). Conversely, introgression
51 can have negative effects on tree species, such as genetic erosion of native populations
52 (Johnson *et al*., 2016) and hindering conservation attempts (Edmands, 2007).
53

54 The genus *Castanea* includes six economically and environmentally important tree
55 species, all distributed in temperate deciduous forests of the northern hemisphere (Poljak *et*
56 *al*., 2017). The sweet chestnut (*Castanea sativa* Mill.) is the only native species of the genus
57 *Castanea* in Europe, and is widely distributed and cultivated for its edible nuts and for wood
58 (Figure 1a) (Míguez-Soto *et al*., 2016; Camisón *et al*., 2021). *Castanea sativa* reproduces
59 almost exclusively through seeds by wind pollination (Manino *et al*., 1991). When cultivated
60 near other chestnut species, including Chinese chestnut (*C. mollissima* Blume, native to
61 China, Taiwan and Korea) and Japanese chestnut (*C. crenata* Sieb. & Zucc., native to Japan
62 and Korea), *C. sativa* easily cross-pollinates with them to form hybrids (Jaynes, 1964;
63 Fernández-Cruz and Fernández-López, 2016). Chinese and Japanese chestnut trees have
64 been found to have high resistance to *Phytophthora cinnamomi* Rands, the causal agent of

65 ink disease (Santos *et al*., 2015; Camisón *et al*., 2019b; Perkins *et al*., 2019). Because of this,
66 breeding programmes were developed in France, Spain and Portugal (Hennion 2010;
67 Fernández-López, 2011; Santos *et al*., 2015; Míguez-Soto *et al*., 2016; Fernandes *et al*.,
68 2020) to confer protection to *C. sativa* by transferring resistant genes from Asian chestnuts.
69 The main objective of these chestnut breeding programmes was to produce vigorous *P.*
70 *cinnamomi*-resistant hybrids as rootstock, with good compatibility with traditional *C. sativa*
71 cultivars (Pereira-Lorenzo and Fernández-López, 1997) or as direct nut or timber producers
72 (Fernández-López, 2011; Míguez-Soto *et al*., 2016).

73

74 From 1917 to 1945, *P. cinnamomi*-resistant seedlings of *C. crenata* and *C. mollissima*
75 were extensively planted in northern Spain (Elorrieta and Artaza, 1949; Fernández-López,
76 2011). However, introduced Asian species did not adapt well to Atlantic climate conditions,
77 and their different agronomic and forestry characteristics made them unsuitable for replacing
78 disease-susceptible *C. sativa* trees (Lafitte, 1946). The first controlled crosses of these
79 chestnut species with the indigenous *C. sativa* were carried out in Spain in 1926, and since
80 the 1940s artificial *C. crenata* × *C. sativa* hybrids resistant to *P. cinnamomi* have been
81 cultivated and marketed in the north of the country (Elorrieta and Artaza, 1949; Fernández-
82 López, 2011). As a consequence of these plantations, the presence of Asian alleles has been
83 reported in several wild populations of *C. sativa* in northwest Spain (Fernández-Cruz and
84 Fernández-López, 2016). However, the extent of introgression of Asian chestnut germplasm
85 has not been monitored intensely in a particular forest stand or in adult and juvenile trees
86 simultaneously.

87

88 *Castanea crenata* × *C. sativa* and *C. mollissima* × *C. sativa* hybrid clones are not
89 suitable for restoring and replacing decimated *C. sativa* populations, because they are

90 sensitive to late frost, do not tolerate drought stress, and usually have low timber value and
91 poor compatibility for grafting with *C. sativa* varieties (Elorrieta and Artaza, 1949; Pereira-
92 Lorenzo and Fernández-López, 1997; Míguez-Soto *et al*., 2019). In 2015, a breeding
93 initiative was started at the Faculty of Forestry, University of Extremadura, to select native
94 pure *C. sativa* trees resistant to *P. cinnamomi* and/or tolerant to drought and heat stress.
95 Initial results revealed that selection of *C. sativa* trees for drought tolerance was possible
96 (Alcaide *et al*., 2019; Camisón *et al*., 2020, 2021). Moreover, surprisingly high *P. cinnamomi*
97 resistance was observed in chestnut seedlings from the Bergondo population (Alcaide *et al*.,
98 2020), in northwest Spain. It was hypothesized that resistant genes from Asian germplasm
99 was behind this resistance (Alcaide *et al*., 2020), although the hypothesis was not tested
100 experimentally.
101
102 It is extremely difficult to identify interspecific chestnut hybrids based on morphology
103 and phenology traits (Elorrieta and Artaza, 1949; Pereira-Lorenzo *et al*., 2010; Fernández-
104 López, 2011). However, genomic tools are capable of properly examining hybridization and
105 introgression. In particular, microsatellites or simple sequence repeats (SSRs), probably the
106 most frequently used molecular markers to assess diversity in *C. sativa* (Martín *et al*., 2012,
107 2017; Mattioni *et al*., 2017; Jarman *et al*., 2019; Pereira-Lorenzo *et al*., 2020; Castellana *et*
108 *al*., 2021), have proven useful in identifying chestnut interspecific hybrids (Pereira-Lorenzo
109 *et al*., 2010; Fernández-López, 2011; González *et al*., 2011; Fernández-Cruz and Fernández-
110 López, 2012, 2016; Pereira-Lorenzo *et al*., 2017). The objectives of the present study were
111 (1) to assess introgression in a natural *C. sativa* forest, and (2) to develop a fast and reliable
112 method to detect the presence of Asian germplasm of *Castanea* spp.
113
114 **Methods**

115 *Study site and sample collection*

116 The study was conducted in Bergondo, Galicia, Spain (43°18'N, 8°14'W; 68 m a.s.l.) (Figure
117 1), 1.5 km from the Atlantic coast. The forest covers approximately 40 ha and is
118 characterized by a dense layer of *C. sativa* trees, the presence of *Quercus robur* and
119 *Eucalyptus globulus* (< 10%), and an understory of *Rubus ulmifolius* and *Pteridium*
120 *aquilinum*. Adult trees were 40 to 80 years old and up to 15 m tall. The forest has an Atlantic
121 climate with a mean annual rainfall of 1,070 mm, mainly from October to April (climate
122 data from climate-data.org). Mean minimum and maximum temperatures are in January
123 (6.3°C) and August (23.2°C), respectively. The area is traditionally used for mushroom and
124 nut gathering. The forest was selected because (i) it contains abundant healthy *C. sativa* adult
125 trees, (ii) *C. crenata* and *C. mollissima* seedlings were probably planted during the 1940s,
126 (iii) an experimental plot including *C. sativa* × *C. crenata* hybrids was established in the
127 1940s in Guísamo, 3 km from the forest, (iv) natural regeneration occurs in most areas of
128 the forest, and (v) information was available about drought and *P. cinnamomi* tolerance in
129 offspring of several trees (Alcaide *et al*., 2019, 2020).

130

131 To evaluate the presence of Asian chestnut germplasm in the uneven-aged forest, 100
132 adult chestnut trees (> 10 cm in diameter) in the middle of the forest were randomly
133 preselected. Non-healthy trees (apparently infected by *Cryphonectria parasitica* and
134 *Phytophthora* spp.) and trees less than 60 m apart were not included. The final 34 adult trees
135 selected (Figure 1b and Table 1) were distant from each other to minimize the chances of
136 intercrossing. In a small area in which recruitment was abundant and tree crowns did not
137 overlap, nine out of the 34 adult trees selected were marked. Within a circle of about 5-7 m
138 in radius around the trunk of each marked tree (Bacilieri *et al*., 1996), the offspring of these
139 trees was also selected, comprising three to five saplings per adult tree (Figure 1b;

140 Supplementary Table S1). Digging confirmed that the juvenile material, 0.3 to 2.5 m high,
141 did not include root suckers. In total, 42 open pollinated saplings were selected. Additionally,
142 as a reference of exotic germplasm, accessions of *C. mollissima* and *C. crenata* were used
143 (Table 1). From the lower crown of each tree, five healthy green leaves were sampled.
144
145 *DNA isolation and microsatellite analysis*
146 Total genomic DNA was extracted from 18-20 mg of lyophilized leaves according to the
147 QIAGEN DNeasy^{TM} Plant Mini Kit protocol. To assess introgression in *C. sativa*, 13 neutral
148 and polymorphic microsatellite (SSR) markers (*CsCAT1*, *CsCAT2*, *CsCAT3*, *CsCAT6*,
149 *CsCAT14*, *CsCAT16*, *CsCAT41*, *EMCs2*, *EMCs14*, *EMCs15*, *EMCs25*, *EMCs38* and *OCI*)
150 belonging to different linkage groups (Barreneche *et al*., 2004) were used. These markers
151 were developed for *C. sativa* (Buck *et al*., 2003; Marinoni *et al*., 2003; Gobbin *et al*., 2007)
152 and had been used in previous studies (Fernández-Cruz and Fernández-López, 2016; Martín
153 *et al*., 2017; Mattioni *et al*., 2017), including the European chestnut database (Pereira-
154 Lorenzo *et al*., 2017) (Table 2).
155
156 The following four multiplex-PCR mixtures were designed: (A), comprising *CsCAT1*,
157 *CsCAT2*, *CsCAT14* and *EMCs25* primers; (B), comprising *CsCAT3*, *CsCAT6*, *CsCAT16* and
158 *EMCs38* primers; (C), comprising *CsCAT41*, *EMCs2* and *OCI* primers; and (D), comprising
159 *EMCs14* and *EMCs15* primers (Table 2). For the 13 SSR primer pairs, 5' end of forward
160 primers were labelled with fluorescent dye tags (6-FAM or VIC or NED) (Applied
161 Biosystem^{TM}). PCR reactions were performed using a GeneAmp 2700 Thermal Cycler
162 (Applied Biosystem^{TM}). PCR reaction mixture comprised 12.5 µL total volume containing
163 4 µL template DNA (~20 ng), 6.25 µL 2x Type-it Multiplex PCR Master Mix, 1.25 µL 10X
164 Primer Mix (2 µM of each primer) and 1 µL RNase-free water, following the Type-It®

165 Microsatellite PCR Kit (QIAGEN, Valencia, CA, USA) protocol. Cycling parameters were
166 15 min initial denaturation at 95°C, 30 cycles of 30 s at 94°C, 90 s at 57°C, 1 min elongation
167 at 72°C, and a final step of 30 min extension at 72°C. Amplification products (1 μL) were
168 mixed with 20 μL formamide and 0.3 μL GeneScan 500 LIZ dye size standard (Applied
169 Biosystem™) and denatured at 95°C for 5 min. Samples were run on an ABI Prism 3130xL
170 Avant Genetic Analyzer DNA sequencer (Applied Biosystem™). The resulting raw data
171 were collected by applying the GeneMapper v. 4.0 software (Applied Biosystem™). Alleles
172 were determined by automated binning and checked by visual inspection.

173

174 *Genetic diversity of the forest*

175 To assess the genetic diversity of the forest, the extracted DNA of all individuals sampled
176 (adults and offspring) was used. The number of total alleles per locus (Na), allele frequency
177 distribution, observed and expected heterozygosity (H_O and H_E), and unbiased expected
178 heterozygosity (uH_E) were estimated with GenAlEx v. 6.503 software (Peakall and Smouse,
179 2012). The polymorphic information content (PIC) values were estimated using MS Tools
180 for Excel (Park, 2001).

181

182 *Assignment of individuals and detection of introgression*

183 STRUCTURE software applied to microsatellite data appears to be the best and most
184 efficient tool to assess introgression and the presence of hybrids in wild populations (Brunet
185 *et al*., 2013; Hamilton *et al*., 2015; Fernández-Cruz and Fernández-López, 2016; An *et al*.,
186 2017). To cluster all the individuals selected from the chestnut population (Table 1) into
187 similar groups, the STRUCTURE v. 2.3.4 software (Falush *et al*., 2003) was used. To
188 estimate the appropriate optimal number of clusters (*K*), we ran STRUCTURE, varying the
189 value for *K* from 1 to 5. Twenty independent runs for each *K* value and a burn-in period of

190 10,000 steps followed by 10^5 Markov chain Monte Carlo replicates were used. The analysis
191 relates trees to each *K* cluster through a membership coefficient (*Q*-value). Individuals with
192 a high probability of belonging to one of the genetic clusters ($Q \geq 0.90$) were classified as
193 pure species. Individuals with $Q < 0.90$ for one genetic cluster and $Q \leq 0.10$ for each of the
194 remaining clusters were also considered pure species, according to Lepais *et al*. (2009).
195 Individuals with $0.90 > Q > 0.10$ for multiple clusters were considered hybrids. To identify
196 the number of clusters (*K*) that best explained the data, the rate of change on L(*K*) (ΔK)
197 between successive *K* values was calculated according to Evanno *et al*. (2005) using
198 STRUCTURE HARVESTER software (Earl and vonHoldt, 2012). The 20 runs for each
199 simulation were averaged using CLUMPP v. 1.1.2 software (Jakobsson and Rosenberg,
200 2007) and represented graphically with DISTRUCT (Rosenberg, 2004).
201
202 To confirm and quantify Asian-to-native gene flow, i.e., allele introgression, private
203 alleles exclusive to *C. sativa*, *C. mollissima* and *C. crenata* were identified using GenAlEx
204 v. 6.503 software (Peakall and Smouse, 2012). Identification of these alleles was carried out
205 firstly within the three groups of pure *C. sativa*, *C. mollissima* and *C. crenata* accessions
206 only, according to STRUCTURE. The presence of the private alleles was then checked in
207 all 76 individuals sampled (34 adult trees and 42 saplings).
208
209 To determine in offspring whether introgression was biased towards the native or the
210 exotic species, we examined the level of genetic differentiation of the hybrid saplings
211 relative to the parental species (Brunet *et al*., 2013). Pairwise F_{ST} values between *C. sativa*
212 adults, *C. crenata* adults, and *C. sativa* × *C. crenata* hybrid saplings were calculated using
213 the AMOVA method of Arlequin 3.11 software (Excoffier *et al*., 2005). Significances were
214 obtained with permutation tests with 10,000 iterations.

215

216 **Results**

217 *Suitability of the selected SSR markers*

218 The 13 SSR loci selected were highly polymorphic (Table 2). For the 76 individuals
219 sampled, 115 different alleles were detected. Each locus had on average 8.9 alleles. The most
220 informative loci were *CsCAT2* and *CsCAT6* with unbiased expected heterozygosity (uH_E) of
221 0.83 and PIC values of 0.80 (Table 2). The least discriminating locus was *EMCs14* (Table
222 2). Allele frequencies were distributed unevenly within the loci. Thirty-six alleles, observed
223 in *CsCAT1*, *CsCAT2*, *CsCAT14*, *CsCAT3*, *CsCAT6*, *CsCAT16*, *EMCs38*, *CsCAT41* and
224 *EMCs15* loci, were classified as rare, with frequencies below 5%. The selected SSR markers
225 were able to differentiate the three species (*C. sativa*, *C. mollissima* and *C. crenata*) (Figure
226 2). The polymorphic information content (PIC) values in nine loci were higher than 0.70
227 (Table 2), indicating the utility of markers for linkage analysis.

228

229 *Genetic structure of the adult chestnut population and relationship between individuals*

230 STRUCTURE software identified $K = 4$ as the most probable genetic division of individuals
231 of the chestnut forest, with the strongest support in terms of log-likelihood (Figure 2). Based
232 on $K = 4$, Bayesian clustering identified two groups for *C. sativa* individuals (genetic *Q*1 and
233 *Q*2 clusters, in orange and green, respectively; Figure 2), one group for *C. mollissima* trees
234 (*Q*3, in yellow), and one group for *C. crenata* trees (*Q*4, in brown; Figure 2).

235

236 The admixture coefficient (*Q*) made it possible to assign adult individuals to the three
237 *C. sativa*, *C. mollissima* and *C. crenata* species (Figure 2 and Supplementary Table S2).
238 Within the chestnut population studied, 24 trees, i.e., 70.6% of the adult individuals, were
239 classified as pure *C. sativa* species. Of these, 17 had a high probability of belonging to *C.*

240 *sativa* ($Q \geq 0.90$) and seven showed minor signs of introgression (i.e., they had $Q < 0.90$ for
241 *C. sativa* clusters and $Q \leq 0.10$ for either *C. crenata* or *C. mollissima* clusters). Five adult
242 individuals of the population were *C. crenata* (14.7%) (Supplementary Table S2). Four adult
243 individuals were identified as *C. sativa* × *C. crenata* hybrids (11.8%), and one individual
244 was identified as a *C. sativa* × *C. crenata* × *C. mollissima* hybrid (2.9%) (Supplementary
245 Table S2).
246
247 *Introgression of Asian germplasm in offspring*
248 Twelve chestnut saplings were classified as pure *C. sativa*, i.e., 28.6% of the offspring
249 studied (Supplementary Table S2). Nine saplings were classified as pure *C. crenata* (21.4%).
250 Seventeen saplings were classified as *C. sativa* × *C. crenata* hybrids (40.5%), of which five
251 were predominantly *C. sativa* ($Q < 0.90$ and $Q > 0.10$ for *C. sativa* and *C. crenata* clusters,
252 respectively) and six were predominantly *C. crenata* ($Q < 0.90$ and $Q > 0.10$ for *C. crenata*
253 and *C. sativa* clusters, respectively) (Supplementary Table S2). Four saplings were classified
254 as *C. crenata* × *C. mollissima* hybrids (9.5%), all predominantly *C. crenata* ($Q < 0.90$ and
255 $Q > 0.10$ for *C. crenata* and *C. mollissima* clusters, respectively) (Supplementary Table S2).
256
257 Fifteen private alleles were identified, nine specific to *C. crenata* from *CsCAT1*,
258 *CsCAT6*, *CsCAT16*, *EMCs38*, *CsCAT41B* and *OCI* markers, four specific to *C. mollissima*
259 from *CsCAT41B* and *EMCs14* markers, and two present in the two Asian species only (Table
260 3). In total, 26.5% of adults and 54.8% of saplings (Table 4) presented private alleles of
261 Asian species. Among the nine mother trees whose offspring leaves were sampled, three
262 individuals (A_{13}, A_{14} and A_{15}) classified as pure *C. sativa* did not have private alleles of
263 Asian species. However, in the offspring of these individuals, three out of 13 saplings
264 (23.1%) showed private alleles of *C. crenata* (Table 4). Within the offspring of A_{11} and A_{12}

265 hybrid trees, four out of 10 saplings (40%) showed private alleles of *C. crenata* (Table 4).
266 Within the offspring of A_6, A_7, A_8 and A_9 *C. crenata* trees, 14 out of 19 saplings (73.7%)
267 showed private alleles of *C. crenata* (Table 4).

268

269 The pattern of introgression in offspring of the chestnut stand was not biased towards
270 *C. sativa* or *C. crenata*. Pairwise genetic differentiation (F_{ST}) based on the 13 microsatellite
271 loci revealed that hybrid saplings were similarly differentiated from each of the two parental
272 species (F_{ST} = 0.03 with *C. sativa* and 0.02 with *C. crenata*; $P < 0.05$).

273

274 **Discussion**

275 This study obtains three significant results: firstly, it quantifies, for the first time in Europe,
276 introgression of Asian germplasm into a natural *C. sativa* forest; secondly, it examines the
277 extent of this introgression in offspring, which is much higher than the extent of
278 hybridization detected in the adult trees; and thirdly, it provides a method to detect exotic
279 germplasm in *C. sativa* forests and identify pure *C. sativa* individuals.

280

281 The presence of *C. sativa* × *C. crenata* and *C. sativa* × *C. crenata* × *C. mollissima*
282 hybrids in Bergondo forest is in accordance with other studies reporting interspecific
283 chestnut hybrids and backcrossing of Asian germplasm with *C. sativa* in forests (Pereira-
284 Lorenzo *et al*., 2010; Fernández-López, 2011; González *et al*., 2011; Fernández-Cruz and
285 Fernández-López, 2016). The chestnut adults of Asian origin were probably planted during
286 the 1940s, as reported elsewhere (Elorrieta y Artaza, 1949; Fernández-López, 2011). Asian
287 germplasm in the area may also have come from pollen transported from chestnut plantations
288 of qualified plant material, which frequently include *C. crenata* × *C. sativa* clones for timber
289 and fruit production, and *C. mollissima* × *C. sativa* clones for timber production and as

290 rootstock for grafting (Ramiro Martínez, 'O Pomariño' nursery, A Coruña, Spain, pers.
291 comm.).
292
293 The extent of hybridization detected is substantiated by the high *P. cinnamomi*
294 tolerance of one-year-old seedlings from Bergondo (Alcaide *et al*., 2020). Resistance to
295 *Phytophthora* sp. is usually higher in *C. crenata* × *C. sativa* hybrids than in *C. sativa*
296 (Miranda-Fontaíña *et al*., 2007; Camisón *et al*., 2019b). Using SSR markers, González *et al*.
297 (2011) observed that vigorous chestnut trees selected in the field as free of ink disease were
298 mostly *C. crenata* × *C. sativa*. Human introduction of exotic trees planted to replace native
299 tree species decimated by an epidemic disease creates considerable potential for
300 hybridization between native and introduced species. Hybridization has also been reported
301 between native elm species (*Ulmus minor* in Europe and *U. americana* in the United States)
302 and the Asian *U. pumila*, introduced in various countries in response to the Dutch elm disease
303 pandemics (Brunet *et al*., 2013; Solla *et al*., 2014). A strong pattern of biased introgression
304 towards *U. pumila* was attributed to the low abundance of healthy *U. minor* across Spain and
305 of healthy *U. rubra i*n the United States (Brunet *et al*., 2013).
306
307 Introgression is an important source of genetic variation. This source is adaptive when
308 introgressed alleles are maintained by selection and become fixed (Suarez-Gonzalez *et al*.,
309 2018). Our findings suggest that chestnut hybrid saplings from Bergondo will, in the long
310 term, tolerate ink disease better than purebred *C. sativa* adults. Because Bergondo is a
311 *Phytophthora*-infested site (Paloma Abad-Campos, pers. comm.), the forest will benefit
312 from the transfer of alleles involved in resistance. In contrast, introgression in Bergondo
313 forest leads to genetic integrity loss of *C. sativa* and may have permitted introduction of
314 undesirable adaptive traits such as increased susceptibility to drought and frost. Exotic

315 chestnut species are known to be less tolerant than *C. sativa* to warm and dry conditions
316 (Pereira-Lorenzo *et al*., 2010). Probably because *C. crenata* flowers 20-30 days earlier than
317 *C. sativa* (Elorrieta y Artaza, 1949), it does not tolerate late frost (Pereira-Lorenzo *et al*.,
318 2010). After subjecting chestnut plants from several sites to water deprivation, seedlings
319 from Bergondo showed the highest 'leaf wilting' values of all sites (Alcaide *et al*., 2019). In
320 consequence, Bergondo forest will probably suffer the effects of climate change more than
321 forests of pure *C. sativa* trees. Introgression of adaptive genetic variation has been well
322 documented in a number of plant species (Suarez-Gonzalez *et al*., 2018; Burgarella *et al*.,
323 2019), but less in trees. The impact of both *P. cinnamomi* and drought stressors in Bergondo
324 trees will probably generate conflicting selection pressures related to the amount of native
325 *vs* Asian germplasm, and may in turn constrain regeneration success at the seedling stage
326 (Camisón *et al*., 2019a). Given the partial overlap of flowering periods of *C. sativa*, *C.*
327 *mollissima* and *C. crenata* (Jaynes, 1964; Fernández-Cruz and Fernández-López, 2016), it
328 is expected that the extent of hybridization may increase over time. The susceptibility of
329 Asian chestnut species to drought, heat and early frosts will probably contribute, in the long
330 term, to a reduction in the range of habitats where chestnut hybrids can become established
331 in Spain compared to the *C. sativa* native range. To check this, regular inventorying and
332 monitoring of the genetic structure in chestnut forests should be carried out (Schwartz *et al*.,
333 2007). In this respect, together with the 13 SSRs selected in this study, it should be useful to
334 include SNP markers recently developed in *Castanea* spp. due to their high abundance and
335 relatively even distribution across the genome (Ji *et al*., 2018; Kang *et al*., 2019).
336

337 For adult and sapling trees from Bergondo classified as pure *C. sativa* by
338 STRUCTURE, mean observed heterozygosity (H_O) and expected heterozygosity (H_E) were
339 0.70 and 0.73, respectively, higher than heterozygosity values described in other *C. sativa*

340 populations in northwest Spain (Martín *et al*., 2012; Fernández-Cruz and Fernández-López,
341 2016) and similar to those reported for the chestnut populations of the whole country (Martín
342 *et al*., 2012; Pereira-Lorenzo *et al*., 2019). Likewise, the 13 SSR markers used showed an
343 average of 7.3 alleles per locus for Bergondo, higher than that detected in *C. sativa* forests
344 in Galicia (Martín *et al*., 2012; Fernández-Cruz and Fernández-López 2016). In northwest
345 Spain, *C. sativa* forests are particularly rich in genetic diversity and exhibit a different gene
346 pool from forests in the central and southern Iberian Peninsula (Fernández-Cruz and
347 Fernández-López, 2016; Míguez-Soto *et al*., 2019). Genetic diversity richness can be
348 explained by the existence of several glacial refugees in southern areas of Europe as a
349 consequence of cold periods during the Pleistocene (Krebs *et al*., 2004; Mattioni *et al*., 2013;
350 Castellana *et al*., 2021). Genetic diversity in Galicia is also attributed to a strong
351 domestication process of the species in this region (Fernández-Cruz and Fernández-López,
352 2016; Míguez-Soto *et al*., 2019; Pereira-Lorenzo *et al*., 2019).
353
354 It is noteworthy that 54.8% of saplings showed private alleles of the Asian species,
355 evidencing an increased exotic germplasm presence in offspring in comparison to adult trees
356 (26.5%). Ten private alleles were found in offspring, eight of which (176 for *CsCAT1* locus,
357 138 for *CsCAT6*, 239 and 257 for *EMCs38*, 228 and 232 for *CsCAT41B*, and 151 and 165
358 for *OCI*) were exclusive to *C. crenata* and two (186 for *CsCAT1* and 136 for *CsCAT14* loci)
359 were found in both *C. crenata* and *C. mollissima* species. Allele 138 was previously reported
360 for *C. crenata* (González *et al*., 2011), alleles 186 and 136 were previously reported for both
361 *C. crenata* and *C. mollissima* species (González *et al*., 2011), and allele 232 was previously
362 reported by Pereira-Lorenzo *et al*. (2010) although not as private. Alleles 176, 239, 257, 228,
363 151 and 165 are first reported here as private to *C. crenata*. In some saplings (e.g., $S_{6\text{-}5}$, $S_{8\text{-}2}$,
364 etc; 65.2%; Table 4) the Asian private alleles detected were not present in the mother trees,

365 evidencing the relevance of pollination in introgression. Moreover, allele 176, present in
366 4.4% of saplings (Table 4), was absent in adults, indicating the presence of additional trees
367 of Asian origin in the forest.
368

369 Molecular markers could be applied to germplasm verification and commercial
370 seedling certification and to screen out plants of exotic origin. Several European countries
371 have expressed a need to certify the origin and genetic quality of plant material used in
372 forestry. Use of native species is encouraged, and methods for detecting introgression and
373 hybridization in selected material are needed. Before release onto the market, selected *C.*
374 *sativa* chestnuts from Bergondo must be registered in the Spanish 'National Catalogue of
375 Forest Base Material'. In a recent study, seedlings obtained from nuts collected from the A_{12}
376 *C. sativa* × *C. crenata* tree showed high *P. cinnamomi* resistance (coded Bergondo 9 in
377 Alcaide *et al.*, 2020). Moreover, seedlings from nuts collected from the A_{14} *C. sativa* tree
378 showed tolerance to both drought and *P. cinnamomi* (coded Bergondo 11 in Alcaide *et al.*,
379 2019, 2020). Selected seedlings from A_{12} and A_{14} trees, those assigned as pure *C. sativa* by
380 the 13 markers reported here, are being propagated and will soon be challenged for a second
381 time.
382

383 This work has improved understanding of introgression in chestnut populations. The
384 study shows the utility of SSRs to assess the extent of hybridization in chestnut forests, and
385 provides 13 markers and six new private alleles applicable to germplasm verification and
386 detection of chestnuts of exotic origin. To assess evolutionary and ecological processes in
387 chestnut in areas where ink-disease resistant clones have been planted, genetic monitoring
388 of adult and juvenile trees is recommended.
389

390 **Supplementary data**

391 Supplementary data are available at *Forestry* online.

392

393 **Table S1** Mother trees and saplings genotypes used in the study. The table confirms that the
394 progenies shared 50% of their alleles with their putative mother trees.

395

396 **Table S2** Genetic profile and taxa assignation of 34 adult trees (A) and their offspring (S)
397 sampled from a chestnut forest in northwest Spain. Estimated membership coefficients (Q)
398 of each individual to each group (K = 4) detected by STRUCTURE are shown. Values in
399 bold exceed 70% group membership.

400

401

402 **Data Availability Statement**

403 The data underlying this article will be shared on reasonable request to the corresponding

404 author.

405

406 **Acknowledgments**

407 We are very grateful to Álvaro Camisón, Mireya Salas, Rubén González and Simone
408 Castellana for technical assistance, Ramiro Martínez for information about the most frequent
409 chestnut clones planted in Galicia, and Jane McGrath for English editing of the manuscript.

410

411 **Conflict of interest statement**

412 None declared.

413

414 **Funding**

415 This work was supported by the Spanish Ministry of Economy and Competitiveness [grant
416 number AGL2014-53822-C2-1-R], the Regional Government of Extremadura [grant
417 number IB18091, VI R&D Regional Plan], and the European Union's European Regional
418 Development Fund (ERDF).

419

420 **References**

421 Alcaide, F., Solla, A., Mattioni, C., Castellana, S. and Martín, M.Á. 2019 Adaptive diversity and
422 drought tolerance in *Castanea sativa* assessed through EST-SSR genic markers. *Forestry* **92** (3),
423 287–296.
424 Alcaide, F., Solla, A., Cherubini, M., Mattioni, C., Cuenca, B., Camisón, Á., *et al* 2020 Adaptive
425 evolution of chestnut forests to the impact of ink disease in Spain. *J. Syst. Evol.* **58** (4), 504–516.
426 An, M., Deng, M., Zheng, S.S., Jiang, X.L. and Song, Y.G. 2017 Introgression threatens the genetic
427 diversity of *Quercus austrocochinchinensis* (Fagaceae), an endangered oak: a case inferred by
428 molecular markers. *Front. Plant Sci*. **8**, 229.
429 Bacilieri, R., Ducousso, A., Petit, R.J. and Kremer, A. 1996 Mating system and asymmetric
430 hybridization in a mixed stand of European oaks. *Evolution* **50** (2), 900–908.
431 Barreneche, T., Casasoli, M., Russel, K., Akkak, A., Meddour, H., Plomion, C., *et al* 2004
432 Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs).
433 *Theor. Appl. Genet*. **108** (3), 558–566.
434 Brunet, J., Zalapa, J.E., Pecori, F. and Santini, A. 2013 Hybridization and introgression between the
435 exotic Siberian elm, *Ulmus pumila*, and the native Field elm, *U. minor*, in Italy. *Biol. Invasions*
436 **15** (12), 2717–2730.
437 Buck, E.J, Hadonou, M., James, C.J., Blakesley, D. and Russell, K. 2003 Isolation and
438 characterization of polymorphic microsatellites in European chestnut (*Castanea sativa* Mill.).
439 *Mol. Ecol. Notes* **3** (2), 239–241.
440 Buck, R., Hyasat, S., Hossfeld, A. and Flores-Rentería, L. 2020 Patterns of hybridization and cryptic
441 introgression among one- and four-needled pinyon pines. *Ann. Bot.* **26** (3), 401–411.
442 Burgarella, C., Barnaud, A., Kane, N.A., Jankowski, F., Scarcelli, N., Billot, C., *et al* 2019 Adaptive
443 introgression: an untapped evolutionary mechanism for crop adaptation. *Front. Plant Sci.* **10**, 4.
444 Camisón, Á., Martín, M.Á., Oliva, J., Elfstrand, M. and Solla, A. 2019a Increased tolerance to
445 *Phytophthora cinnamomi* in offspring of ink-diseased chestnut (*Castanea sativa* Miller) trees.
446 *Ann. For. Sci*. **76** (4), 1–13.

447 Camisón, Á., Martín, M.Á., Sánchez-Bel, P., Flors, V., Alcaide, F., Morcuende, D., *et al* 2019b
448 Hormone and secondary metabolite profiling in chestnut during susceptible and resistant
449 interactions with *Phytophthora cinnamomi*. *J. Plant Physiol*. **241**, 153030.
450 Camisón, Á., Martín, M.Á., Dorado, F.J., Moreno, G. and Solla, A. 2020 Changes in carbohydrates
451 induced by drought and waterlogging in *Castanea sativa*. *Trees* **34,** 579–591.
452 Camisón, Á., Martín, M.Á., Flors, V., Sánchez-Bel, P., Pinto, G., Vivas, M., *et al* 2021 Exploring
453 the use of scions and rootstocks from xeric areas to improve drought tolerance in *Castanea sativa*
454 Mill. Env. Exp. Bot. **187**, 104467.
455 Castellana, S., Martín, M.Á., Solla, A., Alcaide, F., Villani, F., Cherubini, M., *et al* 2021 Signatures
456 of local adaptation to climate in natural populations of sweet chestnut (*Castanea sativa* Mill.)
457 from southern Europe. *Ann. For. Sci*. **78**, 27.
458 Crowl, A.A., Manos, P.S., McVay, J.D., Lemmon, A.R., Lemmon, E.M., *et al* 2020 Uncovering the
459 genomic signature of ancient introgression between white oak lineages (*Quercus*). *New Phytol.*
460 **226** (4), 1158–1170.
461 Earl, D.A. and vonHoldt, B.M. 2012 STRUCTURE HARVESTER: a website and program for
462 visualizing structure output and implementing the Evanno method. *Conserv. Genet. Resour.* **4** (2),
463 359–361.
464 Edmands, S. 2007 Between a rock and a hard place: evaluating the relative risks of inbreeding and
465 outbreeding for conservation and management. *Mol. Ecol*. **16** (3), 463–475.
466 Elorrieta y Artaza, J. 1949 El castaño en España. Instituto Forestal de Investigaciones y Experiencias
467 (IFIE), Madrid.
468 Evanno, G., Regnaut, S. and Goudet, J. 2005 Detecting the number of clusters of individuals using
469 the software STRUCTURE: a simulation study. *Mol. Ecol.* **14** (8), 2611–2620.
470 Excoffier, L., Laval, G. and Schneider, S. 2005 Arlequin (version 3.0): an integrated software
471 package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50.
472 Falush, D., Stephens, M. and Pritchard, J.K. 2003 Inference of population structure using multilocus
473 genotype data: linked loci and correlated allele frequencies. *Genetics* **164** (4), 1567–1587.

474 Fernandes, P., Tedesco, S., Vieira da Silva, I., Santos, C., Machado, H. and Lourenço Costa, R. 2020
475 A new clonal propagation protocol develops quality root systems in Chestnut. *Forests*, **11** (8),
476 826.
477 Fernández-Cruz, J. and Fernández-López, J. 2012 Morphological, molecular and statistical tools to
478 identify *Castanea* species and their hybrids. *Conserv. Genet.* **13** (6), 1589–1600.
479 Fernández-Cruz, J. and Fernández-López, J. 2016 Genetic structure of wild sweet chestnut (*Castanea*
480 *sativa* Mill.) populations in northwest of Spain and their differences with other European stands.
481 *Conserv. Genet*. **17** (4), 949–967.
482 Fernández-López, J. 2011 Identification of the genealogy of interspecific hybrids between *Castanea*
483 *sativa*, *Castanea crenata* and *Castanea mollissima*. *For. Syst.* **20** (1), 65–80.
484 Geraldes, A., Farzaneh, N., Grassa, C.J., McKown, A.D., Guy, R.D., Mansfield, S.D., *et al* 2014
485 Landscape genomics of *Populus trichocarpa*: the role of hybridization, limited gene flow, and
486 natural selection in shaping patterns of population structure. *Evolution* **68** (11), 3260–3280.
487 Gobbin, D., Hohl, L., Conza, L., Jermini, M., Gessler, C. and Conedera, M. 2007 Microsatellite-
488 based characterization of the *Castanea sativa* cultivar heritage of southern Switzerland. *Genome*
489 **50** (12), 1089–1103.
490 González, M.V., Cuenca, B., López, M., Prado, M.J. and Rey, M. 2011 Molecular characterization
491 of chestnut plants selected for putative resistance to *Phytophthora cinnamomi* using SSR markers.
492 *Sci. Hortic.* **130** (2), 459–467.
493 Hamilton, J.A., De la Torre, A.R. and Aitken, S.N. 2015 Fine-scale environmental variation
494 contributes to introgression in a three-species spruce hybrid complex. *Tree Genet. Genomes* **11**
495 (1), 817.
496 Hennion, B. 2010 Chestnut production in France: review, perspectives. *Acta Hortic*. **866**, 493–497.
497 Jakobsson, M. and Rosenberg, N.A. 2007 CLUMPP: a cluster matching and permutation program
498 for dealing with label switching and multimodality in analysis of population structure.
499 *Bioinformatics* **23** (14), 1801–1806.

500 Jarman, R., Mattioni, C., Russell, K., Chambers, F.M., Bartlett, D., Martin, M.A., *et al* 2019 DNA
501 analysis of *Castanea sativa* (sweet chestnut) in Britain and Ireland: Elucidating European origins
502 and genepool diversity. *PloS One* **14** (9), e0222936.
503 Jaynes, R.A. 1964 Interspecific crosses in the genus *Castanea. Silvae Genet*. **13** (5), 146–154.
504 Ji, F., Wei, W., Liu, Y., Wang, G., Zhang, Q., Xing, Y., *et al* 2018 Construction of a SNP-based
505 high-density genetic map using genotyping by sequencing (GBS) and QTL analysis of nut traits
506 in Chinese chestnut (*Castanea mollissima* Blume). *Front. Plant Sci*. **9**, 816.
507 Johnson, M.G., Lang, K., Manos, P., Golet, G.H. and Schierenbeck, K.A. 2016 Evidence for genetic
508 erosion of a California native tree, *Platanus racemosa*, via recent, ongoing introgressive
509 hybridization with an introduced ornamental species. *Conserv. Genet.* **17** (3), 593–602.
510 Kang, M.J., Shin, A.Y., Shin, Y., Lee, S.A., Lee, H.R., Kim, T.D., *et al* 2019 Identification of
511 transcriptome-wide, nut weight-associated SNPs in *Castanea crenata*. *Sci. Rep.* **9**, 13161.
512 Krebs, P., Conedera, M., Pradella, M., Torriani, D., Felber, M. and Tinner W. 2004 Quaternary
513 refugia of the sweet chestnut (*Castanea sativa* Mill.): an extended palynological approach. *Veg.*
514 *Hist. Archaebot.* **13** (3), 145–160.
515 Lafitte, G. 1946 Le châtaignier japonais en Pays Basque. *Mendionde* **69**.
516 Lepais, O., Petit, R.J., Guichoux, E., Lavabre, J.E., Alberto, F., Kremer, A., *et al* 2009 Species
517 relative abundance and direction of introgression in oaks. *Mol. Ecol.* **18** (10), 2228–2242.
518 Manino, A., Patetta, A., Marletto F. 1991. Investigations on chestnut pollination. *Acta Hortic*. **288**,
519 335–339.
520 Marinoni, D., Akkak, A., Bounous, G., Edwards, K.J. and Botta, R. 2003 Development and
521 characterization of microsatellite markers in *Castanea sativa* (Mill.). *Mol. Breed.* **11** (2), 127–
522 136.
523 Martín, J.A., Solla, A., Oszako, T. and Gil, L. 2021 Characterizing offspring of Dutch elm disease-
524 resistant trees (*Ulmus minor* Mill.). *Forestry* **94** (3), 374–385.
525 Martín, M.A., Mattioni, C., Molina, J.R., Alvarez, J.B., Cherubini, M., Herrera, M.A., *et al* 2012
526 Landscape genetic structure of chestnut (*Castanea sativa* Mill.) in Spain. *Tree Genet. Genomes* **8**
527 (1), 127–136.

528 Martín, M.A., Monedero, E. and Martín, L.M. 2017 Genetic monitoring of traditional chestnut
529 orchards reveals a complex genetic structure. *Ann. For. Sci.* **74** (1), 15.
530 Mattioni, C., Martín, M.A., Pollegioni, P., Cherubini, M. and Villani, F. 2013. Microsatellite markers
531 reveal a strong geographical structure in European populations of *Castanea sativa* (Fagaceae):
532 Evidence for multiple glacial refugia. *Am. J. Bot.* **100** (5), 951–961.
533 Mattioni, C., Martin, M.A., Chiocchini, F., Cherubini, M., Gaudet, M., Pollegioni, P., *et al* 2017
534 Landscape genetics structure of European sweet chestnut (*Castanea sativa* Mill): indications for
535 conservation priorities. *Tree Genet. Genomes* **13** (2), 39.
536 McKay, J.W. 1942 Self-sterility in the Chinese chestnut (*Castanea mollissima*). *P. Am. Soc. Hortic.*
537 *Sci.* **41**, 156–160.
538 Míguez-Soto, B., López-Villamor, A. and Fernández-López, J. 2016 Additive and non-additive
539 genetic parameters for multipurpose traits in a clonally replicated incomplete factorial test of
540 *Castanea* spp. *Tree Genet. Genomes* **12** (3), 47.
541 Míguez-Soto, B., Fernández-Cruz, J. and Fernández-López, J. 2019 Mediterranean and Northern
542 Iberian gene pools of wild *Castanea sativa* Mill. are two differentiated ecotypes originated under
543 natural divergent selection. *PloS One* **14** (2), e0211315.
544 Miranda-Fontaíña, M.E., Fernández-López, J., Vettraino, A.M. and Vannini, A. 2007 Resistance of
545 *Castanea* clones to *Phytophthora cinnamomi*: testing and genetic control. *Silvae Genet.* **56**, 11–
546 21.
547 Park, S.D.E. 2001 Trypanotolerance in West African cattle and the population genetic effects of
548 selection. PhD thesis, University of Dublin, Dublin, Ireland. Available from
549 http://animalgenomics.ucd.ie/sdepark/ms-toolkit/.
550 Peakall, R. and Smouse, P.E. 2012 GenAlEx 6.5: genetic analysis in Excel. Population genetic
551 software for teaching and research-an update. *Bioinformatics* **28** (19), 2537–2539.
552 Pereira-Lorenzo, S. and Fernández-López, J. 1997 Propagation of chestnut cultivars by grafting:
553 Methods, rootstocks and plant quality. *J. Hortic. Sci.* **72** (5), 731–739.

554 Pereira-Lorenzo, S., Costa, R.M.L., Ramos-Cabrer, A.M., Ribeiro, C.A.M., da Silva, M.F.S.,
555 Manzano, G., *et al* 2010 Variation in grafted European chestnut and hybrids by microsatellites
556 reveals two main origins in the Iberian Peninsula. *Tree Genet. Genomes* **6** (5), 701–715.
557 Pereira-Lorenzo, S., Ramos-Cabrer, A.M., Barreneche, T., Mattioni, C., Villani, F., Díaz-Hernández,
558 M.B., *et al* 2017 Database of European chestnut cultivars and definition of a core collection using
559 simple sequence repeats. *Tree Genet. Genomes* **13** (5), 1–6.
560 Pereira-Lorenzo, S., Ramos-Cabrer, A.M., Barreneche, T., Mattioni, C., Villani, F., Díaz-Hernández,
561 M.B., *et al* 2019 Instant domestication process on European chestnuts cultivars. *Ann. Appl. Biol.*
562 **174** (1), 74–85.
563 Pereira-Lorenzo, S., Bischofberger, Y., Conedera, M., Piattini, P., Crovadore, J., Chablais, R., *et al*
564 2020 Reservoir of the European chestnut diversity in Switzerland. *Biodivers. Consevr.* **29** (7),
565 2217–2234.
566 Perkins, M.T., Robinson, A.C., Cipollini, M.L. and Craddock, J.H. 2019 Identifying host resistance
567 to *Phytophthora cinnamomi* in hybrid progeny of *Castanea dentata* and *Castanea mollissima*.
568 *HortScience* **54** (2), 221–225.
569 Poljak, I., Idžojtić, M., Šatović, Z., Ježić, M., Ćurković-Perica, M., Simovski, B., *et al* 2017 Genetic
570 diversity of the sweet chestnut (*Castanea sativa* Mill.) in Central Europe and the western part of
571 the Balkan Peninsula and evidence of marron genotype introgression into wild populations. *Tree*
572 *Genet. Genomes* **13** (1), 18.
573 Rosenberg, N.A. 2004 DISTRUCT: a program for the graphical display of population structure. *Mol.*
574 *Ecol. Notes* **4** (1), 137–138.
575 Santos, C., Machado, H., Correia, I., Gomes, F., Gomes-Laranjo, J. and Costa, R. 2015 Phenotyping
576 *Castanea* hybrids for *Phytophthora cinnamomi* resistance. *Plant Pathol.* **64** (4), 901–910.
577 Solla, A., López-Almansa, J., Martín, J. and Gil, L. 2014 Genetic variation and heritability estimates
578 of *Ulmus minor* and *Ulmus pumila* hybrids for budburst, growth and tolerance to *Ophiostoma*
579 *novo-ulmi*. *IFOREST* **8** (4), 422–430.
580 Suarez-Gonzalez, A., Lexer, C. and Cronk, Q.C.B. 2018 Adaptive introgression: a plant perspective.
581 *Biol. Lett.* **14** (3), 20170688.

582 Schwartz, M.K., Luikart, G., Waples, R.S. 2007 Genetic monitoring as a promising tool for
583 conservation and management. *Trends Ecol. Evol.* **22** (1), 25–33.

584 **Table 1** Trees sampled in Bergondo chestnut forest (NW Spain), comprising 34 adult individuals and 42 open pollinated saplings, and *Castanea*
585 *mollissima* and *C. crenata* accessions used

Accession	Identification code
Bergondo chestnut population	
Adult trees (saplings from offspring)[1]	A_1, A_2, A_3, A_4, A_5, A_6 (S_{6-1}, S_{6-2}, S_{6-3}, S_{6-4}, S_{6-5}), A_7 (S_{7-1}, S_{7-2}, S_{7-3}, S_{7-4}, S_{7-5}), A_8 (S_{8-1}, S_{8-2}, S_{8-3}, S_{8-4}, S_{8-5}), A_9 (S_{9-1}, S_{9-2}, S_{9-3}, S_{9-4}), A_{10}, A_{11} (S_{11-1}, S_{11-2}, S_{11-3}, S_{11-4}, S_{11-5}), A_{12} (S_{12-1}, S_{12-2}, S_{12-3}, S_{12-4}, S_{12-5}), A_{13} (S_{13-1}, S_{13-2}, S_{13-3}, S_{13-4}, S_{13-5}), A_{14} (S_{14-1}, S_{14-2}, S_{14-3}, S_{14-4}, S_{14-5}), A_{15} (S_{15-1}, S_{15-2}, S_{15-3}), A_{16}, A_{17}, A_{18}, A_{19}, A_{20}, A_{21}, A_{22}, A_{23}, A_{24}, A_{25}, A_{26}, A_{27}, A_{28}, A_{29}, A_{30}, A_{31}, A_{32}, A_{33}, A_{34}
Asian germplasm collection	
Castanea crenata	*C. crenata* 1, *C. crenata* 2[2], *C. crenata* 3[2], *C. crenata* 4[2]
Castanea mollissima	*C. mollissima* 1, *C. mollissima* 2, *C. mollissima* 3

586 [1]F_1 saplings from 9 selected adult mother trees.
587 [2]Provided by TRAGSA chestnut breeding programme (González *et al*., 2011).
588

589 **Table 2** Characteristics of 13 SSR markers grouped in multiplexes A, B, C and D used to detect introgression of Asian germplasm into a natural
590 chestnut forest in northwest Spain

Locus	Motif	Primer sequence (5′–3′)	Size (bp)	L_G	Na	H_O	H_E	uH_E	PIC
Multiplex A									
CsCAT1	$(TG)_5TA(TG)_{24}$	F: NED-GAGAATGCCCACTTTTGCA R: GCTCCCTTATGGTCTCG	174-221	8	12	0.34	0.57	0.57	0.56
CsCAT2	$(AG)_{16}$	F: 6-FAM-GTAACTTGAAGCAGTGTGAAC R: CGCATCATAGTGAGTGACAG	187-245	10	12	0.70	0.82	0.83	0.80
CsCAT14	$(CA)_{22}$	F: 6-FAM-CGAGGTTGTTGTTCATCATTAC R: GATCTCAAGTCAAAAGGTGTC	130-164	2	12	0.74	0.80	0.81	0.78
EMCs25	$(GA)_{12}$	F: VIC-ATGGGAAAATGGGTAAAGCAGTAA R: AACCGGAGATAGGATTGAACAGAA	140-160	-	5	0.54	0.76	0.77	0.72
Multiplex B									
CsCAT3	$(AG)_{20}$	F: 6-FAM-CACTATTTTATCATGGACGG R: CGAATTGAGAGTTCATACTC	186-264	12	11	0.39	0.75	0.76	0.72
CsCAT6	$(AC)_{24}AT(AC)_4$	F: VIC-AGTGCTCGTGGTCAGTGAG R: CAACTCTGCATGATAAC	158-196	1	10	0.86	0.82	0.83	0.80
CsCAT16	$(TC)_{20}$	F: 6-FAM-CTCCTTGACTTTGAAGTTGC R: CTGATCGAGAGTAATAAAG	125-171	6	6	0.48	0.65	0.66	0.59
EMCs38	$(AG)_{31}$	F: NED-TTTCCCTATTTCTAGTTTGTGATG R: ATGGCGCTTTGGATGAAC	229-270	4	14	0.36	0.82	0.82	0.79
Multiplex C									
CsCAT41A	$(AG)_{20}$	F: VIC-AAGTCAGCCAACACCATATGC R: CCCACTGTTCATGAGTTTCT	186-258	8	12	0.84	0.80	0.80	0.77
EMCs2	$(CGG)_7$	F: 6-FAM-GCTGATATGGCAATGCTTTTCCTC R: GCCCTCCAGCCTCACCTTCATCAG	140-174	6	5	0.91	0.65	0.65	0.60
OCI	$(GT)_8$	F: NED-GGAATAAGTGGGGTGGGTTT R: GGGCCAAAGCATCACATTAC	145-170	-	7	0.92	0.78	0.78	0.75
Multiplex D									
EMCs14	$(GAG)_7$	F: VIC-GTGCTTCAGGGACCTTTCTTCTC R: GCCGCCGCCTCCTGCTGCTC	128-152	5	3	0.84	0.53	0.54	0.45
EMCs15	$(CAC)_9$	F: 6-FAM-CTCTTAGACTCCTTCGCCAATC R: CAGAATCAAAGAAGAGAAAGGTC	71-95	9	6	0.93	0.77	0.78	0.73

591 L_G, linkage group; Na, number of different alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; uH_E, unbiased expected heterozygosity; PIC, polymorphic
592 information content.

593 **Table 3** Privative alleles of Asian *Castanea* species obtained from SSR markers tested in chestnut trees from Bergondo forest (NW Spain) and in
594 *C. mollissima* and *C. crenata* accessions. Allele sizes are given in base pairs

	SSR markers							
	CsCAT1	*CsCAT14*	*CsCAT6*	*CsCAT16*	*EMCs38*	*CsCAT41B*	*OCI*	*EMCs14*
C. crenata	176, 186[1]	136[1]	138	130	239, 257	228, 232	151, 165	
C. mollissima	186[1]	136[1]				251, 254, 256		133

595 [1]Alleles present in both Asian species only, not in *C. sativa*.

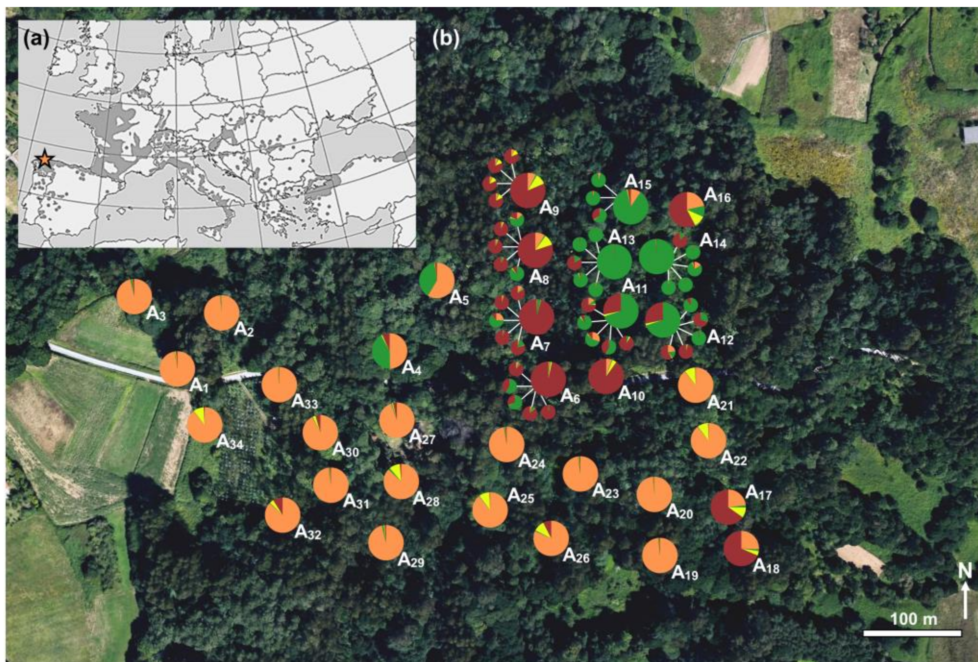
Table 4 Genetic profile and private alleles of Asian origin within nine mother trees (A_6-A_9, A_{11}-A_{15}) and their offspring (S) sampled from a chestnut forest in northwest Spain. Estimated membership coefficients (Q) of each individual to each group (K = 4) detected by STRUCTURE are shown. Values in bold exceed 70% group membership. Private allele sizes are given in base pairs

ID[1]	Q-value (K = 4)				SSR loci[2]					
	$Q1$	$Q2$	$Q3$	$Q4$	*CsCAT1*	*CsCAT14*	*CsCAT6*	*EMCs38*	*CsCAT41B*	*OCI*
A_6	0.011	0.009	0.013	**0.967**	186		138		232	165
S_{6-1}	0.064	0.038	0.008	**0.890**						151
S_{6-2}	0.008	0.573	0.004	0.415						165
S_{6-3}	0.014	0.658	0.010	0.318			138			
S_{6-4}	0.044	0.068	0.013	**0.875**	176					151
S_{6-5}	0.005	0.028	0.003	**0.964**				257		165
A_7	0.007	0.030	0.004	**0.959**			138	257		151, 165
S_{7-2}	0.025	0.005	0.014	**0.956**			138			151, 165
S_{7-4}	0.015	0.014	0.007	**0.964**			138			151, 165
A_8	0.089	0.010	0.100	**0.801**				239		
S_{8-2}	0.024	0.009	0.031	**0.936**					228	165
S_{8-4}	0.009	0.039	0.012	**0.940**			138	257		
S_{8-5}	0.012	**0.856**	0.024	0.108						165
A_9	0.076	0.011	0.099	**0.814**				257	228, 232	
S_{9-1}	0.070	0.006	0.110	**0.814**				257	228	
S_{9-2}	0.061	0.020	0.108	**0.811**				239, 257	228	
S_{9-3}	0.059	0.009	0.112	**0.820**				239	228	
S_{9-4}	0.057	0.015	0.108	**0.820**					228	
A_{11}	0.004	0.696	0.017	0.283		136				
S_{11-3}	0.298	0.570	0.020	0.112		136				
S_{11-4}	0.011	0.545	0.009	0.435		136				
S_{11-5}	0.004	0.051	0.034	**0.911**	186	136	138			151
A_{12}	0.004	0.696	0.017	0.283		136				
S_{12-1}	0.014	**0.876**	0.004	0.106			138			
S_{12-2}	0.013	0.281	0.006	0.700				257		
S_{12-4}	0.007	0.013	0.022	**0.958**	186		138			151
A_{13}	0.002	**0.992**	0.002	0.004						
S_{13-3}	0.005	0.136	0.011	**0.848**			138			151
A_{14}	0.002	**0.992**	0.002	0.004						
S_{14-1}	0.011	0.095	0.006	**0.888**			138			151
A_{15}	0.094	**0.870**	0.008	0.028						
S_{15-3}	0.021	0.616	0.013	0.350					232	

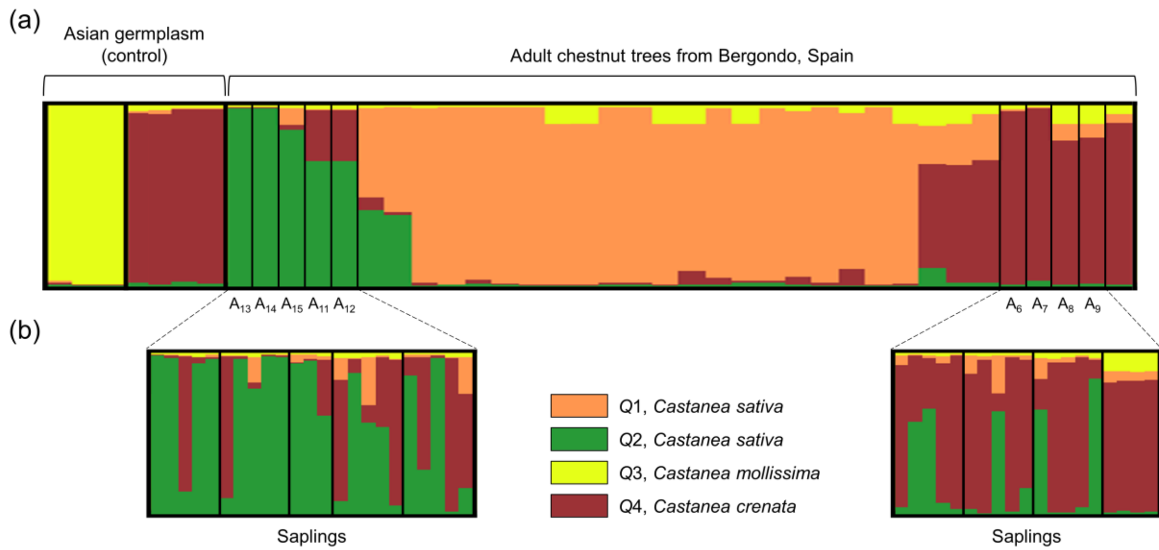
[1]Identification codes of adults (A) and saplings (S) from the chestnut forest.
[2]SSR loci showing private alleles found in the Asian *Castanea* species.

Figure legends

Figure 1 Distribution area of *Castanea sativa* according to EUFORGEN (a). The orange star indicates the location of the Bergondo chestnut forest studied, in northwest Spain. Location of the 34 adult trees and 42 saplings of chestnut sampled (b). The circular diagrams represent the genotypes, and colours indicate the proportion of estimated membership of trees detected by STRUCTURE analysis.



609 **Figure 2** Bayesian genotype assignment revealing introgression of Asian germplasm into a
610 chestnut forest in northwest Spain. The study comprised DNA from 3 *Castanea mollissima*,
611 4 *C. crenata* and 34 chestnut adult trees (a), and 42 chestnut saplings from nine mother trees
612 (b). STRUCTURE software based on 13 SSR data revealed $K = 4$. Each column represents
613 an individual tree, and colours in each column indicate the proportion of estimated
614 membership of trees to genetic clusters $Q1$ and $Q2$ (in orange and green, respectively,
615 assigned to *C. sativa*), $Q3$ (in yellow, assigned to *C. mollissima*) and $Q4$ (in brown, assigned
616 to *C. crenata*).



617