Domestication process on European chestnut cultivars

Santiago Pereira-Lorenzo[1]*, Ana M. Ramos-Cabrer[1], Teresa Barreneche[2], Claudia Mattioni[3], Fiorella Villani[3], María B. Díaz-Hernández[1], Luís M. Martín[4], Aurelio Robles-Loma[5], Yonatan Cáceres[6], Ángela Martín[4]

[1]Universidade de Santiago de Compostela. Departamento de Producción Vegetal y Proyectos de Ingeniería, Escola Politécnica Superior, Campus de Lugo, 27002 Lugo, Spain. e-mail adress Santiago Pereira Lorenzo: santiago.pereira_lorenzo@usc.es; Ana M. Ramos Cabrer: ana.ramos@usc.es; María B, Díaz Hernández: belen.diaz@usc.es

[2]INRA-Université de Bordeaux. UMR 1332 de Biologie du Fruit et Pathologie. F-33140 Villenave d'Ornon, France. e-mail adress: teresa.barreneche@inra.fr

[3]Istituto di Biologia Agroambientale e Forestale. Consiglio Nazionale delle Ricerche. Viale Guglielmo Marconi, 2, 05010 Porano TR, Italia. e-mail adress Claudia Mattioni: claudia.mattioni@ibaf.cnr.it; Fiorella Villani: fiorella.villani@ibaf.cnr.it

[4]Universidad de Córdoba. Escuela Técnica Superior de Ingenieros Agrónomos y de Montes. Edificio Gregor Mendel (C-5), Campus de Rabanales, 14071 Córdoba, Spain. e-mail adress: Luís M. Martín: lm.martin@uco.es; Ángela Martín: angela.martin@uco.es

[5]Rescue and conservation of fruit genetic resources. C/ Juan de Garay Nº 5. 01012 Vitoria-Gasteiz, Alava, Spain. E-mail adress: linoneka@yahoo.es

[6]Ingeniería Forestal y del Medio Natural, Universidad de Extremadura, Avenida Virgen del Puerto 2, 10600 Plasencia, Spain. ycaceres@unex.es

*corresponding author: santiago.pereira.lorenzo@usc.es Phone number +34982823128. Fax number +34982823001

Abstract

† Background and Aims This study presents the results of the first extensive sampling and systematic genetic analysis of ancient chestnut trees in Italy and the Iberian Peninsula in order to better understand the effect of grafting on the domestication process on chestnut from the earliest origins of clonally propagated European chestnut trees (*Castanea sativa* Mill.) and to investigate the impact of early selection and improvement process on the genetic diversity retained.

† Methods We have evaluated 105 giant trees from Italy, Spain and Portugal. They were compared with the EU database of chestnut cultivars (Pereira-Lorenzo et al., 2017) by using a set of 24 SSRs. We have measured the circumference (girth) at the diameter at breast height (dbh). Samples from both the canopy and the roots of each tree were analysed to distinguish

which trees were self-rooted and which were grafted. Diversity was compared using standard metrics and model-based approaches based on expected heterozygosity (He) at equilibrium.

† Key Results We could differentiate 91 new genotypes, from which only 9.6% matched known chestnut cultivars. We found the first evidences of cultivation, meaning grafting to produce "instant domestication" with unknown cultivars in Galicia and in the Douro Valley in trees of 14 m of perimeter (XV century) and in Basque Country (first report in that area) in a tree of 11.5 m of circumference (XVI century). In Italy, cultivar 'Marrone Fiorentino' was found in some giant trees between 8 and 9 m of perimeter (XVII-XVIII centuries) in Toscana and Umbria. Those findings matched with written references in Portugal XVI century in Portugal (Fernandes, 2012) and later on in the XVIII century in Spain (Sarmiento, 1986).

† Conclusions "Instant domestication" from wild could be dated after XV century and it was related to the wild populations existing in the same areas where cultivars are being propagated, without a different genetic structure for wild chestnut trees and with a high diversity maintained through the initiation of domestication.

Introduction

Large old trees are key structures that provide various ecological functions in many different environments, having major influence on hydrological regimes, nutrient cycles and distribution and abundance of populations of their own or other species (Lindenmayer et al. 2012). Moreover, the studies on centennial trees can be useful for interpreting the history of the species which is not only due to biological and environmental causes but also can be the result of anthropogenic drivers. In this context studying the genetic variability of ancient trees of domesticated species can give insights to reconstruct the domestication history of species. Domestication is a conscious or unconscious sampling and selection of the wild to selective propagate lineages with agronomical traits (Gross et al., 2014), which produces produces bottlenecks, with alleles fixed or lost at chance with the result in genetic changes that can be neutral or produce genetic and phenotypic differentiation of the crop from the wild relative, different to the one produced by directed selection.

European chestnut (*Castanea sativa* Mill.) is present in southern Europe as wild and cultivated. The first evidence of active chestnut cultivation is located in the eastern part of European range (Anatolian Peninsula, Northeastern Greece and Southeastern Bulgaria) in the third millennium before Christ (Conedera et al., 2004; Pereira-Lorenzo et al., 2012), being the Greeks first and the Romans later who diffused the chestnut to the west (Conedera et al., 2004). In the Middle Ages the cultivation of chestnut for timber and nut production was a common component of the traditional farming system in the areas where it is cultivated nowadays. Like other perennial crops, chestnut is clonally propagated, and many historical cultivars recorded in the literature as early as the XVI century in Portugal (Fernandes, 2012) and in the XVIII century in Spain (Sarmiento, 1986) still exist as orchard trees.

Giant trees can be a reservoir of genetic diversity as it was demonstrated in olive (Díez et al., 2011), being some of them unknown traditional cultivars that remained uncharacterized, and representing early stages in the cultivation processes, supporting the existence of several multilocal selection events in olive as it was demonstrating previously on chestnut (Pereira-Lorenzo et al., 2011).

This study presents the results of the first extensive sampling and systematic genetic analysis of ancient chestnut trees in Italy and the Iberian Peninsula. In a previous study, most of the cultivars from southern Europe (mostly Italy and Spain) were compared by using a common set of SRRs to report the first European database (Pereira-Lorenzo et al., 2017) with two different genetic backgrounds, one in Italy and a second one in the Iberian Peninsula. The main objective of this study was to evaluate chestnut giant trees from Italy and the Iberian Peninsula and to compare them with the genotypes of the European database in order to enhance our knowledge about the cultivation process and the impact of the giant trees on the genetic diversity in the European chestnut.

Material and Methods

Plant Material

We have collected a total of 102 giant trees; 84 from Spain, 16 from Italy and 2 from Portugal (Table S1, Figure S1). Only for 3 Spanish and 1 Italian trees the morphometric measures are notreported.

More than one sample was collected per tree, when it was possible, including at least one from shoots and from sprouts at the base of the trunk, and one over 1.50 m of the ground, to evaluate if they were grafted or not. For those giant trees with more than one trunk, each one was sampled independently. A total of 238 samples were collected from giant trees (Table S1).

We measured the circumference (girth) at breast height (dbh) of 1.30m. The putative age of the chestnut trees was estimated considering an annual slow ring growth of 0.4 cm (Fonti, 2002; Fonti and Sell, 2003). This putative range of growth is in validate considering the measures of the giant tree "Sietepernadas" which was planted in 1518 in Tenerife, Canary Islands(Pereira-Lorenzo et al., 2009).

Methods

Genetic analyses

A set of 24 SSR markers developed in *C. sativa* (Buck et al., 2003; Marinoni et al., 2003; Gobbin et al., 2007), or in *Quercus* (Steinkellner et al., 1997; Kampfer et al., 1998), were used for the analysis (Table S2). Methods are described by Pereira-Lorenzo et al. (2017).

We have evaluated separately the crown and the shoots from the base of the trunk when it was possible (Figure S2), finding two different situations: i) when they presented the same genotype, they were considered as wild,; ii) when they showed different genetic profiles, they were grafted and we differentiated the cultivar (crown) from the base (rootstock). Those genotypes were

compared to the genotypes of the EU database of chestnut cultivars (Pereira-Lorenzo et al., 2017).

A Bayesian analysis was performed with the STRUCTURE software (Pritchard et al., 2000) by using the admixture model with unlinked loci and correlated allele frequencies as defined in Pereira-Lorenzo et al. (2010) and Porras-Hurtado et al. (2013), recommending a minimum of 20 iterations to estimate the ancestry membership proportions of a population. We computed K = 1 to 15 unknown RPPs (reconstructed panmictic populations) of genotypes with the options use popinfo = 0, popflag = 0, which considers that the sampled genotypes were of unidentified origin, assigning them probabilistically to RPPs based on a qI (probability of membership) of 80%, while a lower probability meant an admixed genotype. The second order change of the likelihood function divided by the standard deviation of the likelihood (ΔK) was also estimated to find the best K value supported by the data (Evanno et al., 2005) by using Structure Harvester (Earl and Vonholdt, 2012).

Measures of expected (He) and observed heterozygosity (Ho) were estimated by using GenAlEx program (Peakall and Smouse, 2006) and inbreeding coefficient (Fis) (Weir and Cockerham, 1984) was calculated in the program GENODIVE (Meirmans and Van Tienderen, 2004). Measures of allelic richness (the mean number of distinct alleles per locus in a standardized sample size, Ar) were calculated using FSTAT (Goudet, 2001). Tests for recent bottlenecks, based on an excess of heterozygotes compared with the expected frequency under equilibrium H_e were conducted in the program BOTTLENECK (Cornuet and Luikart, 1996), simulated using the recommended two-phase model with 95% single-step mutations and 5% multiple-step mutations and a variance among multiple steps of 15 ((Piry et al., 1999).Sign test, standardized differences test and Wilcoxon sign rank test were performed under three different mutation models, infinite alleles model (IAM), stepwise mutation model (SMM), and two-phase model (TPM). IAM assumes that a mutation can involve any number of tandem repeats resulting in a new allele state in a population. SMM considers that, by insertion or deletion, a mutation can result in a change in one repeat unit, and implies that two alleles differing by one repeat are more closely related than alleles that differ by many repeats. Both models of mutation (SMM and IAM) cannot be appropriate for those loci. Therefore, there is and intermediate model between IAM and SMM (Di Rienzo et al., 1994) called the TPM. Furthermore, a qualitative test of mode shift was performed to evaluate the frequency distribution of alleles at different microsatellite loci using Bottleneck program (Piry et al., 1999; Ganapathi et al., 2012).

Diversity through time, taking into account different perimeters of the giant trees, was compared based on expected heterozygosity (He) at equilibrium as reported by Gross et al. (2014). Giant trees were classified in groups according to the perimeter of the trunk: i) 11-22 m (29 genotypes); ii) 8-10 m (39 genotypes); and iii) 4-7 m (30 genotypes). We have used the gene diversity He from wild trees (checked as wild comparing the crown with the base of the trunk) to compare with gene diversity He in RPPs and with different groups according to the perimeter of the trunks, by determining the diversity retained in % (Gross et al., 2014).

Results

Based on the results obtained with 24 SSRs, we differentiated 102 genotypes in this study, 98 of them corresponding to giant trees and 4 more without data of their perimeter (Tables 1 and S1). 11 genotypes (20 samples) from giant trees out of the 102 genotypes were coincident with 11 cultivars in the EU database. Therefore, we could differentiate 91 new genotypes in this study (175 samples, Table S1) when we compared them with those 123 genotypes reported in the EU database (without hybrids).

Diameters of the giant trees and grafting
Minimum perimeter recorded was 3.8 cm in the Basque Country (Spain) and the maximum were respectively in Sicily (Italy) with 22.0 m for the well-known 'Cento Cavalli' (the total collar circumference is 57.0 m) and close with 19.0 m in Ávila (Spain) (Tables 1 and S1). The giant tree 'Cento Cavalli' was a unique genotype on the crown (three samples from the upper part of the crown were tested and showed the same genotype though no shoots from the root could be checked), and they are part from an existing base of 57.0 m of perimeter. The widest giant tree in Spain is named 'El Abuelo' from El Tiemblo (Ávila), with 19.0 m of perimeter and classified as wild. Two trees with (13.9 and 14.4 m of perimeter), previously reported as 'Amarelante' and 'Campano' in the EU database of chestnut cultivars (Pereira-Lorenzo et al., 2017), were checked with the shoots from the base of the trunk and they showed the same genotype than in the crown.

We found 29 giant trees (72 samples) with a range of perimeter between 11 and 22 m whith an average of 13.8m, 39 between (98 samples) 8 and 10 m (62 samples) with an average of 8.8m, and 30 between 4 and 7 m with an average of 6.2m (Table 1). Among all these trees only nine genotypes presented circumferences over 15 m, three from Italy and six from Spain.

Grafting was firstly found in one giant tree in southern Galicia with 14.8 m of perimeter. In northern Portugal we found a grafted tree with a perimeter of 13.5 m (Table S1). Both of them were unreported or unknown cultivars. In the Basque Country we have found an unknown grafted cultivar in a giant tree of 11.5 m of perimeter, which crown samples had a different genotypes than the samples of the base of the trunk; this unknown cultivar was found in two more other trees from the Basque Country with 4.6 m of perimeter, both without shoots from the base of the trunk. Six more Spanish giant trees were grafted, which were unidentified cultivars; one from Castilla-León, and five from Galicia.

The oldest giant tree grafted with a known cultivar, the rare cv. 'Lebre' (Pereira-Lorenzo et al., 2006, 2011), was found in Galicia with 11.2 m of perimeter. The cv. 'Courela' was found in a tree with a 10.1 m of perimeter without shoots from the crown to check if it was grafted. 'Rapada' was also another reported cv. found in a grafted tree of 10 m of perimeter in Galicia, with another unknown graft over the same rootstock. 'Negral', a relevant cultivar in Galicia, was found in a grafted giant tree with 9.2 m of circumference. In Extremadura (central Iberian Peninsula), we found cv. 'Verata' grafted in a tree with 8.2 m of perimeter. Some other reported cultivars found were 'Luguesa' (7.0 m of perimeter), and 'Parede' (5.7 m of perimeter) in northwestern Spain.

In Italy, the first cultivar found was 'Marrone Fiorentino' in a tree with 8.5 m of perimeter without shoots from the base of the trunk to check if it was grafted and in three more grafted trees with 8 m of perimeter, one out of them in Umbria with a base of 16.7 m.

A total of 55 giant trees were checked to be wild (same genotype from the base of the trunk and the upper part). Wild giant trees have a perimeter varying from 3.8 to 19.0 m (9.3 m on average). Cultivated giant trees (135) varied from 5.7 to 14.8 m (10.6 m on average). Finally, 27 giant trees could not be checked with the sample from the base of the trunk (not sprouts).

Putative age of the giant trees

If we consider 0.40 cm of average growth for every tree, we have that the biggest chestnut tree reported for chestnut ('Cento Cavalli') with 22.0 m of perimeter (same genotype than two more other trunks from a crown of 57.0 m) was from 1141 (Tables 1 and S1). But if we reduced the annual growth of the rings to a minimum of 0.15 cm, for example, the tree came from 252 BC. Moreover, the crown is 57 m perimeter, which means that this Sicilian chestnut tree could have more than 2000 years old for 0.4 cm, 6000 for 0.15 cm and over 9000 for 0.10 cm of the growth of the annual rings, respectively.

First grafted giant trees with unknown cultivars had 14 m of circumference (one tree in Galicia and a second one in the Douro Valley), which were from 1479 for an average of 0.40 cm of the growth of the annual rings, and 584 from an average of 0.15 cm.

The first reported cultivar 'Lebre' was found in a grafted giant tree of 11.2 cm of perimeter and it could be from 1570 for 0.40 cm as an average of the annual growth of the ring, and 828 for 0.15 cm. Moreover, 'Rapada' and 'Courelá', grafted on trees of 10 and 10.1 m of perimeter, could be originated around 1600 considering 0.40 cm or in 944-955 for 0.15 cm. Consequently, 'Negral' could be dated on 1652 or 1945, respectively.

In Italy, 'Marrone Fiorentino' was found in a tree with a circumference of 8.5 cm, which corresponded to 1662 for 0.4 cm and 1061.

Structure and geographical distribution of giant trees in southern Europe

For the 98 genotypes differentiated among the giant trees, the most likely number of clusters considering those 18 SSRs without null alleles (Figure S3, Table S1), according to the ΔK criterion (Figure S3), gave the highest value for K = 2 and 3 and, in a second order, K = 5. For K =2, 52 genotypes only from the Iberian Peninsula were in RPP1, and 28 were in RPP2, including 12 from Italy and 16 from Spain (Galicia, Castilla-Leon, Vasque Country and Andalusia). Finally, 18 more genotypes were admixed. When K=3, two groups were distinguished in RPP1, one including cvs. 'Parede', 'Lebre' and 'Rapada', all of them from Galicia; and a second one grouping cvs. 'Amarelante' and 'Verata from Galicia and Extremadura, respectively.

When we analysed the giant trees with the EU chestnut database (Pereira-Lorenzo et al., 2017), with a total of 214 genotypes, the most likely number of clusters gave the highest value for K = 2 (Figure S3). Considering structure for K = 2, 144 genotypes were grouped with a $ql > 80\%$ (67.3%), two RPPs were thus identified, RRP1 grouping cultivars from the Iberian Peninsula and

Canary Islands (90), and RPP2 those cultivars (54) from Italy (28), Spain (22), France (4). There were 70 admixed genotypes (Figure 1, Tables 1 and S1).

We also tested K = 4, in which Spanish cultivars were separated in three different groups (RPP1a, b and c). This second level of sub-structure classified 106 genotypes (227 accessions) with a $ql>$ 80% (50.0%) (Figure 1, Tables 1 and S1). RPP1a included mainly genotypes (24) from Central Spain related to the most important cultivar in the Iberian Peninsula 'Longal' and the main cultivar 'Verata' found in a giant tree of 8.2 m of perimeter. RPP1b related 22 genotypes from northern Spain, mainly Asturias but also from Cantabria, Basque Country, Castilla-León and Galicia, with the main cultivar 'Parede' found in a giant tree of 5.7 m of circumference. RPP1c included genotypes (20) from Galicia (northwestern Spain), with the cultivars 'Lebre' and 'Rapada' found in two giant trees with 11 and 10 m of perimeter, respectively. RRP2 separated Italian cultivars (25), 13 Spanish cultivars (one from the North, 5 from Extremadura and seven from the South), and two French cultivars ('Marron Comballe' and 'Camberoune') with two relevant cultivars found in giant trees, one 'Marrone Fiorentino' Umbria (Italy) in a tree of 8.8 m of perimeter and a second one 'Luguesa' from Galicia (Spain) in a tree of 7 m of perimeter. Finally, a total of 70 genotypes were admixed when K = 2 and 108 when K = 4. Among the admixed genotypes were found two cultivars in giant trees from Galicia (northwestern Spain), 'Courelá' in a giant tree of 10.1 m of perimeter, and a second one 'Negral' in a giant tree of 9.2 m of circumference. Most of the admixed trees were found in Spain (66 when K = 2 and 98 when K = 4). Wild trees (confirmed the same genotype from the crown and the base of the trunk) were distributed in every RPP and admixed (they were only from Spain for K = 2 and K = 4) (Figure 1, Tables 1 and S1).

Giant trees for the two main RPPs when K = 2 varied on perimeter from 4.0 to 15.9 m (average 9.3 m) for RPP1 (Iberian Peninsula), from 6.3 to 22 m (average 9.9 m) for RPP2 (mainly Italian group) and 3.8 to 19.0 m (average 9.4 m) for admixed; with a minimum age for an annual growth of the ring of 0.4 cm of the RPP1 in 1384, RPP2 in 1141 and 1260 for the admixed.

Giant trees for the four main RPPs when K = 4 varied on perimeter from 8.2 to 13.5 m (average 10.3 m) for RPP1a, from 4.6 to 12.5 m (average 8.2 m) for RPP1b, from 5.7 to 14.8 m (average 9.8 m) for RPP1c, from 6.3 to 22.0 m (average 9.3 m) for RPP2, and 3.8 to 20.0 m (average 9.6 m) for admixed; with a minimum age for an annual growth of the ring of 0.4 cm of the RPP1a in 1479, RPP1b in 1517, RPP1c in 1427, RPP2 in 1141 and 1220 for the admixed.

The results obtained with STRUCTURE for K = 2 were confirmed by the representation of PCoA analysis and the Neighbor-joining (NJ) tree based on a genetic distance matrix using DARwin 6.0.010 software, similarly as it happened in a previous study of European chestnuts (Pereira-Lorenzo et al., 2017) (Figure 2). Moreover, in this study it is noticed the fixation of genotypes by grafting or "instant domestication" defined by Harris et al. (2002) in both RPPs from wild trees, some of them of the same ages as the cultivars cultivated nowadays, such as "Lebre" with a circumference of 11.2 m in RPP1 and 'Marrone Fiorentino' with 8.5 cm in RPP2.

Biggest giant grafted trees (unknown) with 14.8 m of circumference was one admixed from Galicia, Spain, and one from the Douro Valley (Portugal) in RPP1 with 13.5 m. With 15 m and over we found seven giant trees, one in RPP1 (Cantabria, Spain), four in RPP2 (three from Italy in Sicily and Umbria and one from Andalucía, Spain) and two admixed (Castilla-León, Spain).

Most of the admixed trees, including giant trees were in between RPP1 and RPP2 (Figure 2 AFC), indicating they could be originated from both groups, considering the old introgressions of Italian genotypes in the Iberian Peninsula (22 Spanish genotypes were grouped in RPP2). In this study, oldest giant trees admixed were found in Castilla-León (15.4 and 19.0 m of perimeter).

The pairwise FST values among STRUCTURE clusters was 0.068 (P < 0.001) among the Iberian Peninsula cultivar cluster (RPP1) and the Italian cultivar cluster (RPP2), 0.019 (P < 0.001) between admixed cultivars and RPP1 and 0.027 (P < 0.001) between admixed cultivars and RPP2 (Table S3). At the substructure level, a maximum FST value of 0.153 (P < 0.001), was obtained between RPP1c (north-western Spain, Galicia) and central-southern Spain (RPP1a); and lower between northern Spain (RPP1b) and central-southern Spain (RPP1c) with 0.112 (P < 0.001); and close to 0.103 (P < 0.001) between northern Spain (RPP1b) and Italy-France (RPP2). All RPPs showed values with admixed of 0.05 and lower. The pairwise FST value between wild and cultivated was of 0.008 (P < 0.001) (Table S4).

Chestnut diversity and domestication through time

We have used the gene diversity He = 0.685 from wild trees (checked as wild comparing the crown with the base of the trunk, including rootstocks) to compare with gene diversity and, we found out that RPP1 retained 94.5% of the diversity and RPP2 99.6% for K =2, with a 100.9% in the admixed (Table 2). When K = 4, a minimum of the 84.4% of the diversity retained was found in RPP1b ('Parede' group), followed for the 100% in RPP1a ('Verata' group), 101.7% in RPP1c ('Lebre' group), 102.6% for RPP2 ('Marrone Fiorentino' group), and with 97.8% in the admixed.

For ranges of perimeters, the biggest giant trees (11-22 m) retained 98.5% of the diversity, giant trees between 8-10 m the 100.7%, the giant trees between 4-7 m the 97.1%, meanwhile those chestnut trees without any information about the perimeter retained the 98.4%. Finally, cultivated trees (we excluded those that could not be checked with the base of the trunk) retained the 98.5% of the diversity respecting the wild chestnut trees.

Allelic richness showed a slightly different pattern, depending on the groups evaluated, with a slight increase on the cultivated group respecting to the wild chestnut group, very stable average number in the different groups of chestnut trees classified by the perimeter of the trunk. Average was also stable for RPPs when K = 2 and 4 for 24 SSRs (all loci included) and lower in RPP1 and higher in RPP2 for 18 SSRs (those without deviations due to null alleles).

In order to test the mutation drift equilibrium in RPPs for K = 2 and K =4, giant trees classified for perimeters, and cultivated and wild chestnut trees using multilocus genotypic data, three mutation models of microsatellite evolution were assumed, viz., IAM, TPM, and SMM (Table S5). The results of sign test showed, in general, a higher number of loci with He excess only under IAM, and the contrary for TPM and SMM. Thus, sign test revealed significant heterozygosity excess (P<0.05) indicating the deviation of chestnut groups from mutation drift equilibrium under IAM of microsatellite evolution. Standardized differences tests were performed by comparing the calculated T2 statistics (the differences between observed and expected gene diversities divided by standard deviation of the corresponding distributions of gene diversities), which were, in general, significantly positive (P<0.01) under IAM and significantly negative in chestnut giant trees classified by the perimeter and cultivated versus wild chestnut trees for TPM and SMM. Finally, the one-tail Wilcoxon sign rank test for gene diversity excess revealed

significant deviation (P<0.01) of the different chestnut groups from mutation drift equilibrium under IAM and never under TPM and SMM.

Furthermore, tests for mode shift in frequency distribution of different alleles, a qualitative method for detection of genetic bottleneck, were performed for different RPPs, chestnut groups classified by perimeters and cultivated vs. wild trees (Figure S4). In non bottlenecked populations, a large proportion of rare alleles are expected, meanwhile in bottleneck events rare alleles to become less abundant in the population than alleles with intermediate frequencies. In such cases, the plotting of proportion of different alleles against allele frequency classes will cause mode shift from the normal L-shaped distribution. In the present study, the normal L-shaped distribution was present in RPPs for K =2 and K = 4, in groups of giant trees classified by perimeters (meaning different times) and cultivated vs. wild trees, suggesting not bottleneck events.

Discussion

When cultivation begun in chestnut culture

In this study, we could relate wild chestnut trees and cultivars in the same areas where they were originated, and differently to olive (Díez et al., 2011), in which the origin of olive cultivars is still uncertain. 135 ancient olive trees classified according to their trunk diameter between 0.6 m (1.9 m of perimeter) to a maximum of 2.72 m (8.5 m of perimeter) revealed 10 known cultivars (9.6%), meanwhile in chestnut giant trees from 3.8 m to 22 m of perimeter we identified 11 cultivars (11.2%). Some unknown cultivars were found in olive giant trees as in our chestnut study. The percentage of grafted trees was higher among the olive trees with the largest trunk diameter, which suggests that this technique was used more frequently in the past to convert pre-existing wild olives into cultivated ones. On chestnut, grafting is compulsory for propagating cultivars due to the difficulty of self-rooting (Pereira-Lorenzo et al., 2012) and the biggest grafted tree had 14.8 m of perimeter (year 1427), but the percentage of grafted trees varied from 38.5% between 8-10m up to 50-52% for both 4-7 m and 11-22 m of perimeter, respectively. Even though grafting, described as 'instant domestication', was developed 3800 years ago (Harris et al., 2002), in our study it seems that it occurred in the XV century in the Iberian Peninsula and in the XVII century in Italy, which it does not match with the spread of the cultivation from the East to the West, which can be due to the lost of many giant trees due to pest and diseases, which firstly affect in the Eastern Europe (Pereira-Lorenzo et al., 2012). In other perennial crop tree as apple, the first cultivars were cited in 1200s (Gross et al., 2014), but apple tree lifespan is not long enough to certainly know if cultivars cited and that time correspond with genotype conserved in germplasm banks. In olive, the existence of several multilocal selection events was proposed based on molecular studies (Claros et al., 2000; Besnard et al., 2001), supported by the huge diversity of different clonally propagated cultivars and the presence of the wild olive the ancestor of the cultivated olive. On chestnut, locally selection of cultivars was found in the Iberian Peninsula identifying the origin of the main cultivars (Pereira-Lorenzo et al., 2010, 2011).

The oldest chestnut tree in the world is the "Cento Cavalli" with a main trunk of 22 m of perimeter over a crown of 57 m. Apart from this old tree, both main areas for chestnut culture,

Italy and the Iberian Peninsula, presented quite similar number of old trees less than 20 m of circumference.

First evidences of grafting "instant domestication" (Zohary and Spiegel Roy 1975; Harris et al., 2002) were found in Galicia and in the Douro Valley with two unknown cultivars in trees of 14 m of perimeter (1479) and in Basque Country, with an unknown cultivar grafted in a tree of 11.5 m of circumference (1558).

Reported cultivars in the EU chestnut database by Pereira-Lorenzo et al. (2017) were identified in three giant trees from Galicia with with respectively 11, 10 and 9 m of perimeter. In Italy, first reported cultivar 'Marrone Fiorentino' was found in a tree with 9 m.

Estimations of the ages for the giant trees considering 0.4 cm for annual growth of the rings (Fonti, 2002; Fonti and Sell, 2003) fit with the well-known giant tree in Tenerife, Canary Islands, supposedly planted at the beginning of the XVth century (Pereira-Lorenzo et al., 2009). In 1518 'Adelantado', Alonso Fernández de Lugo wrote about the plantation of 3000 chestnut trees in La Orotava. Considering the 0.4 cm, the oldest giant tree 'Cento Cavalli' in Sicily could be dated in XII century for the main trunk (22 m of perimeter). Following, grafting of unknown, or unreported, cultivars were then dated on XV th century in Galicia and Douro Valley, and the first reported grafted trees in Galician from XVI century onwards in Galicia and XVII century in Umbria. Cultivated chestnut was reported firstly in Greece, introduced in Sicily to be spread later in Italy, and progressively to the west (Conedera et al., 2004; Pereira-Lorenzo et al., 2012); however, the oldest grafted trees survived in southwestern Europe, which cannot be explained by the dynamics of the modern pests and diseases affecting chestnut groves in southern Europe. Ink disease (*Phytophthora* spp.) was firstly reported in Portugal in 1838 (Vannini and Vettraino, 2001). Blight (*Cryphonectria parasítica*) was firstly observed in Europe in Genoa, ltaly in 1938, and it spread quickly through ltaly and other European countries (Robin and Heiniger, 2001), and less so in southem UK, the Netherlands, central and southern Spain, and the Canary lslands. In 2002, gall wasp *Dryocosmus kuriphilus* was reported for the first time in Europe in northwest ltaly (Pereira-Lorenzo et al., 2012).

In Portugal we found the oldest written reference of chestnut cultivar in 1531-1532 (Fernandes, 2012) naming cv. 'Longal' and 'Reborda'. 'Longal' is the main cultivar in the Iberian Peninsula (Pereira-Lorenzo et al., 2011) spread as a clone from southern Galicia and northern Portugal to Extremadura in central Spain, and further in Castilla-León (Ávila) surrounding the Sierra de Gredos.

Grafted giant tree with cultivar 'Rapada' from Galicia (10 m of perimeter and dated in 1618) was firstly cited in between 1746-1775 in Galicia (Sarmiento, 1986), with some other cultivars such as 'Parede'.

In Canary Islands, the presence of the first chestnut tree was in 1518 (Pereira-Lorenzo et al., 2009) but not grafted. In Basque Country, cultivars were described by Elorrieta (1949) and in this work we found an unknown in a giant tree of 11.5 m of perimeter (1558).

In Italy, cultivar 'Marrone Fiorentino' was found in some giant trees between 8 and 9 m of perimeter (between 1650 and 1700) in Toscana and Umbria. The use of 'Marrone' referred to

the best quality of chestnut cultivars was explained by Breviglieri (1955) in the "Squeda Castagnografica" as a word evolved from the latin "maro", introduced in France in XVI century.

Structure and geographical distribution of giant trees

The biggest old European chestnut was found out in Sicily as expected. In the Iberian Peninsula the second biggest trunk was found in Central Spain (19 m of circumference). However, evidences of the oldest grafted trees were found in the Iberian Peninsula against previous references indicating that cultivation was introduced from Greece into Sicily and later on in the rest of Italy (Conedera et al., 2004; Pereira-Lorenzo et al., 2012).

Both main groups of cultivars, one located in the Iberian Peninsula and a second one in Italy were originated from different wild populations since we found some wild trees grouped in each RPP, indicating a different genetic background from which each RPP evolved. The Italian group (RPP2) presented some genotypes in the Iberian Peninsula, even some wild giant trees up to 9 m, indicating an earlier introduction from Italy in Spain, mainly in Andalucía and Extremadura, but also in Castilla-León and Galicia, with the cultivar 'Luguesa' belonged to this RPP2 and found in a giant tree of 7 m of perimeter.

More difficult is to explain admixed trees, because some of them up to 20 m of perimeter (when K = 4) could hardly be originated by the hybridisation of both main RPPs. However, some of those admixed genotypes (when K = 2 and K = 4) were not clustered in a different RPP, indicating a different origin or, more likely, an independent population, from which some cultivars were also selected in Spain later on than in RPP1 and earlier than in RPP2. On the other hand, many ancient chestnut trees could be already dead because of the many threatens of this species (Pereira-Lorenzo et al., 2012). Genetic differentiation between all RPPs and admixed were low, which could support the early hybridisation between the Italian RPP with the Iberian group.

In a similar study on olive giant trees, genetic structure for K = 2 differentiated wild trees from cultivated (Díez et al. 2011). In our results wild could not be separated from cultivated, which could indicate a more recent grafting "instant domestication" on chestnut from wild, even though evidences found on chestnut giant trees of cultivars grafted were older than in olive according to the size of the trees, 8.5 m of perimeter on olive versus 14.8 m on chestnut. The very low Fst value between wild and cultivated chestnut trees also confirmed the genetic closeness.

Chestnut diversity and effect of the domestication through time

Here, we discuss domestication and improvement bottlenecks in cultivated chestnut and the comparison of improvement bottlenecks across annual and perennial fruit crops, considering the oldest wild population detected by SSRs in this study to be compared with those cultivated cultivars in southern Europe. *He* is an useful measure because it incorporates the number and frequency of the alleles within a population (Gross et al., 2014); therefore, it reflects how likely it is for an allele to be lost from a population) and because it is a standard measure that can be compared across studies, mostly in perennial crops that are non-model systems, with less population level data.

Even though the higher *He* values can be due to high mutation rates at SSR loci that allow crops to regain diversity through the generation of novel or homoplasious alleles, the more severe values in annuals than in perennials, suggests that the trend is robust regardless of the measurement (Gross et al., 2014). Simulations of *He* based on allelic frequencies can also be used to detect recent genetic bottlenecks, indicated by an excess of heterozygotes compared with the simulated equilibrium expectations, as a result of a sudden decrease in allelic diversity that has not yet stabilized in the genotypes of individuals in the population (Cornuet and Luikart, 1996 ; Piry et al., 1999).

As it happened in other perennial crop trees such as apple, *Malus* $\times$ *domestica*, treated as a species group, retained 95.1% of the diversity present in *M. sieversii* (wild ancestor) based on the 9-SSR data set, and 96.7% based on the 19-SSR data set, as calculated from *H* e (Gross et al., 2014), similar to the 98.5 % of the diversity present in the cultivated chestnut group vs. the wild trees.

Both 18-SSR and the 24-SSR data set showed a significant excess of heterozygotes compared with the equilibrium expectation for the IAM mutation model of microsatellite evolution, but not for TPM, and SMM. Only when we evaluated giant trees classified by the perimeters of the trees, the three mutation models were significant. For any given data set, IAM predicts lower equilibrium gene diversity than TPM and SMM, and hence, it is more likely to indicate significant heterozygosity excess. Furthermore, the test for mode shift in frequency distribution of different alleles was performed in the different chestnut groups, a qualitative method for detection of genetic bottleneck, showing a the normal L-shaped distribution in all groups, even for those by the perimeter of the trunk. Therefore, if there was a genetic bottleneck in the past it was mild as happened on apple tree crop (Gross et al., 2014), explained by the high diversity maintained through the initiation of domestication, providing a broad genetic base for the development of new cultivars, leading to a mild improvement bottleneck.

Conclusions

We found the first evidences of cultivation in 20 (20.4%) giant trees (11 known and 9 unknown cultivars), meaning grafting to produce "instant domestication" in Galicia with unreported cultivars in a tree of 15 m of perimeter (1427), in the Douro Valley on a tree of 14 m of perimeter (1479) and in the Basque Country, with an unknown cultivar grafted in a tree of 11.5 m of circumference (1558). In Italy, cultivar 'Marrone Fiorentino' was found in some giant trees between 8 and 9 m of perimeter (between 1650 and 1700) in Toscana and Umbria. The reason why oldest evidences were found in western Europe instead in Italy could be related with the dynamics of gall wasp and blight (but not the ink disease) from the East to the West, which could produce the lost of many giant trees in those areas. Moreover, the existence of ancient trees in the Iberian Peninsula grouped in the Italian group gives evidence of an early introduction of Italian chestnuts in the Iberian Peninsula, which hybridised producing the admixed group.

"Instant domestication" (grafted giant trees) from wild (non-grafted giant trees) was a relatively recent event in the same areas where cultivars are being propagated, without a different genetic structure for wild chestnut trees since we found grafted and non-grafted giant trees in the same RPP. The high diversity maintained through the initiation of cultivation provided a broad genetic

base for the development of new cultivars, and that diversity was maintained by grafting in a high number of cultivars, which retained most of the diversity from wild. Promoting the use of that clonal diversity, so as the diversity conserved in the giant trees, must be a priority in order to maintain genetic variability of the orchards and, in general, of chestnut in Europe.

References

Besnard, G., Baradat, P., Berville, A. (2001) Genetic relationships in the olive (Olea europaea L.) reflect multilocal selection of cultivars. Theoretical and Applied Genetics 102: 251–258.

Breviglieri, N. (1951) Research on the flower and fruit biology in *Castanea sativa* and *Castanea crenata* in Vallombrosa territory. Publication n.1, Centro di Studio sul Castagno (C.N.R.) La Ricerca Scientifica (Suppl.) 21: 15-49

Buck, E.J., Hadonou, M., James, C.J., Blakesley, D., Russell, K. (2003) Isolation and characterization of polymorphic microsatellites in European chestnut (*Castanea sativa* Mill.) Mol Ecol Notes 3: 239-241. doi: 10.1046/j.1471-8286.2003.00410.x

Claros, M.G., Crespillo, R., Aguilar, M.L., Canovas, F.M. (2000) DNA fingerprinting and classification of geographically related genotypes of olive-tree (*Olea europaea* L.). Euphytica 116: 131–142.

Conedera, M., Krebs, P., Tinner, W., Pradella, M. & Torriani, D. (2004). The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. Veg Hist Archaeobot 13, 161–179.

Cornuet, J.M.,Luikart, G. (1996) Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks from Allele Frequency Data. Genetics, 144, 2001-2014.

Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M., Friemer, N.B. (1994) Mutational processes of simple sequence repeat loci in human populations. Proceedings of the National Academy of Sciences of the United States of America 91:3166–3170

Díez, C., Trujillo, I., Barrio, E., Belaj, A., Barranco, D., Rallo, L. (2011) Centennial olive trees as a reservoir of genetic diversity. Annals of Botany 108: 797–807, 2011. doi:10.1093/aob/mcr194

Earl, D.A., Vonholdt, B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4: 359-361

Elorrieta, J. (1949) El castaño en España. MAPA, Madrid, 303 pp.

Evanno, G., Regnaut, S., Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molec Ecol 14: 2611–2620. doi:10.1007/s10722-006-0020-6

Fernandes, R. (2012) Descrição do terreno ao redor de Lamego duas léguas, 1531-1532; ed., estudo introdutório e apêndice documental Amândio Morais Barros. - 1ª ed. - Casal de Cambra: Caleidoscópio, 2012. - 126 p. : il. ; 24 cm. - ISBN 978-989-658-188-6.

Fonti, P. (2002) Investigations into ring shake of chestnut. A dissertation (Diss. ETH No. 14732) submitted to the Swiss Federal Institute of Technology Zürich for the degree of Doctor of Natural Sciences, 103 pp.

Fonti, P., Sell, J. (2003) Radial split resistance of chestnut earlywood and its relation to ring width. Wood and Fiber Science, 35(2): 201–208.

Ganapathi, P., Rajendran, R., Kathiravan, P. (2012) Detection of occurrence of a recent genetic bottleneck event in Indian hill cattle breed Bargur using microsatellite markers. Tropical Animal Health and Production, 44:2007-2013. DOI 10.1007/s11250-012-0171-8

Gobbin, D., Hohl, L., Conza, L., Jermini, M., Gessler, C., Conedera, M. (2007) Microsatellite-based characterization of the *Castanea sativa* cultivar heritage of southern Switzerland. Genome 50: 1089-1103

Goudet, J. (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet 1995

Gross, B.L., Henk, A.D., Richards, C.M., Fazio, G., Volk, G.M. (2014). Genetic diversity *in Malus × domestica* (Rosaceae) through time in response to domestication. American Journal of Botany. 101 (10): 1770‐1779.

Harris, S.A., Robinson, J.P., Juniper, B.E. (2002) Genetic clues to the origin of the apple. Trends in Genetics 18: 426–430.

Kampfer, S., Lexer, C., Glossl, J., Steinkellner, H. (1998) Characterization of (GA)n microsatellite loci from *Quercus robur*. Hereditas 129: 183-186. doi: 10.1111/j.1601-5223.1998.00183.x

Lindenmayer, D.B., Gibbons, P., Bourke, M., Burgman, M., Dickman, C.R., Ferrier, S., Fitzsimons, J., Freudenberger, D., Garnett, S.T., Groves, C., Hobbs, R.J., Kingsford, R.T., Krebs, C., Legge, S., Lowe, A.J., McLean, R., Montambault, J., Possingham, H., Radford, J., Robinson, D., Smallbone, L., Thomas, D., Varcoe, T., Vardon, M., Wardle, G., Woinarski, J., Zerger, A. (2012). Improving biodiversity monitoring. Austral Ecology 37:285-294. http://dx.doi.org/10.1111/j.1442-9993.2011.02314.x

Marinoni, D., Akkak, A., Bounous, G., Edwards, K.J., Botta, R. (2003) Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). Mol Breeding 11: 127-136. doi:10.1023/A:1022456013692

Meirmans, P.G., Van Tienderen, P.H. (2004) GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. Mol. Ecol. Notes. 4:792–794. doi:10.1111/j.1471-8286.2004.00770.x.

Peakall, R., Smouse, P.E. (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288–295. doi: 10.1111/j.1471-8286.2005.01155.x

Pereira-Lorenzo, S., Díaz-Hernández, M.B., Ramos-Cabrer, A.M. (2006) Use of highly discriminating morphological characters and isozymes in the study of Spanish chestnut cultivars. J Am Soc Hortic Sci 131:770–779

Pereira-Lorenzo, S., Ríos-Mesa, D., González-Díaz, A.J., Ramos-Cabrer, A.M. (2009) Canary Islands. In: Following chestnut footprints (*Castanea sativa* L.) (Damiato and Bounous eds.). Scripta Horticulturae n° 9, 134-142. International Society for the Horticultural Science. ISBN: 978-90-6605-632-9

Pereira-Lorenzo, S., Costa, R., Ramos-Cabrer, A.M., Ribeiro, C., Serra da Silva, C., Manzano, G., Barreneche, T. (2010) Variation in grafted European chestnut and hybrids by microsatellites reveals two main origins in the Iberian Peninsula. Tree Genet Genomes 6: 701-715. doi: 10.1007/s11295-010-0285-y

Pereira-Lorenzo, S., Lourenço Costa, R.M., Ramos-Cabrer, A.M., Ciordia-Ara, M., Marques Ribeiro, C.A., Borges, O., Barreneche, T. (2011) Chestnut cultivar diversification process in the Iberian Peninsula, Canary Islands, and Azores. Genome 54: 301-315. doi: 10.1139/g10-122

Pereira-Lorenzo, S., Ballester, A., Corredoira, E., Vieitez, A.M., Agnanostakis, S., Costa, R., Bounous, G., Botta, R., Beccaro, G.L., Kubisiak, T.L., Conedera, M., Krebs, P., Yamamoto, T., Sawamura, Y., Takada, N., Gomes-Laranjo, J., Ramos-Cabrer, A.M. (2012) Chestnut Breeding (729-770). In: Fruit breeding / Springer Series "Handbook of Plant Breeding" (M.L. Badenes and D. Byrne eds.), XVI, Springer (2012), 875 pp. ISBN: 978-1-4419-0762-2

Pereira-Lorenzo, S., Ramos-Cabrer, A.M., Barreneche, T., Mattioni, C., Fiorella, V., Díaz-Hernández, M.B., Martín, L.M., Martín, A. (2017). Database of European chestnut cultivars and definition of a core collection using simple sequence repeats. Tree Genetics & Genomes 13: 114. https://doi.org/10.1007/s11295-017-1197-x

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, A., Carracedo, A., Lareus, M.V. (2013). An overview of STRUCTURE: applications, parameter settings, and supporting software. Frontiers in Genetics: 4, 98. doi: 10.3389/fgene.2013.00098.

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

Piry, S., Luikart, G., Cornuet, J.M. (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90: 502–503.

Robin, C., Heiniger, U. (2001) Chestnut blight in Europe: Diversity of *Cryphonectria parasitica*, hypovirulence and biocontrol. For. Snow Landsc. Res. 76, 3: 361–367.

Sarmiento, F.M. (1986). Catálogo de voces vulgares y en especial de voces gallegas de diferentes vegetales. J.L. Pensado (Ed). Salamanca: Ediciones Universidad Salamanca.

Steinkellner, H., Fluch, S., Turetschek, E., Lexer, C., Streiff, R., Kremer, A., Burg, K., Glössl, J. (1997) Identification and characterization of (GA/CT)(n)-microsatellite loci from *Quercus petraea*. Plant Mol Biol 33: 1093-1096. doi: 10.1023/A:1005736722794

Vannini, A., Vettraino, A.M. (2001) Ink disease in chestnuts: impact on the European chestnut. Forest Snow and Landscape Research 76:345–350.

Weir, B.S., Cockerham, C.C. (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370. doi: 10.1111/j.1558-5646.1984.tb05657.x.

Table 1. Classification of 214 genotypes (440 accessions) of *Castanea sativa* by structure (Pritchard et al., 2000) using 18 SSR loci in K =2 and K =4 reconstructed populations, perimeter of the trunks and estimated age for an annual growth of the ring of 0.4 cm and cultivated versus wild trees.

RPP	qI^a	Nº (% respecting to the total 214) of genotypes	Nº (%) of samples	No. of giant trees	Minimum perimeter of the group	Maximum perimeter of the group	Average perimeter of the group	Minimum age of the group for an annual growth of the ring of 0.4cm	Maximum age of the group for an annual growth of the ring of 0.4cm	Average age of the group for an annual growth of the ring of 0.4cm	Number of cultivars referred in the EU database (Pereira-Lorenzo et al., 2017)	Number of unknown cvs.	Main cultivars found in giant trees (perimeters in brackets)	Origen of the genotypes (number in brackets)	Number of trees (% respecting to the total in the group) not grafted (checked as wild comparing the crown with the base of the trunk, including rootstocks)
Giant trees groups per perimeter of the trunk															
11-22 m		29 (13.6)	72 (16.4)	29	10.8	22.00	13.8	1141	1586	1468	4	5	Amarelante', 'Lebre'	Italy (3), Portugal (1), Spain (26)	**14 (51.7)**: Beira Valente (1), Cantabria (3), Castilla-Leon (4), Galicia (5)
8-10 m		39 (18.2)	98 (22.3)	39	7.5	10.45	8.8	1600	1718	1663	5	3	Marrone Fiorentino', 'Courela', 'Negral', 'Rapada', 'Verata'	Italy (5), Spain (34)	**24 (38)**: Cantabria (1), Castilla-Leon (2), Extremadura (12), Galicia (4), Pais Vasco (3), Umbria (2)
4-7		30 (14.0)	62 (14.1)	30	3.8	7.47	6.2	1719	1865	1771	2	1	Luguesa', 'Parede'	Italy (7), Spain (23)	**15 (50)**: Umbria (1), Cantabria (1), Castilla-Leon (5), Extremadura (1), Galicia (1), Pais Vasco (4), Toscana (1)

Genotypes without information about the perimeter		116 (54.2)	209 (47.5)								114	2	All the other cultivars from the EU database (Pereira-Lorenzo et al., 2017)	France (4), Italy (16), Portugal (2), Spain (94)	**2:** Castilla-Leon (2)
At K = 2															
RPP1	>0.8	90 (42.1)	171 (38.9)	39	4	15.9	9.3	1384	1847	1646	55	10	Da Lebre' (11 m), 'Rapada' (10 m), 'Verata' (8 m)	Portugal (3), Spain (87)	**14:** Cantabria (4), Castilla-Leon (1), Extremadura (1), Galicia (1), Galicia (2), Pais Vasco (5)
RPP2	>0.8	54 (25.2)	141 (32.0)	26	6.3	22	9.9	1141	1678	1620	29	1	Marrone Fiorentino' (8.8 m), 'Luguesa' (7 m)	France (4), Italy (28), Spain (22)	**11:** Castilla-Leon (1), Extremadura (8), Toscana (1), Umbria (1)
Admixed	<0.8	70 (32.7)	128 (29.1)	33	3.8	19.0	9.4	1260	1865	1643	39	6	Courelá' (10.1 m), 'Negral' (9.2 m), 'Parede' (6 m)	Italy (3), Portugal (1), Spain (66)	Only considered those admixed for K = 2 and K = 4 (**14**): Castilla-León (9), Extremadura (3), Galicia (1), Pais Vasco (1)
At K = 4															
RPP1a	>0.8	24 (11.2)	41 (9.3)	8	8.2	13.5	10.3	1479	1688	1607	18	2	Verata' (8.2 m)	Andalucia (2), Canary Islands (4), Castilla-León (3), Extremadura (6), Galicia (7), Portugal (2)	**2:** Galicia (2)
RPP1b	>0.8	22 (10.3)	42 (9.5)	11	4.6	12.5	8.2	1517	1833	1691	9	2	Parede' (5.7 m)	Asturias (12), Castilla-Leon (1), Cantabria (2), Galicia (1), Pais Vasco (6)	**9:** Castilla-Leon (1), Cantabria (2), Galicia (1), Pais Vasco (5)

Group	qI[a]	N (%)	N (%)										Examples	Origin	Origin 2
RPP1c	>0.8	20 (9.3)	34 (7.7)	8	5.7	14.8	9.8	1427	1788	1627	15	4	Da Lebre' (11 m), 'Rapada' (10 m)	Asturias (2), Castilla-Leon (1), Galicia (16)	3: Castilla-Leon (1), Galicia (2)
RPP2	>0.8	40 (18.7)	109 (24.8)	20,0	6.3	22,0	9.3	1141	1678	1645	21		Marrone Fiorentino' (8.8 m), 'Luguesa' (7 m)	France (2), Italy (25), Andalucía (7), Extremadura (5), Galicia (1)	9: Umbria (3), Toscana (1), Extremadura (5)
Admixed	<0.8	108 (50.5)	214 (48.6)	51,0	3.8	20,0	9.7	1220	1865	1631	60	8	Courelá' (10.1 m), 'Negral' (9.2 m)	France (2), Italy (6), Portugal (2), Spain (98)	Only considered those admixed for K = 2 and K = 4 (14): Castilla-León (9), Extremadura (3), Galicia (1), Pais Vasco (1)
Cultivated		132 (61.7)	283 (64.3)	20	5.7	14.8	10.6	1427	1788	1595		17	All cultivars from the EU database (Pereira-Lorenzo et al., 2017) and 9 unknown cvs. not cited before but checked to be grafted		
Wild (checked as wild comparing the crown with the base of the trunk, including rootstocks)		55 (25.7)	118 (26.8)	53	3.8	19	9.3	1260	1865	1646					55: Toscana (1), Umbria (4), Cantabria (5), Castilla-Leon (14), Extremadura (13), Galicia (10), Pais Vasco (8)

[a]Probability of membership to an RPP when qI > 80%

Table 2. Number of alleles, average allele number, He , Ho , allelic richness (Ar) with sample sizes used in Ar calculations, and FIS values for 18 SSRs (without null alleles) and 24 SSRs evaluated in this study.

	RPP	qI^a	Nº (%) of genotypes	Nº (%) of accessions	No. of giant trees	18 SSRs Average number of alleles per locus	Ho	He	Ar	Fis		% Retained of the diversity respecting to the wild trees	24 SSRs N. alleles	Ho	He	Ar	Fis		% Retained of the diversity respecting to the wild trees
maximum sample size is 2 x the N value).	**At K = 2**																		
	RPP1	>0.8	90 (42.1)	171 (38.9)	39	6.2	0.621	0.648	5.7	-0.037	**	94.5	144	0.591	0.596	6.1	0.083	**	92.0
	RPP2	>0.8	54 (25.2)	141 (32.0)	26	8.3	0.704	0.682	8.3	0.04	*	99.6	190	0.620	0.688	6.0	0.055	**	106.1
	Admixed	<0.8	70 (32.7)	128 (29.1)	33	7.3	0.684	0.692	7.0	-0.003	ns	100.9	170	0.638	0.651	6.0	0.041	**	100.5
	At K = 4																		
	RPP1a	>0.8	24 (11.2)	41 (9.3)	8	4.7	0.685	0.570	4.5	-0.18	***	100.0	109	0.634	0.548	2.2	-0.136	** *	84.5
	RPP1b	>0.8	22 (10.3)	42 (9.5)	11	4.1	0.569	0.578	4.0	0.008	ns	84.4	97	0.522	0.530	2.2	0.038	ns	81.7
	RPP1c	>0.8	20 (9.3)	34 (7.7)	8	4.3	0.697	0.590	4.3	-0.156	***	101.7	97	0.638	0.550	2.2	-0.133	** *	84.9
	RPP2	>0.8	40 (18.7)	109 (24.8)	20,0	7.9	0.703	0.703	6.6	0.013	ns	102.6	181	0.635	0.689	2.7	0.092	** *	106.3
	Admixed	<0.8	108 (50.5)	214 (48.6)	51,0	7.5	0.669	0.670	5.4	0.003	**	97.8	176	0.615	0.640	2.5	0.044	** *	98.8

Giant trees groups per diameter of the trunk																	
11-22 m	29 (13.6)	72 (16.4)	29	6.2	0.628	0.675	6.2	0.087	***	98.5	146	0.577	0.642	4.9	0.122	***	99.1
8-10 m	39 (18.2)	98 (22.3)	39	6.8	0.692	0.690	6.4	0.01	ns	100.7	158	0.628	0.655	5.1	0.05	**	101.1
4-7	30 (14.0)	62 (14.1)	30	6.4	0.648	0.665	6.4	0.043	*	97.1	148	0.600	0.633	4.9	0.069	**	97.7
Genotypes without information about the perimeter	116 (54.2)	209 (47.5)		8.2	0.680	0.674	6.4	-0.005	ns	98.4	187	0.621	0.650	5.0	0.051	***	100.4
Cultivated	132 (61.7)	283 (64.3)	20	8.4	0.679	0,674	7.2	-0.002	ns	98.5	191	0.622	0.650	5.507	0.048	***	100.4
Wild (checked as wild comparing the crown with the base of the trunk)	55 (25.7)	118 (26.8)	53	6.9	0.665	0,685	6.8	0.038	*		162	0.603	0.648	5.487	0.078	***	

[a]Probability of membership to an RPP when qI > 80%

Figure 1. Classification of 214 European chestnut genotypes including 98 corresponding to giant trees (perimeter of the trunk in m in the upper part of the figure) in reconstructred populations (RPPs) when K =2 (in the middle) and 4 and 8 (in the bottom) based on data of 18 SSR loci.
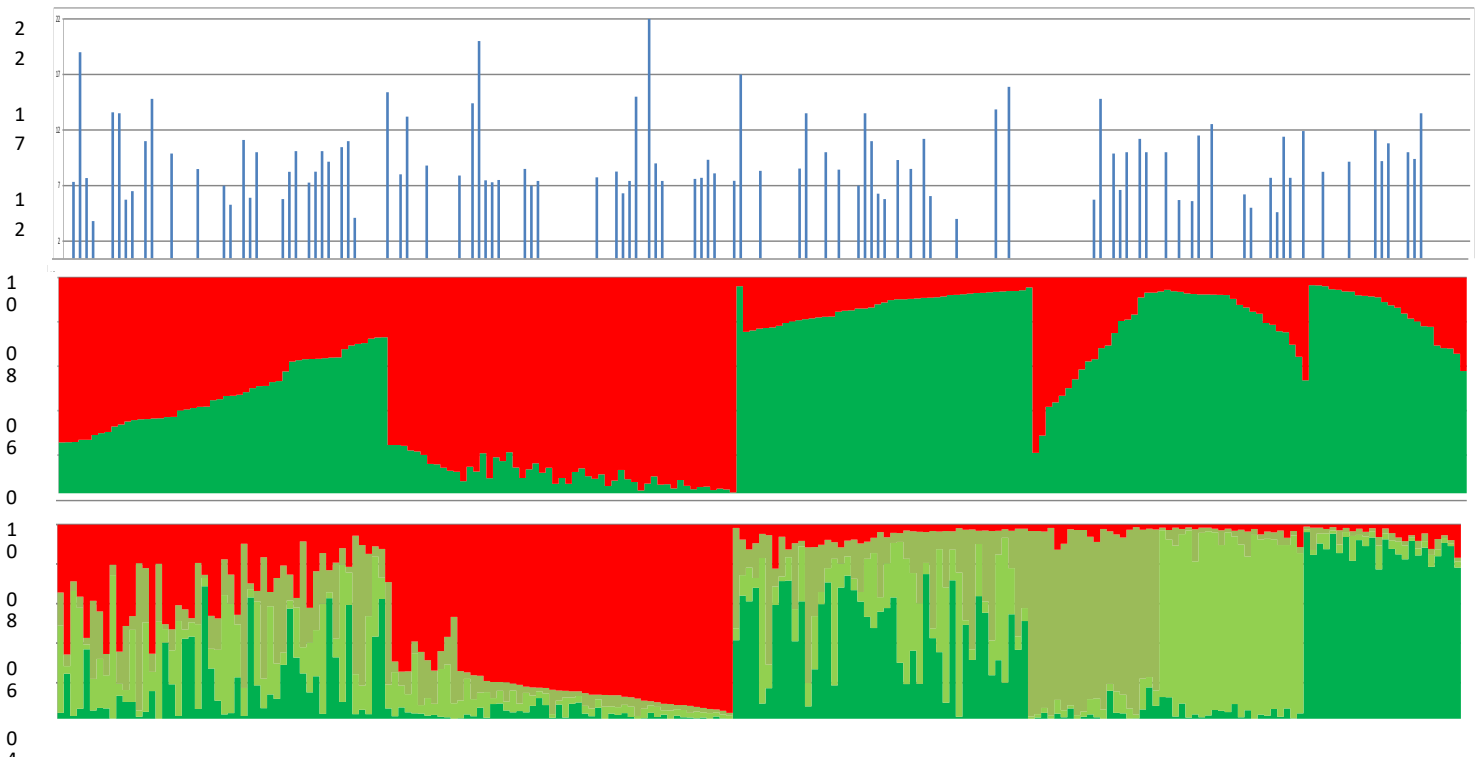
Figure 2. Neighbor-Joining Trees (left) and principal coordinates analysis (PCoA, right) using 18 SSRs in the set of 214 European chestnut genotypes showing structure (K = 2). In red, Italian (I) and French (F) genotypes; in green genotypes from the Iberian Peninsula; and in black, admixed genotypes. Numbers are the perimeters of the giant chestnut trees in m.