Plant Mol Biol Rep (2015) 33:1712–1720
DOI 10.1007/s11105-015-0867-9

ORIGINAL PAPER

Western European Wild and Landraces Hazelnuts Evaluated by SSR Markers

Sandra Martins · Fernanda Simões · Diogo Mendonça · José Matos · Ana Paula Silva · Valdemar Carnide

Published online: 14 March 2015

Abstract Hazelnut (*Corylus avellana* L.), a member of the family Betulaceae and the order Fagales, is native to Europe, Turkey, and the Caucasus. It has been cultivated for more than 5000 years. The majority of cultivars were selected from local wild populations over the past few centuries. Hazelnut is native to northern Portugal, with local production from landrace clonal selections. In recent years, hazelnut production has declined, and landraces are in danger of being lost. In this study, we used 16 microsatellite markers to investigate genetic diversity in wild populations, Portuguese landraces, and reference cultivars. A high level of polymorphism was found, as indicated by mean values for expected heterozygosity (0.74), observed heterozygosity (0.71), and polymorphism information content (0.78). A neighbor-joining dendrogram showed a clear separation of the wild genotypes from the landraces and reference cultivars. Bayesian analysis placed individuals in three groups, showing differentiation of wild populations, landraces, and reference cultivars. Our studies provide new insights on the origin of Portuguese landraces.

Keyword Hazelnut · Microsatellites (SSR) · Genetic diversity · Biodiversity

S. Martins (✉) · V. Carnide
Department of Genetics and Biotechnology, University of Trás-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal
e-mail: sandram@utad.pt

F. Simões · D. Mendonça · J. Matos
Strategic Unit for Biotechnology and Genetic Resources, National Institute for Agrarian and Veterinarian Research, I.P., Avenida da República, Quinta do Marquês Edifício da Ex-Estação Florestal Nacional, 2184-505 Oeiras, Portugal

A. P. Silva
Centre for the Research and Technology of Agro-Environment and Biological Sciences, University of Trás-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal

V. Carnide
Institute for Biotechnology and Bioengineering, Centre of Genomics and Biotechnology, University of Trás-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal

Introduction

European hazelnut (*Corylus avellana* L.) is diploid ($2n=2x=22$), monoecious, dichogamous, and wind pollinated. Sporophytic incompatibility enforces cross-pollination. *C. avellana* is native to Europe and adjacent areas in Asia, including Turkey and the Caucasus Mountains. Black Sea countries account for the majority of world hazelnut production: Turkey (660,000 t), Azerbaijan (29,624 t), and Georgia (24,700 t) (FAO Production Yearbook 2012). Other important producer countries are Italy (85,232 t), USA (30,000 t), and Spain (13,900 t) (FAO Production Yearbook 2012). Over many centuries, the cultivars in Europe and Turkey were selected from local wild populations. Identification of hazelnut cultivars is primarily based on analysis of nuts, husks, and other morphological traits. However, these traits are often unreliable or imprecise indicators of plant genotype, being influenced by environmental factors. Thus, discrimination among closely related cultivars and clones is often extremely difficult. Microsatellites or simple sequence repeats (SSRs) are molecular markers consisting of tandemly repeated units of short nucleotide motifs, 1- to 6-bp long. These regions occur frequently and randomly throughout the genomes of plants and animals and typically show extensive variation within and among individuals, populations, and species (Jarne and Lagoda 1996). SSR markers have become valuable molecular tools for fingerprinting genotypes, assessment of genetic

diversity in collections and linkage mapping, having the advantage of being often transferable among closely related species and sometimes related genera. This feature of SSRs combined with their hypervariability make them ideal for a wide range of applications in fruit and nut crops such as cultivar identification, management of germplasm collections, identification of duplicate accessions, and evaluation of genetic diversity. SSR markers identified in *C. avellana* (Bassil et al. 2005a, b; Boccacci et al. 2005; Gürcan and Mehlenbacher 2010a, b; Gürcan et al. 2010a, b; Campa et al. 2011) were used for linkage mapping (Mehlenbacher et al. 2006; Gürcan et al. 2010a, b) to assess genetic relationships among cultivars (Boccacci et al. 2006, 2008; Ghanbari et al. 2005; Gökirmak et al. 2009; Boccacci and Botta 2010; Gürcan et al. 2010a, b) and to fingerprint cultivars in collections, identify synonyms, and determine parentage (Botta et al. 2005; Gökirmak et al. 2009; Sathuvalli and Mehlenbacher 2011). Cross-species transferability of a number of SSRs was demonstrated in *Corylus* (Bassil et al. 2005a; Boccacci et al. 2005) and within the Betulaceae family maintaining the polymorphic feature of the molecular marker (Gürcan and Mehlenbacher 2010b). In the modern agriculture of recent decades, the planting of cultivars with high yield and good fruit quality has led to the loss of landraces. To avoid loss of local germplasm, efforts are needed on a worldwide scale to explore, collect, characterize, and preserve genetic diversity, especially in the vegetatively propagated crops including fruits and nuts (Erfatpour et al. 2011). In Portugal, hazelnut production has suffered a sharp decline since the 1990s, which entailed an abandonment of the culture in favor of other agronomic cultures, leading to the loss of landrace genotypes. This loss of landrace genotypes may reduce the genetic diversity of residual populations, which is potentially catastrophic since genetic diversity is the basis of all biodiversity and is widely recognized as a key requirement for the long-term survival of species on an evolutionary timescale. Genetic diversity provides the template for adaptation, evolution, and survival of populations and species, especially in environments that are subject to climate changes or to the introduction of new pests, pathogens, or competitors. Landraces and reference cultivars are the most frequently studied specimens. However, wild populations may still be an important source of genetic diversity. Genetic diversity has been studied in wild populations in specific areas of Spain (Campa et al. 2011) and Italy (Boccacci et al. 2013). In Portugal, the number of landraces has decreased in recent decades and few have been located, identified, or characterized and their origins have not been documented. In this work, Portuguese wild hazelnuts were for the first time located and sampled, and SSR markers used to elucidate their genetic relationships with Portuguese landraces and reference cultivars.

Material and Methods

Plant Material

A total of 58 accessions were used in this study: 19 landraces and 13 wild accessions from Portugal and 26 reference cultivars that originated in seven countries (England, USA, Spain, Italy, Germany, Bulgaria, and Turkey) (Table 1). The 19 landraces were collected within an area of approximately 27, 000 km^2, located between Northern and Central Portugal (latitude 40.30 to 41.70 N and longitude 6.29 to 8.35 W) (Fig. 1). This area was recognized as an area of hazelnut production in olden times. The identification and location of Portuguese landrace genotypes was carried out by collecting oral testimonials from farmers and older people in the region. Testimony confirmed that the Portuguese landraces had existed in the region for >100 years. From each site, Portuguese landrace showing different morphological characteristics were collected. We searched for wild trees in natural (Alvão and Douro International) and national (Peneda-Gerês) parks with the collaboration of forest keepers. Such parks were preferred sampling sites because the genotypes were known to be wild and contribute to the conservation value of parks in general.

DNA Isolation

Immature catkins were collected in the field, kept cool during the transport, and frozen in a −80 °C freezer. Total genomic DNA was extracted by using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. DNA concentration was determined by both spectrophotometry at 260 nm and gel electrophoresis, using a known quantity of Lambda DNA digested with *Hind*III.

SSR Analysis

Sixteen SSR loci, three developed in Corvallis, USA [CAC-A102, CAC-B020, CAC-B028,] (Bassil et al. 2005a, b) and 13 developed in Torino, Italy [CaT-A114, CaT-B107, CaT-B501, CaT-B502, CaT-B503, CaT-B504, CaT-B505, CaT-B507, CaT-B508, CaT-B509, CaT-B511, CaT-C001, and CaT-C504] (Boccacci et al. 2005) were selected for genetic diversity evaluation. Forward primers fluorescently labelled with FAM or HEX and reverse primers were purchased from Metabion (Germany). Polymerase chain reactions were carried out in a Biometra (UNO II-thermoblock) thermocycler and mixtures of the reactions contained 12.5 μl of 2× Qiagen multiplex PCR Master Mix (QIAGEN Multiplex PCR Kit), 2.5 μl Q-Solution, 0.125 μl each of forward and reverse primers, and 50 ng of template DNA in a total volume of 25 μl. The PCR program consisted of an initial denaturation of 15 min at 94 °C followed by 35 cycles of a 30 s denaturation step at 94 °C, a 90s annealing step at the optimum

Table 1 List of *Corylus avellana* accessions used in this study

Accessions	Group	Name	Origin
Ca 1	Landrace	Unknown	Portugal
Ca 2	Landrace	Unknown	Portugal
Ca 3	Landrace	Unknown	Portugal
Ca 4	Landrace	Unknown	Portugal
Ca 5	Landrace	'Grada de Viseu'	Portugal
Ca 6	Landrace	Unknown	Portugal
Ca 7	Landrace	'Veiga'	Portugal
Ca 8	Landrace	'Comum'	Portugal
Ca 9	Landrace	Unknown	Portugal
Ca 10	Landrace	Unknown	Portugal
Ca 11	Landrace	'Tubulosa'	Portugal
Ca12	Landrace	'Rubra'	Portugal
Ca 17	Landrace	Unknown	Portugal
Ca 18	Landrace	Unknown	Portugal
Ca 19	Landrace	Unknown	Portugal
Ca 20	Landrace	Unknown	Portugal
Ca 21	Landrace	Unknown	Portugal
Ca 22	Landrace	Unknown	Portugal
Ca 25	Landrace	'Molar'	Portugal
Ca 23	Wild		Portugal
Ca 24	Wild		Portugal
Ca 28	Wild		Portugal
Ca 29	Wild		Portugal
Ca 30	Wild		Portugal
Ca 31	Wild		Portugal
Ca 32	Wild		Portugal
Ca 33	Wild		Portugal
Ca 34	Wild		Portugal
Ca 35	Wild		Portugal
Ca 36	Wild		Portugal
Ca 37	Wild		Portugal
Ca 38	Wild		Portugal
Ca 13	Reference	'Butler'	USA
Ca 14	Reference	'L. d'Espagne'	UK
Ca 16	Reference	'Merveille de Bollwiller'	Germany
Ca 26	Reference	'Ran Trapezunski'	Bulgaria
Ca 27	Reference	'Rimski'	Bulgaria
Ca 43	Reference	'Nottingham'	UK
Ca 44	Reference	'Sivri'	Turkey
Ca 45	Reference	'Garibaldi'	Germany
Ca 46	Reference	'Marxant-1'	Spain
Ca 47	Reference	'San Giovanni'	Italy
Ca 48	Reference	'Martorella'	Spain
Ca 49	Reference	'Karidaty'	Turkey
Ca 50	Reference	'Comun Aleva'	Spain
Ca 51	Reference	'Ennis'	USA
Ca 52	Reference	'Bard'	UK
Ca 53	Reference	'Tonda bianca'	Italy
Ca 54	Reference	'Gunslebert Zellernuss'	Germany
Ca 55	Reference	'Incekara'	Turkey
Ca 56	Reference	'Jardinera'	Spain
Ca 57	Reference	'Rote Zellernuss'	Germany
Ca 58	Reference	'Jemtegaard #5'	USA
Ca 59	Reference	'Butler'	USA
Ca 60	Reference	'Mortarella'	Italy
Ca 61	Reference	'Nocchione'	Italy
Ca 62	Reference	'Pere Mas'	Spain
Ca 63	Reference	'Contorta'	UK

annealing temperature, and a 30s extension step at 72 °C. There was a final 10 min extension step at 72 °C. Amplification and approximate fragment sizes were confirmed by agarose (2 %) gel electrophoresis using TBE buffer and stained with ethidium bromide. PCR products were separated on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). DNA fragment sizes were determined using GeneMapper software (Applied Biosystems).

Data Analysis

The codominant SSR data were analyzed by MICRO-CHECKER software (Van Oosterhout et al. 2004) for detecting null alleles at each locus for each population. Population genetic analysis was performed using GenAlEx software package version 6.5 (Peakall and Smouse 2006) to calculate the number of alleles, expected heterozygosity (H_e), observed heterozygosity (H_o), Shannon's diversity index (I), and the analysis of molecular variance (AMOVA) in order to partition genetic variation among and within populations (Schneider et al. 2000). The significance of each variance component was tested with permutation tests (Excoffier and Slatkin 1992). The PIC value of a locus, which ranges from 0 (monomorphic) to 1 (highly informative), was calculated using The Excel Microsatellite Toolkit (Park 2001). Wright's F_{ST} was used to estimate population differentiation and was calculated using GenAlEx software package version 6.5 (Peakall and Smouse 2006). Genetic distances were estimated according to Nei (1978) and principal coordinate analyses (PCoA) (Gower 1966) were performed to identify genetic variation patterns among the hazelnut genotypes. Genetic similarity matrices based on the proportion of shared alleles and neighbor-joining cluster analysis were used to construct genetic trees using Populations software (Langella 2011). A dendrogram was created using Mega 5.2 software (Tamura et al. 2011). STRUCTURE ver. 2.3 software (Pritchard et al. 2000) was used to investigate the genetic population structure of hazelnut genotypes. This method uses a Markov Chain Monte Carlo (MCMC) algorithm to cluster individuals into

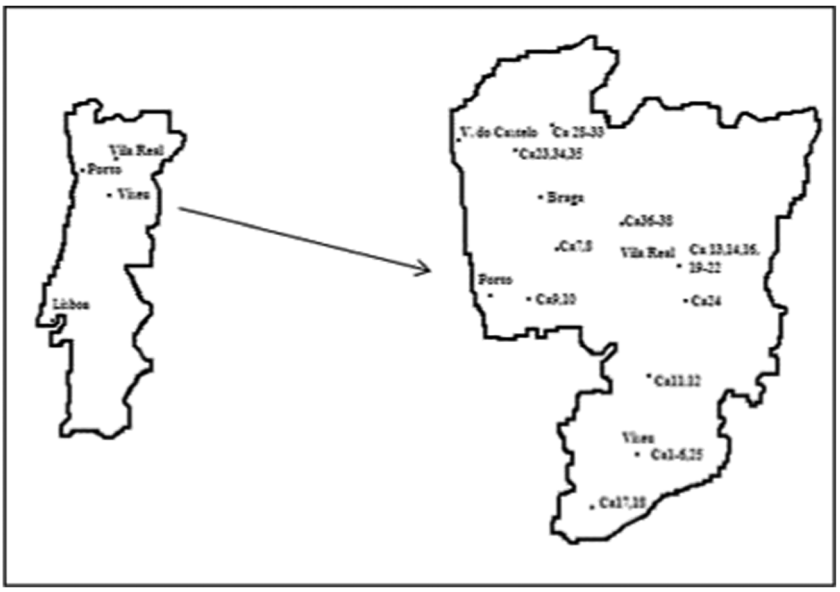

Fig. 1 Map of Northern Portugal showing sampling sites

populations on the basis of multi-locus genotype data (Pritchard et al. 2000; Falush et al. 2003). For STRU CTURE analysis, the accessions were initially assigned to seven groups: Portuguese landraces, Portuguese wild, and five groups of reference cultivars according to their geographic origin (central Europe, southern Europe, Black Sea, UK, and USA). The number of populations (*K*) was estimated by performing ten independent runs for each *K* (from 1 to 8), using 1,000,000 MCMC repetitions and 50,000 burn-in periods. Any prior information about the population of origin was used, and correlated allele frequencies and admixture were assumed. The average of the log-likelihood estimates for each *K* was calculated. The ad hoc statistic DK (Evanno et al. 2005) was used to set the number of populations (*K*).

Results

A total of 58 genotypes were successfully amplified by each of the 16 SSR primer pairs generating a total of 166 alleles. The number of alleles per locus ranged from eight (CaT-B509, CaT-B020, CaT-B501, and CaT-A114) to 17 (CaT-B503) with an average of 10.4. The highest frequency allele was found at locus CaT-B501 (0.719). Fifty-three private alleles were observed, representing 32 % of the total. This high number of private alleles indicates high genetic diversity in the genotypes used for characterization. Seven private alleles were found in the group of the landraces, 15 private alleles in the group of wild hazelnuts, and the remaining private alleles were in the group of reference cultivars. PIC values ranged from 0.879 to 0.500 with an average of 0.778, which indicates that the chosen loci are highly informative (Table 2). The observed heterozygosity (H_o) values for individual loci ranged from 0.499 to 0.880, with an average of 0.711, while expected heterozygosity (H_e) ranged from 0.616 to 0.798 with an average of 0.738 (Table 2). The high level of heterozygosity is a consequence of the self-incompatibility of this species. The highest average number of different alleles for all loci (Na) was found in the group of reference cultivars (8.688) and the lowest value was for the group of landraces (6.313). The group of the landraces had the lowest value of Shannon's information index (*I*) (1.390) and the reference cultivars had

Table 2 Characterization of 16 simple sequence repeat (SSR) loci based on 58 hazelnut accessions

SSR locus	*N*	H_e	H_o	*I*	PIC
CaT-B503	17	0.750	0.725	1.658	0.863
CaT-B502	10	0.737	0.738	1.608	0.727
CaC-B020	8	0.704	0.787	1.424	0.741
CaT-B107	12	0.730	0.832	1.535	0.768
CaT-B507	9	0.780	0.880	1.710	0.850
CaC-A102	10	0.726	0.715	1.585	0.774
CaT-B505	10	0.771	0.632	1.679	0.831
CaT-B511	9	0.768	0.616	1.656	0.795
CaT-A114	8	0.718	0.711	1.497	0.753
CaT-B509	8	0.707	0.720	1.450	0.717
CaT-B508	12	0.681	0.685	1.506	0.740
CaC-B028	11	0.770	0.677	1.686	0.808
CaT-B501	8	0.616	0.530	1.199	0.622
CaT-B504	10	0.762	0.859	1.689	0.808
CaT-C001	10	0.788	0.776	1.719	0.827
CaT-C504	14	0.798	0.499	1.857	0.827
Total	166				
Average	10.375	0.738	0.711	1.591	0.778

N n° of alleles, H_o observed heterozygosity, *I* Shannon's information index, *PIC* polymorphic information content

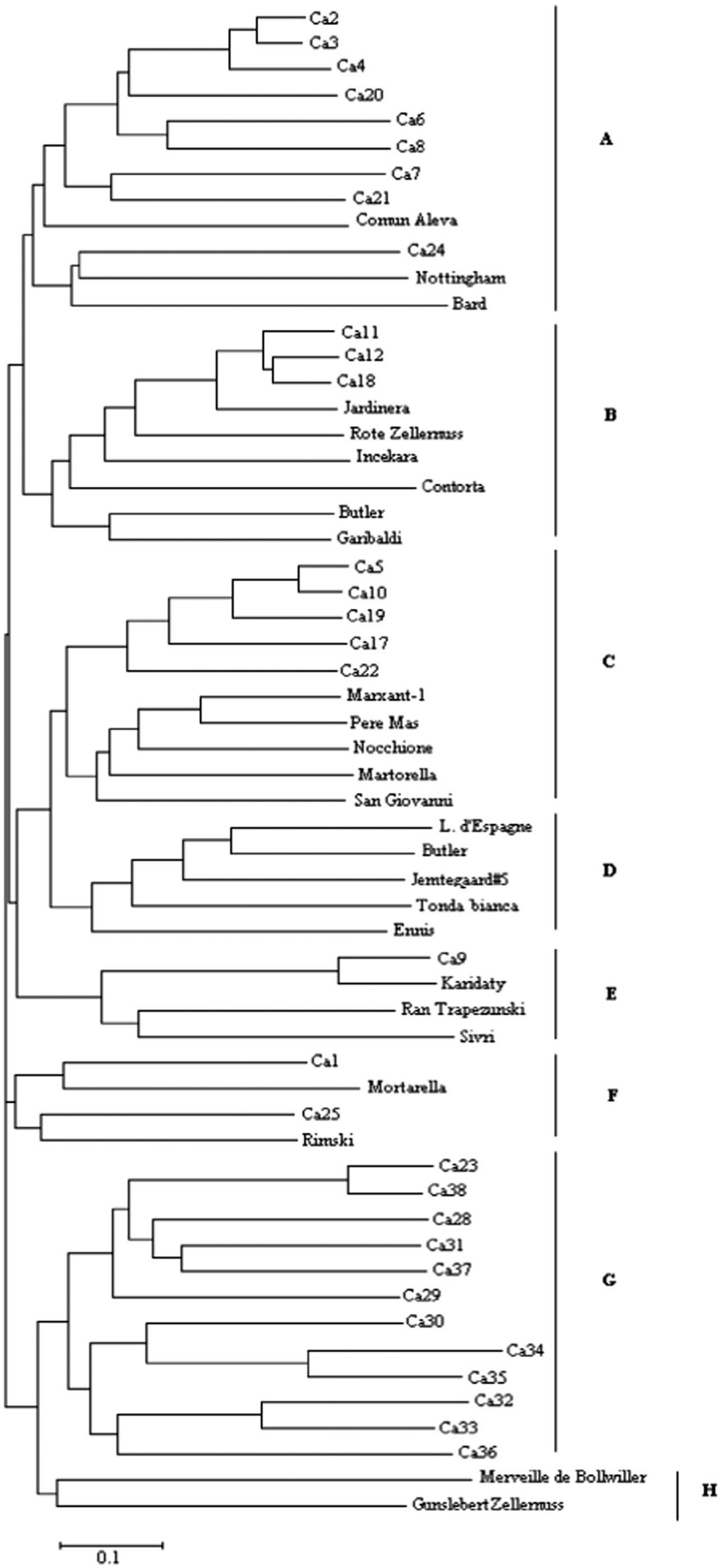

Fig. 2 NJ cluster analysis of hazelnut accessions based on the proportion of shared allele distance for 16 dinucleotide containing SSRs

the highest value (1.792). MICRO-CHECKER detected null alleles at some loci and PCR amplification was repeated for all homozygotes at these loci. In the neighbor-joining dendrogram based on shared allele distance (Fig. 2), the genotypes showed considerable spread and were placed in eight clusters (A to H). The group A contains eight landraces (Ca2, Ca3, Ca4, Ca6, Ca7, Ca8, Ca20, and Ca21), one wild genotype (Ca24) and three reference cultivars: one from Spain ('Comun Aleva, syn. Segorbe') and two British reference cultivars ('Nottingham' and 'Bard'). Group B is a mixture of reference cultivars and landraces from different geographical areas. The three landraces Ca11, Ca12, and Ca18 appeared adjacent to the reference cultivars 'Jardinera' (Spain), 'Rote Zellernuss' from (Germany), and 'Incekara' (Turkey) and *C. avellana* var. contorta. The C group contains five Portuguese landraces (Ca5, Ca10, Ca17, Ca19, and Ca22) and five reference cultivars, two from Italy ('San Giovanni' and 'Nocchione') and three from Spain ('Pere Mas', 'Marxant-1', and 'Martorella'). Three American reference cultivars ('Butler', 'Jemtegaard #5', and 'Ennis'), one British variety ('Longe d'Espagne') and one Italian variety ('Tonda bianca') compose the D group. Group E is composed of 'Sivri' and 'Karidaty' from Turkey, 'Ran Trapezunski' from Bulgaria and landrace Ca9. Group F is composed of 'Mortarella' from Italy, 'Rimski' from Bulgaria, and landraces Ca1 and Ca25. Group G is composed of only wild genotypes. Finally, group H contains the German cultivars 'Gunslebert Zellernus' and 'Merveille de Bollwiller'.

In the PCoA, the first two PCs explained 28.8 % of the total variation. The first coordinate explained 12.8 % of the variation and the second coordinate an additional 9.3 %. The projection of 58 hazelnut accessions on a two-dimensional plane defined by the first two PCs (Fig. 3) also showed a tendency to separate the wild genotypes from cultivated accessions. The pairwise values for F_{ST} ($P<0.001$) ranged from 0.048 (reference cultivars and landraces) to 0.152 (landraces and wild) as shown in Table 3. The landraces and reference cultivars showed the lower genetic differentiation ($F_{ST}=0.048$). However, the genetic differentiation between reference cultivars and the wild group ($F_{ST}=0.088$) was still higher than the

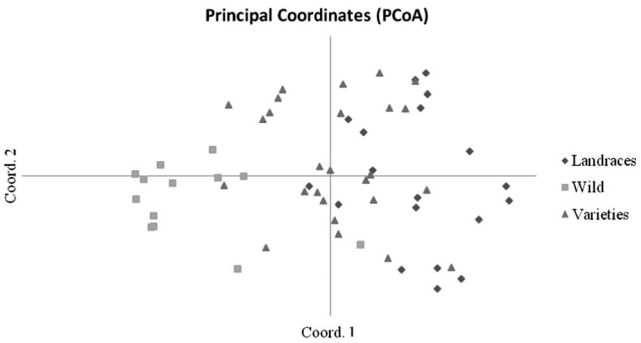

Fig. 3 Principal coordinates analysis of 58 unique hazelnut accessions shown by groups

Table 3 Pairwise values for F_{ST} for three hazelnut groups

	Landraces	Wild
Wild	0.152	0
Reference	0.048	0.088

genetic differentiation between the landraces and reference cultivars showing a moderate genetic differentiation. The AMOVA showed that only 9 % of the genetic variation was among groups, and the remaining 91 % was among individuals within groups. The 58 hazelnut genotypes were further evaluated for population stratification using the STRUCTURE software. SSR data were analyzed increasing the number of subpopulations (K) from 1 to 8. The estimation of ΔK revealed the highest value for $K=3$ ($\Delta K=832.8$), indicating the existence of three groups, composed mainly of landraces, wild, and reference cultivars (Fig. 4). Notably, a unique wild genotype (Ca24) showed a high similarity to landraces and reference cultivars (according to dendrogram clustering). These three initially identified groups remained almost constant, whereas at $K=4$ ($\Delta K=6.375$) several landraces showed the tendency to constitute a subgroup with some reference cultivars and at $K=5$ ($\Delta K=1.211$) the groups formed remained almost constant, but reference cultivars seem to have a tendency to form a subgroup (Fig. 4). Although the reference cultivars have diverse geographic origins, we present results for $K=3$ as the statistic ΔK showed the highest value at $K=3$. Comparing these results with the neighbor-joining dendrogram and the PCA scatterplot, there was general agreement about the clustering of wild genotypes and cultivated forms.

Discussion

One major application of microsatellite markers in plant species (cultivated and wild) is fingerprinting of individuals and thereby distinguishing among them. The level of genetic diversity ($H_e=0.74$), averaged over all groups, was similar to previous reports by Gökirmak et al. (2009) at 21 SSR loci ($H_e=0.72$), Campa et al. (2011) at 13 SSR loci ($H_e=0.73$), Erfani et al. (2012) at 19 SSR loci ($H_e=0.67$), Boccacci et al. (2013) and Boccacci and Botta (2010) at 10 and 16 SSR ($H_e=0.71$ and $H_e=0.72$, respectively). The high level of heterozygosity may be a consequence of the biological features of hazelnut, including sporophytic incompatibility, dichogamy, and the vegetative propagation of superior genotypes (Gürcan et al. 2010a, b). Wright (1978) suggested that F_{ST} values between 0.05 and 0.15 indicated moderate genetic differentiation while values over 0.25 indicated high differentiation. Therefore, the F_{ST} value found between the Portuguese landraces and the reference cultivars (0.048) indicates a low differentiation. The F_{ST} value was expected, since reference cultivars were selected from local populations in

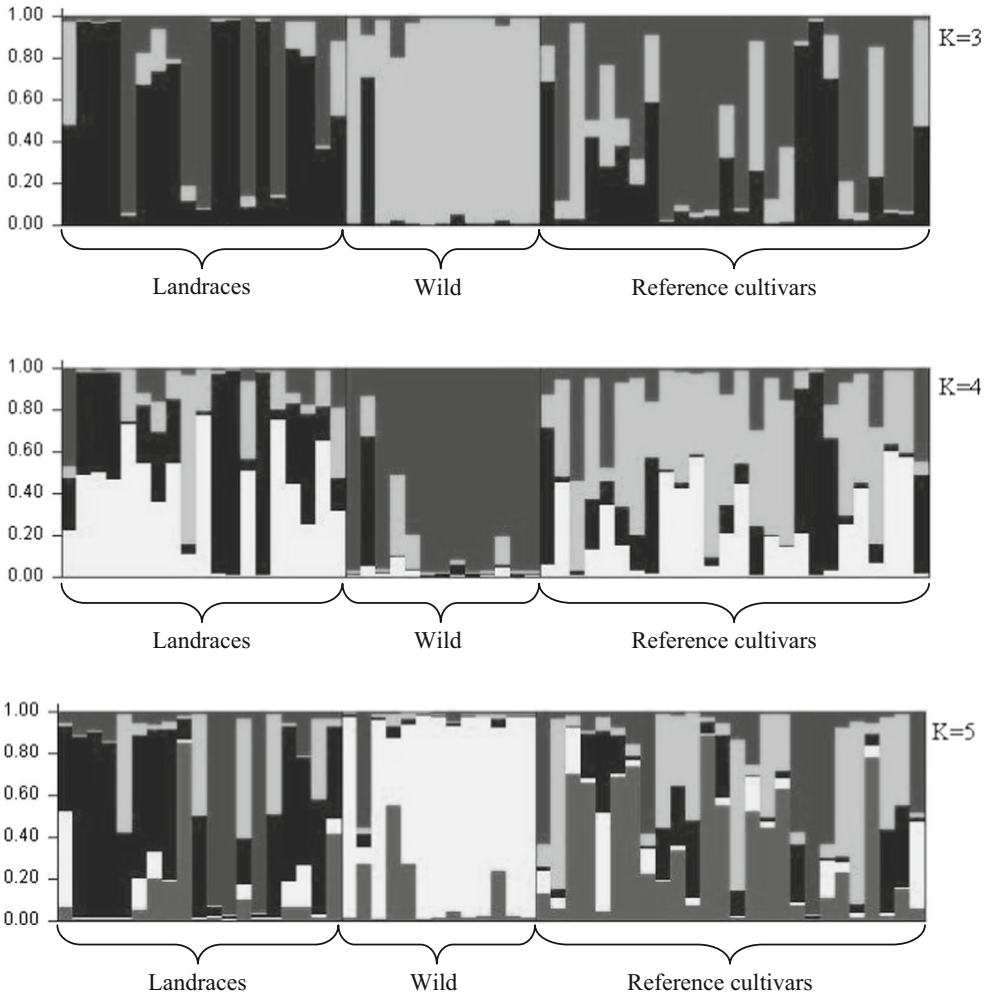

Fig. 4 Hierarchical organization of genetic relatedness of 58 hazelnut genotypes based on 16 SSR markers and analyzed by the STRUCTURE program, with three, four, and five populations (K=3, K=4, and K=5)

various countries. This low differentiation between landraces and reference cultivars may be an indicator of a common origin of some landraces and reference cultivars. The high genetic differentiation (F_{ST}=0.152) between the landraces and wild group was expected and is comparable to the results of Boccacci et al. (2013) and Martins et al. (2014). The Peneda-Gerês and Alvão mountains, where hazelnuts thrive, may act as a natural barrier to plant migration. However, only a moderate genetic differentiation was found between reference cultivars and wild genotypes (F_{ST}=0.088). This moderate value was unexpected, as the Portuguese wild populations are separated by great distance from the countries where the reference cultivars were obtained. However, Martins et al. (2014) with AFLPs and ISSRs reported a high genetic differentiation (F_{ST}=0.503) between Portuguese wild genotypes and cultivars. An explanation for this moderate differentiation can be the sharing of a common genetic background concerning neutral markers while AFLP and ISSR markers span all over the genome and may reflect the influence of adaptive markers. These data are supported by STRUCTURE and the hierarchical analysis. Thus, a significant genetic differentiation among gene pools has occurred, and each group constitutes an independent source of genetic variability and a valuable resource of genetic traits for hazelnut breeders. The hierarchical analysis showed a clear separation of the wild genotypes from the cultivated genotypes. Similarly, Campa et al. (2011), Boccacci et al. (2013), and Martins et al. (2014) found a clear separation of wild genotypes from cultivated forms. In our study, the wild genotypes clustered together, except for Ca24. Martins et al. (2013), based on chloroplast microsatellite loci, observed that most of the wild genotypes had unique haplotypes (H, I, J, and L) but Ca24 shares the most common chloroplast haplotype with the landraces. Thus, the data reinforce the hypothesis that wild genotypes constitute an independent source of genetic variability (hotspot of diversity) and a valuable resource of genetic traits for hazelnut. In group B, we can find two interesting clusters. The landraces and 'Jardinera' (Spanish cultivar) cluster support the idea of gene flow between Portugal and Spain and the 'Butler' and 'Garibaldi' cluster was also reported by Gökirmak et al. (2009). The same applies to five Portuguese landraces from C group that clustered together with Spanish

and Italian cultivar. Also, Gökirmak et al. (2009), Gürcan et al. (2010a, b) and Boccacci and Botta (2010) observed that Spanish and Italian reference cultivars clustered together. In fact, human migrations and trade between the Italian and Iberian Peninsula were frequent during and after the Roman civilization and archeological findings indicate that hazelnut cultivation and consumption were significant in Portugal during the Roman period (Alarcão 1989; Mateus and Queiroz 1993). This can explain the clustering of five of the nineteen Portuguese landraces with the Spanish and Italian reference cultivars.

The population substructure within *C. avellana* was best described through standard structure analysis at $K=3$, where three possible subgroups were detected. Briefly, SSR markers clearly differentiate wild genotypes in a more homogeneous cluster while landraces and reference cultivars of *C. avellana* show some genetic admixture sharing high coefficients of ancestry.

Similar results were obtained by Campa et al. (2011) for 40 wild hazelnuts from northern Spain and 62 locally cultivated accessions using 13 SSR markers, Boccacci et al. (2013) for 118 accessions (landraces, wild, and cultivars) using 10 SSR and by Martins et al. (2014) for 58 accessions (landraces, wild, and cultivars) using AFLPs and ISSRs. We believe that the admixture found among landraces and reference cultivars is due to the introduction of new cultivars from abroad when hazelnut production was booming in Portugal. These results are in agreement with the general idea that most currently important hazelnut cultivars were selected over centuries from local wild populations and a few were spread outside their areas of origin by trade and human migration (Thompson et al. 1996). Nowadays, hazelnut production is declining in Portugal almost to the level of abandonment which may result in loss of landrace genotypes. The study of the genetic relationships and population structure among wild forms, landraces, and cultivars in a geographic area can supply information about the putative domestication events, the evolutionary relationships or the gene flow between them. The neighbor-joining tree (Fig. 2), the PCoA scatter plot (Fig. 3), and the STRUCTURE analyses (Fig. 4) revealed a high level of differentiation between wild and both cultivated forms. The European hazelnut cultivars had a multiple origins in which domestication could have involved a large number of founders over an extended time period and along the entire distribution range of the wild progenitor species. Our results do not reveal that Portugal was one of the domestication centers of hazelnuts. The high genetic diversity found in our work among wild genotypes may indicate that Portugal was also a refuge during the last ice age. In biogeography, refugia are defined as geographical areas in which genetic and species diversity have been preserved despite drastic climatic fluctuations in the past (Petit et al. 2003; Médail and Diadema 2009). Also, Tzedakis et al. (2002) and Petit et al. (2003) suggested that regions showing high levels of diversity could be identified as a potential refuge. In the case of Portuguese wild genotypes, the geographic isolation provided a very low gene flow with surrounding areas. We strongly believe that genetic diversity data from wild and landrace genotypes may provide information useful for conservation of diversity and selections useful for breeding. The low number of wild individuals found must be considered an alert sign for the need of short- and long-term diversity conservation, beyond breeding programs. These results reinforce our idea that Portugal may also have been a natural reservoir for the wild genotypes. Furthers studies using ancient DNA from archeologic samples may provide data to prove the existence of a refuge for wild *C. avellana*. It is widely known that the evaluation of genetic diversity in natural populations is important for the management and utilization of plant collections which play a vital role in the survival and adaptability of species. Knowledge of genetic variation and genetic relationships among genotypes should be considered in the conservation and utilization of hazelnut genetic resources.

Acknowledgments This work was supported by the Ph.D. grant SFRH/BD/40686/2007 attributed by the Portuguese Foundation of Science and Technology (FCT).

Conflict of Interest The authors declare that they have no competing interests.

Authors' Contributions All authors wrote, read, and approved the final manuscript.

References

Alarcão J (1989) O Domínio Romano. O Reordenamento Territorial, Nova História de Portugal. J Serrão e AH de Oliveira Marques. In: Editorial Presença (eds), vol 1, pp 345–382

Bassil NV, Botta R, Mehlenbacher SA (2005a) Microsatellite markers in hazelnut: isolation, characterization and cross-species amplification. J Am Soc Hortic Sci 130:543–549

Bassil NV, Botta R, Mehlenbacher SA (2005b) Additional microsatellite markers of the European hazelnut. Acta Hortic 686:105–110

Boccacci P, Botta R (2010) Microsatellite variability and genetic structure in hazelnut (*Corylus avellana* L.) cultivars from different growing regions. Sci Hortic 124:128–133

Boccacci P, Akkak A, Bassil NV, Mehlenbacher SA, Botta R (2005) Characterization and evaluation of microsatellite loci in European hazelnut (*C. avellana*) and their transferability to other *Corylus* species. Mol Ecol Notes 5:934–937

Boccacci P, Akkak A, Botta R (2006) DNAtyping and genetic relationships among European hazelnut (*Corylus avellana* L.) cultivars using microsatellite markers. Genome 49:598–611

Boccacci P, Botta R, Rovira M (2008) Genetic diversity of hazelnut (*Corylus avellana* L.) germplasm in northeastern Spain. HortScience 43:667–672

Boccacci P, Aramini M, Valentini N, Bacchetta L, Rovira M, Drogoudi P, Silva AP, Solar A, Calizzano F, Erdoğan V, Cristofori V, Ciarmiello LF, Contessa C, Ferreira JJ, Marra FP, Botta R (2013) Molecular and morphological diversity of on-farm hazelnut (*Corylus avellana* L.) landraces from southern Europe and their role in the origin and

diffusion of cultivated germplasm. Tree Genet Gen 9:1465–1480. doi:10.1007/s11295-013-0651-7

Botta R, Akkak A, Boccacci P (2005) DNA-typing of hazelnut: a universal methodology for describing cultivars and evaluating genetic relatedness. Acta Hort 686:117–124

Campa A, Trabanco N, Pérez-Vega E, Rovira M, Ferreira JJ (2011) Genetic relationship between cultivated and wild hazelnuts (*Corylus avellana* L.) collected in northern Spain. Plant Breed 130:360–366

Erfani J, Ebadi A, Abdollahi H, Fatahi R (2012) Genetic diversity of some pear cultivars and genotypes using simple sequence repeat (SSR) markers. Plant Mol Biol Rep 30:1065–1072. doi:10.1007/s11105-012-0421-y

Erfatpour M, Hamidogli Y, Kaviani B, Fatahi R, Falahati M, Javadi D, Hashemabadi D (2011) Assessment of genetic diversity among some Iranian hazelnut genotypes using SSR markers. Aust J Crop Sci 5(10):1286–1291

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Slatkin M (1992) Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol and Evol 12:921–927

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure: extensions to linked loci and correlated allele frequencies. Genetics 164:1567–1587

FAO production yearbook (2012) http://faostat.fao.org/site/339/default.aspx. Accessed 22 August 2014

Ghanbari A, Akkak A, Boccacci P, Talaie A, Vezvaie A, Botta R (2005) Characterization of hazelnut (*Corylus avellana* L.) cultivars using microsatellite markers. Acta Hort 686:111–115

Gökirmak T, Mehlenbacher SA, Bassil NV (2009) Characterization of European hazelnut (*Corylus avellana*) cultivars using SSR markers. Genet Resour Crop Evol 56:147–172

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338

Gürcan K, Mehlenbacher SA (2010a) Transferability of microsatellite markers in the Betulaceae. J Am Soc Hort Sci 135(2):159–173

Gürcan K, Mehlenbacher SA (2010b) Development of microsatellite marker loci for European hazelnut (*Corylus avellana* L.) from ISSR fragments. Mol Breed 26:551–559

Gürcan K, Mehlenbacher SA, Erdoğan V (2010a) Genetic diversity in hazelnut cultivars from Black Sea countries assessed using SSR markers. Plant Breed 129:422–434. doi:10.1111/j.1439-0523.2009.01753.x

Gürcan K, Mehlenbacher SA, Botta R, Boccacci P (2010b) Development, characterization, segregation, and mapping of microsatellite markers for European hazelnuts (*Corylus avellana* L.) from enriched genomic libraries and usefulness in genetic diversity studies. Tree Genet Gen 6:513–531

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. Trends Ecol Evol 11:424–429

Langella O (2011) Populations 1.2.31 CNRS UPR9034

Martins S, Simões F, Mendonça D, Matos J, Silva AP, Carnide V (2013) Chloroplast SSR genetic diversity indicates a refuge for *Corylus avellana* in Northern Portugal. Genet Resour Crop Evol 60:1289–1295. doi:10.1007/s10722-012-9919-2

Martins S, Simões F, Mendonça D, Matos J, Silva AP, Carnide V (2014) Genetic relationship among wild, landraces and cultivars of hazelnut (*Corylus avellana*) from Portugal revealed through ISSR and AFLP markers. Plant Syst and Evol 300:1035–1046. doi:10.1007/s00606-013-0942-3

Mateus J, Queiroz P (1993) Os estudos de vegetação quaternária em Portugal: contextos, balanço de resultados, perspetivas. O quaternário em Portugal. Balanço e perspetivas. APEQ (ed.). Colibri

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. J Biogeogr 36:1333–1345

Mehlenbacher SA, Brown RN, Nouhra ER, Gökirmak T, Bassil NV, Kubisiak TL (2006) A genetic linkage map for hazelnut (*Corylus avellana* L.) based on RAPD and SSR markers. Genome 49:122–133

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Park SDE (2001) Trypanotolerance in West African cattle and the population genetic effects of selection. Ph.D. thesis, University of Dublin

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Note 6:288–295

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Martín JP, Rendell S, Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563–1565

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

Sathuvalli SR, Mehlenbacher SA (2011) Characterization of American hazelnut (*Corylus americana*) accessions and *Corylus americana* × *Corylus avellana* hybrids using microsatellite markers. Genet Resour Crop Evol. doi:10.1007/s10722-011-9743-0

Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis. User manual ver 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Geneva

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol and Evol 28:2731–2739

Thompson MM, Lagerstedt HB, Mehlenbacher SA (1996) Hazelnuts. In: Janick J, Moore JN (eds) Fruit Breeding: Nuts, Vol. 3. Wiley, New York, pp 125–184

Tzedakis PC, Lawson IT, Frogley MR, Hewitt GM (2002) Buffered tree population changes in quaternary refugium: evolutionary implication. Science 292:267–269

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Wright S (1978) Evolution and genetics of populations. Variability within and among natural populations. Vol. 4. University of Chicago Press, Chicago