

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

P. Boccacci · M. Aramini · N. Valentini · L. Bacchetta · M. Rovira · P. Drogoudi ·
A. P. Silva · A. Solar · F. Calizzano · V. Erdoğan · V. Cristofori · L. F. Ciarmiello ·
C. Contessa · J. J. Ferreira · F. P. Marra · R. Botta

Received: 3 November 2012 / Revised: 21 June 2013 / Accepted: 8 July 2013 / Published online: 30 August 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Hazelnut (*Corylus avellana* L.) is a traditional nut crop in southern Europe. Germplasm exploration conducted *on-farm* in five countries (Portugal, Spain, Italy, Slovenia,

and Greece) identified 77 landraces. The present work describes phenotypic variation in nut and husk traits and investigates genetic relationships using ten simple sequence repeat

Communicated by E. Dirlewanger

Electronic supplementary material The online version of this article (doi:10.1007/s11295-013-0651-7) contains supplementary material, which is available to authorized users.

P. Boccacci (✉)
Plant Virology Institute—National Research Council (IVV-CNR),
UOS of Grugliasco, Via Leonardo Da Vinci, 44-10095 Grugliasco,
Turin, Italy
e-mail: p.boccacci@ivv.cnr.it

P. Boccacci · N. Valentini · F. Calizzano · C. Contessa · R. Botta
Department of Agricultural, Forestry and Food Science, University
of Torino, Via Leonardo Da Vinci, 44-10095 Grugliasco, Turin, Italy

M. Aramini · L. Bacchetta
Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo
economico sostenibile (ENEA), Via Anguillarese, 301-00123 Santa
Maria di Galeria, Rome, Italy

M. Rovira
Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Centre
Mas de Bover, Ctra. Reus-El Morell, km 3.8, 43120, Constantí
Tarragona, Spain

P. Drogoudi
Pomology Institute, Hellenic Agricultural Organisation 'Demeter',
DG of Agricultural Research, NAGREF, 38 R.R. Station, 59
035 Naoussa, Greece

A. P. Silva
Centre for the Research and Technology of Agro-Environment and
Biological Sciences (CITAB), Universidade de Trás-os-Montes e
Alto Douro, Apartado 1013, 5001-801 Vila Real, Portugal

A. Solar
Biotehniška Fakulteta, Univerza v Ljubljani, Jamnikarjeva 101,
1000 Ljubljana, Slovenia

V. Erdoğan
Department of Horticulture, Faculty of Agriculture, Ankara
University, Dışkapı, 06110 Ankara, Turkey

V. Cristofori
Department of Agriculture, Forests, Nature and Energy, University of
Tuscia, Via S. Camillo de' Lellis snc, 01100 Viterbo, Italy

L. F. Ciarmiello
Consiglio per la Ricerca e Sperimentazione in Agricoltura
(CRA)—Istituto Sperimentale per la Frutticoltura, via Torrino,
3-81100 Caserta, Italy

J. J. Ferreira
Area de Cultivos Hortofrutícolas y Forestales, Servicio Regional de
Investigación y Desarrollo Agroalimentario (SERIDA), 33300,
Villaviciosa Asturias, Spain

F. P. Marra
DEMETRA, University of Palermo, Viale delle Scienze,
90128 Palermo, Italy

(SSR) markers among these landraces, 57 well-known references cultivars, and 19 wild accessions. Among the 77 landraces, 42 had unique fingerprints while 35 showed a SSR profile identical to a known cultivar. Among the 42 unique landraces, morphological observations revealed high phenotypic diversity, and some had characteristics appreciated by the market such as nut round and caliber. Analysis of genetic relationships and population structure allowed investigation of the origin and spread of the cultivated germplasm in southern Europe. Our results indicate the existence of three primary centers of diversity in the Mediterranean basin: northwestern Spain (Tarragona) and southern Italy (Campania) in the West and Black Sea (Turkey) in the East. Moreover, the data suggest the existence of secondary gene pools in the Iberian (Asturias) and Italian (Liguria and Latium) Peninsulas, where local varieties were recently domesticated from wild forms and/or from introduced ancient domesticated varieties.

Keywords Filbert · In situ conservation · Biodiversity · Simple sequence repeat (SSR) markers · Microsatellite · Domestication

Introduction

Corylus avellana L., the European hazelnut, is diploid ($2n = 2x = 22$), monoecious, dichogamous, and wind-pollinated and has sporophytic incompatibility that enforces cross-pollination. Its geographical distribution extends from Europe and North Africa to the Caucasus region and Asia Minor. It is the source of important cultivars in Europe and Turkey, which were selected over many centuries from local wild populations (Thompson et al. 1996). A few superior cultivars were spread beyond their area of origin by trade and human migration. In spite of the long cultivation history, little remains known about the origin and domestication. European hazelnut is one of the most important tree nut crops in terms of worldwide production. The Black Sea countries account for the majority of world production: Turkey (610,264 tons, average of 2009–2011), Azerbaijan (28,564 tons), and Georgia (20,567 tons). Other important producers are Italy (114,991 tons), the USA (35,079 tons), and Spain (16,988 tons) followed by Iran, China, France, and Greece (FAOstat 2011), with significant new plantings in Chile. About 90 % of the world crop is shelled and sold as kernels, while the remaining 10 % is sold in-shell for fresh consumption. The primary user of kernels, the food industry, requires cultivars that produce nuts with few defects and has precise requirements for morphological, chemical, and physical characteristics of the kernels.

In recent years, efforts to improve the efficiency and effectiveness of agro-biodiversity conservation have been conducted for most crop species as required by the Convention on

Biological Diversity (CBD 1992), the International Treaty on Plant Genetic Resources for Food and Agriculture (FAO 2001), and the Global Plant Conservation Strategy (CBD 2002). Agro-biodiversity includes plant genetic resources (PGR) for food and agriculture: (a) modern cultivars, breeding lines, and genetic stocks that are widely and actively conserved by plant breeders and gene banks and (b) obsolete cultivars, landraces (e.g., farmer populations of crop plants), ecotypes (e.g., natural plant populations), and wild and weedy relatives that still need to be actively conserved (Polegri and Negri 2010). PGR can be used as raw material required for genetic improvement, allowing a crop to adapt to unpredictable environmental changes and resistant to diseases, guaranteeing food security for future generations (FAO 2001). In recent decades, PGR have usually been conserved by ex situ methods. More recently, in situ conservation, sometimes referred to as “on-farm conservation,” has been proposed as a better conservation strategy in order to protecting diversity of cultivated and wild plant species in its natural habitat (Jarvis et al. 2000).

In hazelnut, about 400 clonal cultivars have been described and are maintained in different ex situ germplasm repositories (Thompson et al. 1996). A total of 510 accessions are conserved in 13 European hazelnut collection fields: four in Italy, three in Portugal, two in Spain, and one each in Slovenia, France, and Greece (Rovira et al. 2011). More than 700 *Corylus* accessions were preserved in the major world hazelnut collection located in Oregon (USA), while a collection containing 20 registered cultivars and more than 400 accessions collected from the Black Sea coast is in Turkey (Gürçan et al. 2010a). In contrast, the in situ conservation strategies proposed by Jarvis et al. (2000) have not yet been applied to hazelnut PGR, although a first on-farm exploration was conducted in northern Spain (Asturias) by Ferreira et al. (2010). Over three consecutive years (2008–2010), the EU AGRI GEN RES project SAFENUT (“Safeguard of almond and hazelnut genetic resources: from traditional uses to modern agro-industrial opportunities”) aimed to increase knowledge of genetic diversity in the European hazelnut. Objectives included description of cultivars from different ex situ European collections as well as the on-farm exploration, description, and in situ conservation of local endangered PGR. This characterization was carried out using different set of descriptors: morphological, biochemical, molecular, as well as ecological and cultural aspects (Bacchetta et al. 2011).

Identification of accessions and analysis of genetic diversity in collections (ex situ and in situ) are important in the management and utilization of PGR. Traditional methods to characterize and identify hazelnut accessions or cultivars are based on morphological and phenological descriptors (Biodiversity International 2008). In recent years, DNA markers have proven to be useful for accurately identifying cultivars due to their high discriminating power at a relatively low cost. In *C. avellana*, microsatellite or simple sequence

repeat (SSR) markers have been developed (Bassil et al. 2005a, b; Boccacci et al. 2005; Gürcan and Mehlenbacher 2010a, b; Gürcan et al. 2010b) and mapped (Mehlenbacher et al. 2006; Sathuvalli et al. 2011). Loci have been used to fingerprint accessions in collections, identify synonyms, determine parentage, and assess genetic relationships among cultivars (Boccacci et al. 2006, 2008; Gökirmak et al. 2009; Gürcan et al. 2010a). SSR markers have been also used to investigate the genetic diversity and structure of different populations (Boccacci and Botta 2010; Gökirmak et al. 2009; Gürcan et al. 2010a) or compare local cultivars and wild hazelnuts (Campa et al. 2011).

The present work reports the results of a hazelnut germplasm exploration conducted *on-farm* within the SAFENUT project in five southern European countries (Portugal, Spain, Italy, Slovenia, and Greece). The main aims were to characterize hazelnut landraces using morphological descriptors and SSR markers and to investigate their genetic relationships with wild forms and well-known cultivars. The information will be useful to identify landraces for in situ preservation, further evaluation in ex situ collections, and use in breeding programs. The origin and diffusion of the cultivated germplasm in southern Europe, particularly in the Italian Peninsula, will also be discussed

Materials and methods

Plant material

A total of 153 hazelnut accessions were analyzed in this study: (a) 77 landraces mostly surveyed *on-farm* during the SAFENUT project (2008–2010) (Table 1), (b) 57 reference cultivars from different European and Turkish collections (Supplementary Table 1), and (c) 19 wild hazelnuts sampled in the sites of Vejano (Latium, central Italy) and Benevento (Campania, South Italy), where wild populations are still present. The landraces were surveyed in the traditional areas of hazelnut cultivation in five southern European countries (Table 1). Among them, 5 were collected in northern Portugal, 10 in northern Spain (Asturias), 52 in six Italian regions [6 in Piedmont (northwestern Italy), 10 in Liguria (northwestern), 1 in Marche (central Italy), 12 in Latium (central Italy), 3 in Calabria (southern Italy), and 20 in Sicily], 5 from Slovenia, and 5 from northern Greece. Farmers were contacted explaining the reasons for the project and interviewed about the presence of old endangered cultivars on their farms. Information on agronomic and qualitative traits, as well as use, local names, tradition, and social context, were also collected. The set of reference cultivars included those well-known and grown in the above mentioned countries, together with the most important ones from Black Sea region. In consideration of the high number of landraces and cultivars

from Italy, some wild forms were included to investigate the origin and circulation of hazelnut in the Italian Peninsula.

Morphological observations

A total of 20–50 nuts were collected in situ from each surveyed landrace. Husks or involucre, nuts, and kernels were characterized using 14 qualitative standard descriptors (Table 2), following Thompson et al. (1978), the UPOV (1979), and Bioversity International (2008) guidelines.

DNA extraction and SSR analysis

Total genomic DNA was extracted from 0.25 g of young leaves or immature catkins using the modified procedure of Thomas et al. (1993). A total of ten SSR loci, selected by Boccacci and Botta (2010) for the SAFENUT project, were analyzed: CaT-B107, CaT-B501, CaT-B502, CaT-B503, CaT-B504, CaT-B505, CaT-B507, CaT-B508 (Boccacci et al. 2005), CaC-B020, and CaC-B028 (Bassil et al. 2005a). PCR amplifications were performed in a volume of 15 μ l containing 40 ng DNA, 0.5 U Taq-DNA polymerase (AmpliTaQ Gold, Applied Biosystems, Foster City, CA, USA), 1.5 μ l 10 \times PCR buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl), 2 mM MgCl₂, 200 μ M dNTPs, and 0.5 μ M of each primer. The PCR conditions were: a first denaturation step at 95 °C for 9 min, followed by 26 cycles of denaturation (30 s at 95 °C), annealing (45 s at 55 °C and 50 °C for CaT-B502), and extension (90 s at 72 °C). The final elongation step was carried out at 72 °C for 30 min. Amplification products were analyzed using an ABI-PRISM 3130 Genetic Analyzer capillary electrophoresis instrument (Applied Biosystems, Foster City, CA, USA). Results were processed with GeneMapper software (Applied Biosystems), and alleles were designated by their size in base pairs using a GeneScan-500 LIZ standard (Applied Biosystems).

Data analysis

Microsatellite data obtained at ten SSR loci for 153 hazelnut accessions were processed using the software Identity 4.0 (Wagner and Sefc 1999) to identify accessions with identical SSR profiles. When two or more accessions had identical SSR profiles, only one was retained for further analysis.

The genetic relationships among the different genotypes were investigated using two types of analysis. An unweighted pair-group method using arithmetic average (UPGMA) was used to construct and draw a dendrogram from the genetic similarity matrix using MEGA v. 5.05 (Tamura et al. 2011). Genetic distances (1,000 bootstraps) were computed as: $D = [1 - (\text{proportion of shared alleles})]$, using Microsat software (Minch 1997). A principal coordinate analysis (PCoA) was computed by GenAlEx 6.2 (Peakall and Smouse 2006).

Table 1 Locations of 77 landraces characterized in five southern European countries

Accession/country	Region	Zone/council	Town/locality
Portugal			
Cartuxeira/Tubulosa	Norte	Viseu	Moimenta da Beira
Downton	Norte	Viseu	Moimenta da Beira
Purpurea	Norte	Viseu	Moimenta da Beira
Quinta Vila Nova Do Rego	Norte	Viseu	Povolide
Raul	Norte	Penafiel	Entre-os-Rios
Spain			
Allande-3	Asturias	Allande	Pola de Allande
Barreiras-1	Asturias	Sta. Eulalia de Oscos	Barreiras
Las Cuevas-1	Asturias	Caso	Las Cuevas
Llanazares-2	Asturias	Aller	Casomera
Priero-1	Asturias	Caso	Prieres
Pumares-4	Asturias	Sta. Eulalia de Oscos	Pumares
Robriguedo-2	Asturias	Penamellera Baja	Robriguero
Rubiano-1	Asturias	Grado	Rubiano
San Pedro-1	Asturias	Grado	San Pedro de los Burros
Tuñon-3	Asturias	Santo Adriano	Tuñon
Italy			
Allungata	Latium	Viterbo	Caprarola
Bardina	Liguria	Genova	Mezzanego
Barrettona	Latium	Viterbo	Vico Matrino
Cappello del prete	Latium	Viterbo	Caprarola
Caraffara SIC 1	Sicily	Catania	Etna Natural Park
Caraffara SIC 10	Sicily	Catania	Etna Natural Park
Caraffara SIC 11	Sicily	Catania	Etna Natural Park
Caraffara SIC 14	Sicily	Catania	Etna Natural Park
Caraffara SIC 2	Sicily	Catania	Etna Natural Park
Caraffara SIC 3	Sicily	Catania	Etna Natural Park
Caraffara SIC 9	Sicily	Catania	Etna Natural Park
Ciasetta	Liguria	Genova	Mezzanego
Del Rosso	Liguria	Genova	Pian dei Cunei
Dell'Orto	Liguria	Genova	Mezzanego
Gianchetta	Liguria	Genova	Mezzanego
Itavex	Latium	Viterbo	Caprarola
Lunghera	Liguria	Genova	Mezzanego
Madonnella	Latium	Viterbo	Caprarola
Menoia	Liguria	Genova	Mezzanego
Minnulara/Minnolara	Sicily	Catania	Etna Natural Park
Nocchia rosa	Latium	Viterbo	Caprarola
Nocciola Ada	Latium	Viterbo	Caprarola
Nocciola Benedetta	Latium	Viterbo	Carbognano
Nocciola Centenaria	Latium	Viterbo	Ronciglione
Nocciola della Madonnella	Latium	Viterbo	Caprarola
Nocciola lunga	Latium	Viterbo	Ronciglione
Nocciola Meloni	Latium	Viterbo	Caprarola
Noscello	Liguria	Genova	Pian dei Cunei
San Vicino Vittori	Marche	Ascoli-Piceno	Castigliano
Seigretta	Liguria	Genova	Mezzanego
Selvaggiola agostara	Sicily	Catania	Etna Natural Park

Table 1 (continued)

Accession/country	Region	Zone/council	Town/locality
Selvaggiola lunga	Sicily	Catania	Etna Natural Park
Selvaggiola riccia	Sicily	Catania	Etna Natural Park
Selvaggiola SIC 16	Sicily	Catania	Etna Natural Park
Selvaggiola SIC 17	Sicily	Catania	Etna Natural Park
Selvaggiola SIC 4	Sicily	Catania	Etna Natural Park
Selvaggiola SIC 6	Sicily	Catania	Etna Natural Park
Selvaggiola SIC 7	Sicily	Catania	Etna Natural Park
Selvaggiola SIC 13	Sicily	Catania	Etna Natural Park
Selvaggiola tardiva SIC 8	Sicily	Catania	Etna Natural Park
Selvaggiola tardiva SIC12	Sicily	Catania	Etna Natural Park
Tapparona	Liguria	Genova	Mezzanego
Tonda di Biglini Bi/01	Piedmont	Cuneo	Biglini
Tonda di Biglini Bi/02	Piedmont	Cuneo	Biglini
Tonda di Biglini LeqBer/01	Piedmont	Cuneo	Lequio Berria
Tonda di Biglini LeqBer/02	Piedmont	Cuneo	Lequio Berria
Tonda di Biglini Ver/05	Piedmont	Cuneo	Verduno
Tonda di Biglini Ver/06	Piedmont	Cuneo	Verduno
Tonda di Calabria Ca/01	Calabria	Catanzaro	Torre di Ruggiero
Tonda di Calabria Ca/02	Calabria	Catanzaro	Gagliato
Tonda di Calabria Ca/03	Calabria	Vibo Valentia	Jonadi
Trichette	Sicily	Catania	Etna Natural Park
Slovenia			
CV/1	Lower Styria	Šentjur	Rifnik
CV/2	Lower Styria	Šentjur	Rifnik
T/0	Lower Styria	Slovenske Konjice	Novo Tepanje
T/10	Lower Styria	Slovenske Konjice	Novo Tepanje
T/16	Lower Styria	Slovenske Konjice	Novo Tepanje
Greece			
Argiroupoli	East Macedonia and Thrace	Drama	Drama
Philio	East Macedonia and Thrace	Drama	Drama
Patem large	East Macedonia and Thrace	Drama	Drama
Patem small	East Macedonia and Thrace	Drama	Drama
Polykarpos	East Macedonia and Thrace	Drama	Drama

SSR profiles indicated that 42 are unique genotypes

The program STRUCTURE v. 2.3.3 (Pritchard et al. 2000), a model-based Bayesian clustering method, was used to infer population structure and assign individuals to sub-populations. STRUCTURE was run five independent times for each K value ranging from 1 to 10. The admixture model was applied and allele frequencies were assumed to be correlated. A burn-in period of 100,000 generations and 200,000 Markov chain Monte Carlo replications were used. All individuals were treated as having known origin and assigned to one of eight geographical groups. Among them, accessions from Italy were assigned to different geographical groups in order to investigate the origin and spread of cultivated germplasm in

the Italian Peninsula. Landraces and cultivars were assigned to one of six groups: the Iberian Peninsula (Spain and Portugal, 18 accessions), Sicily (22), southern Italy (7), central Italy (12), northwestern Italy (13), and Balkans–Black Sea (Slovenia, Greece, and Turkey, 27). Wild individuals were from Latium (9) and Campania (10). The statistic ΔK (Evanno et al. 2005) was calculated by STRUCTURE HARVESTER software (Earl and vonHoldt 2012) and used to select the optimal K value.

Genetic diversity and differentiation among the eight geographical populations was investigated. Popgene software (Yeh et al. 1997) was used to calculate observed (N_a) and

Table 2 Proportion of phenotypic classes of morphological descriptors of hazelnut fruits collected from landraces shown to have unique SSR genotype

Descriptors	<i>N</i>	Phenotypic classes (number of samples/proportion %)						<i>H</i>
Predominant nut number per cluster	39	1 (0/0)	1–2 (12/30.8)	2–3 (18/46.2)	3–4 (3/7.7)	>4 (6/15.4)		1.20
Involucre length compared to nut length	38	Shorter (8/21.1)	Equal (12/31.6)	Longer (18/47.4)				1.05
Nut size ^a	42	Very large (1/2.4)	Large (9/21.4)	Medium (15/35.7)	Small (17/40.5)			1.15
Nut shape	42	Oblate (3/7.1)	Globular (14/33.3)	Conical (0/0)	Ovoid (7/16.7)	Short cylindrical (7/16.7)	Long cylindrical (11/26.2)	1.50
Nut shell color	42	Greenish yellow (2/4.8)	Light brown (27/64.3)	Brown (12/28.6)	Dark brown (1/2.4)			0.88
Nut shell striping	42	Absent (0/0)	Few (14/33.3)	Medium (20/47.6)	Many (8/19.0)			1.04
Size of pistil scar	34	Small (15/44.1)	Medium (15/44.1)	Large (4/11.8)				0.97
Presence of double kernels	42	Absent (39/92.9)	Present (3/7.1)					0.26
Kernel size ^b	42	Very large (0/0)	Large (6/14.3)	Medium (19/45.2)	Small (17/40.5)			1.00
Kernel shape	42	Oblate (2/4.8)	Globular (10/23.8)	Conical (2/4.8)	Ovoid (12/28.6)	Short cylindrical (4/9.5)	Long cylindrical (12/28.6)	1.57
Appearance of skin	42	Smooth (4/9.5)	Slightly corky (24/57.1)	Medium corky (11/26.2)	Strongly corky (3/7.1)			1.08
Size of internal cavity of kernel	32	Absent (5/15.6)	Small (17/53.1)	Medium (6/18.8)	Large (4/12.5)			1.20
Percentage of kernel by weight ^c	41	Very low (12/29.3)	Low (12/29.3)	Medium (13/31.7)	High (4/9.8)	Very high (0/0)		1.31
Percentage of kernel caliber > 12 mm ^d	41	Very low (13/31.7)	Low (3/7.3)	Medium (7/17.1)	High (18/43.9)			1.22

N number of landraces characterized, *H* Shannon–Weaver diversity index

^a Nut size: 1=very large (>4 g), 2=large (3.1–4 g), 3=medium (2.1–3 g), 4=small (<2.0 g)

^b Kernel size: 1=very large (>1.65 g), 2=large (1.26–1.65 g), 3=medium (0.86–1.25 g), 4=small (<0.85 g)

^c Percentage of kernel by weight: 1=very low (<40 %), 2=low (40.1–45 %), 3=medium (45.1–50 %), 4=high (50.1–55 %), 5=very high (>55 %)

^d Percentage of kernel caliber >12 mm: 1=very low (0–25 %), 2=low (25.1–50 %), 3=medium (50.1–75 %), 4=high (75.1–100 %)

effective (N_e) number of alleles, observed (H_o) and expected heterozygosity (H_e), Nei's (1978) coefficient of genetic identity (G_i) and genetic distance (G_d), and gene flow (N_m) (Slatkin and Barton 1989). The fixation index (F_{st}) was estimated according to Weir and Cockerham (1984) using the program F-STAT (Goudet et al. 1995). The significance levels of the F_{st} values were determined after 560 permutations.

The Shannon–Weaver index was calculated for phenotypic diversity in each trait in the 42 landraces. The diversity index was calculated as $H = -\sum p_i \ln p_i$, where p_i is the frequency of the phenotypic class i for each trait, as reported in Table 2 (Shannon and Weaver 1949). A principal component analysis (PCA) using the 14 morphological descriptors for 42 landraces and 11 standard cultivars ('Casina,' 'Barcelona,' and 'Negret' from Spain; 'Nocchione,' 'Tonda Gentile Romana,' 'Tonda di Giffoni,' and 'Tonda Gentile delle Langhe' from Italy; 'Tombul' from Turkey; 'Cosford' from England; 'Istrska dolgoplodna' and 'Istrska okrogloplodna' from Slovenia) was performed using PAST v. 2.12 software (Hammer et al. 2001).

Results

Sets of duplicates

Microsatellite analysis identified 42 unique genotypes and ten sets of duplicate accessions among the 77 investigated landraces. Accessions listed as duplicates were similar for nut and husk morphology. Thirty-five landraces showed SSR profiles identical to a reference cultivar (Supplementary Table 1).

Among the landraces surveyed in the Iberian Peninsula, three sets of duplicates were detected. The first set consisted of three accessions from Portugal ('Cartuxeria/Tubulosa,' 'Dawton,' and 'Purpurea') whose SSR profiles were identical to 'Fructo rubro' (syn. 'Pellicule rouge'). All had small, long, thin-shelled nuts and long tubular husks. The second set was the pair 'Raul' and the Turkish cultivar 'Karidaty' (syns. 'Imperiale de Trebizonde,' 'Kargalak'). The third set consisted of six accessions from northern Spain which were identical to 'Casina,' the most common cultivar in this area.

Table 3 Morphological and technological parameters of 42 landraces with unique SSR profiles and standard cultivars ‘Tonda Gentile delle Langhe’ and ‘Negret’

Identification	Code in the PCA	Nut size	Nut shape	Presence of double kernels	Kernel size	Kernel shape	Appearance of skin	Percentage of kernel by weight (%)	Percentage of kernel caliber > 12 mm (%)
Allande-3	All_3	Large	Short cylindrical	Absent	Small	Ovoid	Medium corky	38	70
Allungata	Allung	Medium-small	Long cylindrical	Absent	Medium	Long cylindrical	Slightly corky	47	100
Bardina	Bard	Small	Ovoid	Absent	Small	Ovoid	Medium corky	54	14
Barrettina	Barrett	Medium	Globular	Absent	Medium	Globular	Slightly corky	50	100
Cappello del Prete	Capp_Prete	Medium	Globular	Absent	Medium	Globular	Slightly corky	44	100
Ciasetta	Cias	Small	Long cylindrical	Absent	Small	Long cylindrical	Medium corky	49	0
Dall’Orto	D_Orto	Small	Long cylindrical	Present	Small	Long cylindrical	Slightly corky	45	0
Del Rosso	D_Rosso	Small	Ovoid	Absent	Small	Ovoid	Not corky	45	4
Philio	Philio	Large	Oblate	Absent	Medium	Ovoid	Medium corky	40	100
Gianchetta	Gianch	Small	Long cylindrical	Absent	Small	Long cylindrical	Slightly corky	53	0
Itavex	Itavex	Small	Globular	Absent	Medium	Globular	Slightly corky	44	100
Las Cuevas-1	L_Cuev1	Medium	Globular	Absent	Small	Globular	Slightly corky	45	20
Lunghera	Lungh	Small	Long cylindrical	Absent	Small	Long cylindrical	Slightly corky	45	0
Madonnella	Madon	Small	Globular	Absent	Small	Globular	Slightly corky	45	90
Menolia	Menolia	Small	Ovoid	Absent	Small	Ovoid	Slightly corky	47	5
Minnulara	Minnu	Large	Short cylindrical	Absent	Medium	Short cylindrical	Slightly corky	34	40
Nocciola Ada	Nocc_Ada	Medium	Long cylindrical	Absent	Medium	Long cylindrical	Slightly corky	50	70
Nocciola Benedetta	Nocc_Ben	Medium	Ovoid	Absent	Medium	Ovoid	Slightly corky	40	100
Nocciola Centenaria	Nocc_Cent	Medium	Long cylindrical	Absent	Medium	Long cylindrical	Slightly corky	44	90
Nocciola lunga	Nocc_lung	Medium	Long cylindrical	Absent	Medium	Long cylindrical	Slightly corky	50	100
Noscello	Nosc	Medium	Ovoid	Absent	Medium	Ovoid	Medium corky	49	93
Prieres-1	Pri_1	Small	Globular	Absent	Small	Ovoid	Not corky	-	-
Quinta Vila Nova do Rego	QVNdR	Small	Globular	Absent	Small	Globular	Slightly corky	47	48
Robriguedo-2	Robr_2	Medium	Globular	Absent	Medium	Globular	Medium corky	44	90
San Vicino Vittori	San_Vic_Vitt	Very large	Short cylindrical	Absent	Large	Short cylindrical	Medium corky	49	100
Seigretta	Seigr	Small	Ovoid	Present	Small	Ovoid	Medium corky	51	0
Selvaggiola agostara	Selv_ago	Large	Globular	Absent	Medium	Conical	Strongly corky	30	70
Selvaggiola lunga	Selv_lunga	Small	Long cylindrical	Present	Small	Long cylindrical	Medium corky	45	0
Selvaggiola riccia	Selv_ric	Medium	Globular	Absent	Medium	Globular	Slightly corky	42	55
Selvaggiola SIC13	Selv_SIC13	Large	Globular	Absent	Large	Ovoid	Slightly corky	40	100
Selvaggiola SIC16	Selv_SIC16	Small	Globular	Absent	Small	Globular	Slightly corky	39	45
Selvaggiola SIC17	Selv_SIC17	Medium	Short cylindrical	Absent	Medium	Ovoid	Strongly corky	34	100
Selvaggiola SIC4	Selv_SIC4	Medium	Short cylindrical	Absent	Large	Long cylindrical	Slightly corky	40	60
Selvaggiola SIC6	Selv_SIC6	Large	Globular	Absent	Medium	Conical	Strongly corky	36	70

Table 3 (continued)

Identification	Code in the PCA	Nut size	Nut shape	Presence of double kernels	Kernel size	Kernel shape	Appearance of skin	Percentage of kernel by weight (%)	Percentage of kernel caliber > 12 mm (%)
Selvaggiola SIC7	Selv_SIC7	Medium	Long cylindrical	Absent	Medium	Long cylindrical	Medium corky	39	55
Selvaggiola tardiva SIC12	Selv_SIC12	Small	Short cylindrical	Absent	Small	Short cylindrical	Slightly corky	44	80
Selvaggiola tardiva SIC8	Selv_SIC8	Large	Oblate	Absent	Medium	Oblate	Medium corky	39	100
T/0	T/0	Small	Oblate	Absent	Small	Oblate	Slightly corky	47	17
T/10	T/10	Small	Ovoid	Absent	Small	Ovoid	Not corky	46	8
T/16	T/16	Large	Short cylindrical	Absent	Large	Short cylindrical	Not corky	43	92
Tapparona	Tapp	Medium	Long cylindrical	Absent	Medium	Long cylindrical	Slightly corky	51	3
Tricchete	Trich	Large	Globular	Absent	Large	Globular	Slightly corky	48	85
TGL	TGL	Medium	Globular	Present	Medium	Globular	Medium corky	48	90
Negret	Negret	Medium	Ovoid	Present	Medium	Ovoid	Not corky	49	65

Three sets of accessions with the same SSR profile were found among the accessions collected in Italy. The first set included six accessions of ‘Tonda di Biglini’ (Piedmont) and ‘Tonda Gentile delle Langhe’ (‘TGL’). The latter cultivar represents over 90 % of the orchards in the Piedmont region and is well-known for kernel quality. However, there were phenological and carpological differences between ‘Tonda di Biglini’ and ‘TGL.’ ‘Tonda di Biglini’ nuts matured 10–15 days earlier, had thicker shells, a lower percentage of kernel by weight, and a higher frequency of double kernels (data not showed). The second set was ‘Meloni’ and ‘Nocciola della Madonnella’ from Latium, which were identical to ‘Tonda Gentile Romana’ (‘TGR’), which represents ~85 % of the local nut production. Nut maturity in ‘Meloni’ was about 15 days earlier than ‘TGR.’ The third set consisted of ‘Nocchia rosa’ (Latium), three accessions of ‘Tonda Calabrese’ (Calabria), and six accessions of ‘Caraffara’ (Sicily), all of which were identical to ‘Nocchione.’ All had round-oblate nuts of medium size in short husks. ‘Nocchione’ is the main pollinizer of ‘TGR’ in the Latium region. It is also the most widely grown cultivar in Sicily, where is known under different names: ‘Nostrale,’ ‘Comune,’ or ‘Mansa’ (Catania and Messina provinces); ‘Racinante’ (Enna province); and ‘Santa Maria del Gesù’ (Palermo province). Alberghina (1982) attributed the small morphological differences observed among the above-mentioned cultivars to environmental factors and renamed the group ‘Siciliana.’ Recently, Boccacci et al. (2006) and Gökirmak et al. (2009) confirmed their identical profiles at 24 and 21 SSR loci, respectively.

Accessions ‘CV/1’ and ‘CV/2’ from Slovenia were identical to ‘Barcelona’ (syn. ‘Castanyera’). ‘Barcelona’ is commercially important in the USA and in France where it is known as ‘Fertile de Coutard.’ It is of minor importance in Spain where it is and known as ‘Castanyera’ in the northeast (Tarragona province) and ‘Grande’ in the north. In Portugal, it is called ‘Grada de Viseu’ (Mehlenbacher and Miller 1989).

Among the Black Sea cultivars, three sets of synonyms were noted. The first set was the pair ‘Patem small’ from Greece and ‘Fructo rubro.’ The second set was ‘Argiroupoli’ and ‘Patem large’ from Greece and the cultivar ‘Yassi Badem’ from Turkey. Kernels of ‘Yassi Badem’ resemble almonds in size and shape and are consumed fresh, but are not suitable for processing. Finally, the third set was the pair ‘Polykarpos’ and ‘Tombul Ghiaghli’ from Greece; the latter is commonly cultivated there.

Morphological characterization

Morphological observations revealed high phenotypic diversity among the 42 unique landrace genotypes (Table 2). The *H* index calculated for each of the 14 morphological descriptors averaged 1.1, ranging from 0.26 (‘presence of double kernels’) to 1.57 (‘kernel shape’); the highest values were for

‘nut shape’ (1.50) and ‘kernel shape’(1.57). The predominant number of nuts per cluster was 2–3 (46.2 %) and 1–2 (30.8 %). The majority of the landraces had an involucre longer than the nut (47.4 %). Of the nut characters, the small (40.5 %) and medium (37.5 %) sizes were most common. Nut shape was highly variable, but the globular (33.3 %) and long cylindrical (26.2 %) shapes were the most common. The majority of nuts had a light brown shell color (64.3 %), few (33.3 %) or medium stripes (47.6 %), and small (44.1 %) or medium (44.1 %) size of pistil scar. Among kernel descriptors, almost all accessions (92.9 %) had no double kernels. The majority had medium (45.2 %) or small (40.5 %) size kernels, and the most common shapes were ovoid, long cylindrical, and globular (28.6, 28.6, and 23.8 %, respectively). The appearance of the skin (pellicle) was slightly corky (57.1 %), and the size of the internal cavity was small (53.1 %). Concerning the ‘percentage of kernel by weight,’ 31.7 % of landraces showed medium values (45.1–50.0 %), while 58.6 % had values less than 45.0 %. Finally, 43.9 % of them had a high ‘percentage of kernel caliber >12 mm’ (75.1–100 %). These descriptors can be used to identify accessions suitable for the kernel or in-shell markets. The main morphological and technological traits are reported for the 42 landraces in Table 3.

In the PCA based on 14 morphological descriptors for 42 landraces and 11 reference cultivars, the first two components (PC1 and PC2) explained 38.7 % of the total variation. PC1 accounted for 25.1 % and was positively correlated with nut and kernel size. PC2 accounted for an additional 13.6 % and was mostly associated with nut and kernel shape. The PCA scatter-plot split the samples into three main groups (Fig. 1). Among Italian landraces, the northwestern accessions (Liguria) were separated from those from central (Latium) and southern (Sicily) Italy. Ligurian landraces, with the exception of ‘Noscello,’ were grouped on the right side of the scatter-plot with ‘Casina,’ ‘Istrska dolgoplodna,’ ‘Negret,’ and ‘Tombul.’ The accessions collected in Sicily and Latium, except ‘Selvaggiola lunga,’ clustered in two adjacent groups. The group in the upper left contained: (a) ‘Barrettona,’ ‘Cappello del Prete,’ ‘Itavex,’ and ‘Madonnella’ from Latium; (b) ‘Selvaggiola riccia,’ ‘Selvaggiola SIC6,’ SIC13, and SIC16, ‘Selvaggiola agostara,’ ‘Selvaggiola tardiva SIC8,’ and ‘Trichette’ from Sicily; and (c) the standard cultivars ‘Barcelona,’ ‘Nocchione,’ ‘TGR,’ ‘Tonda di Giffoni,’ and ‘TGL.’ The group in the lower part of the graph included: (a) ‘Allungata,’ ‘Nocciola Ada,’ ‘Nocciola Benedetta,’ ‘Nocciola centenaria,’ and ‘Nocciola lunga’ from Latium and ‘San Vicino Vittori’ from Marche; (b) ‘Minnulara,’ ‘Selvaggiola tardiva SIC12,’ and ‘Selvaggiola’ SIC4, SIC7,

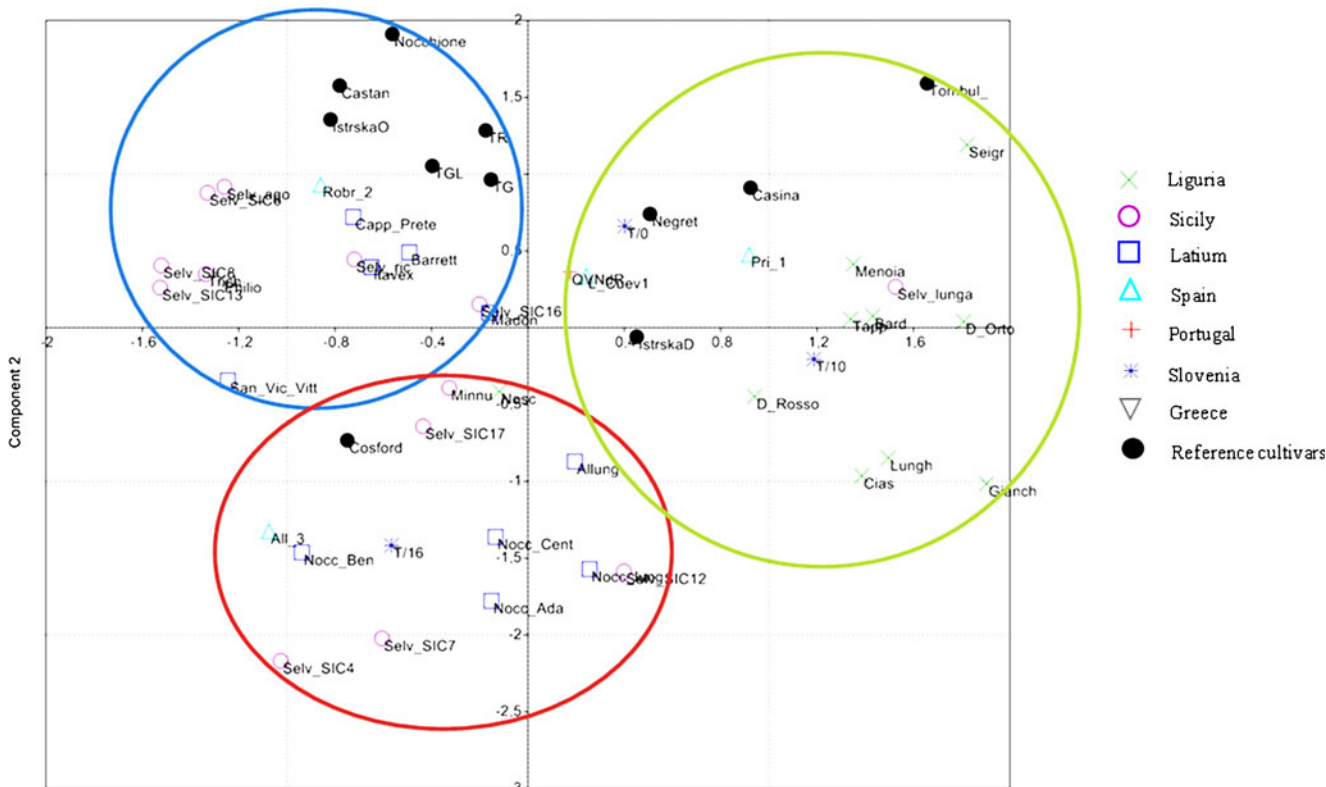


Fig. 1 PCA two-dimensional scatter plot based on the first two principal components (PC1 and PC2) generated for 42 landraces and 11 standard cultivars based on 14 morphological traits

Fig. 2 UPGMA dendrogram based on SSR analysis of 42 unique landrace genotypes (*LR*), 57 cultivars (*CV*), and 19 wild individuals (*W*)

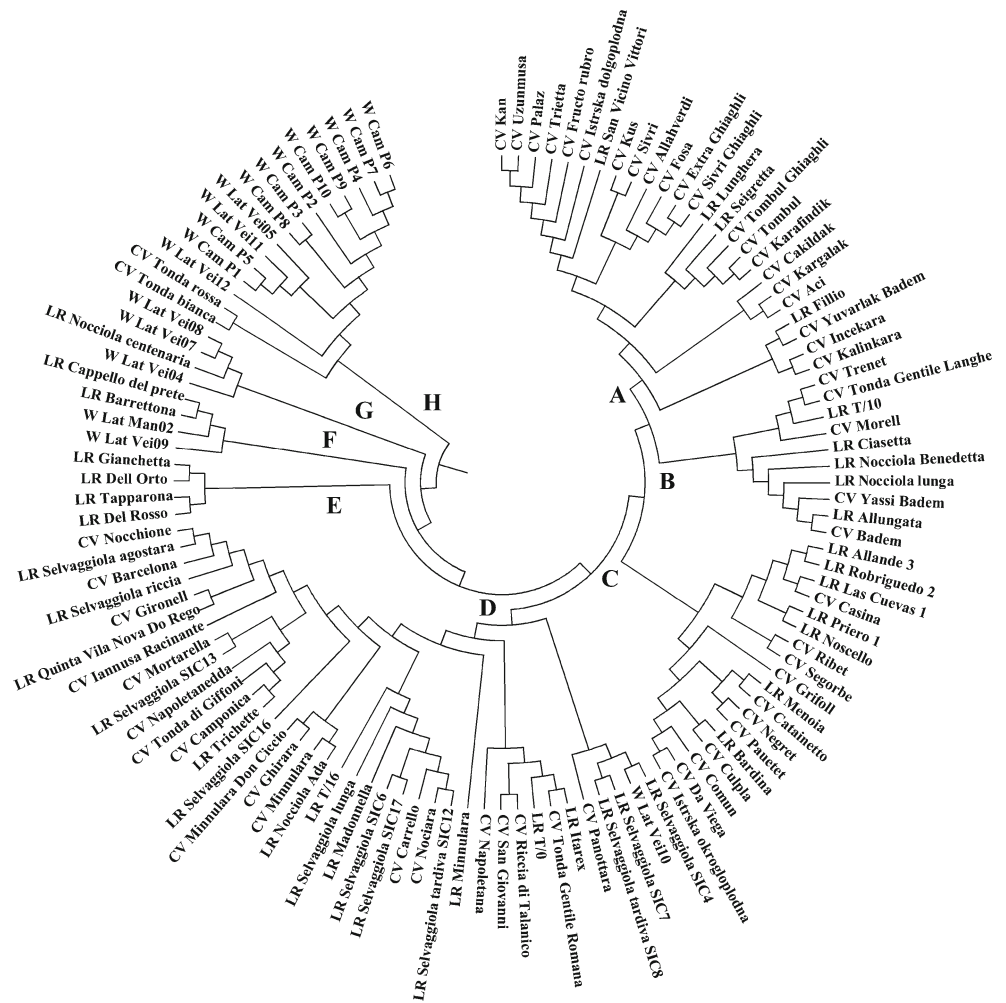
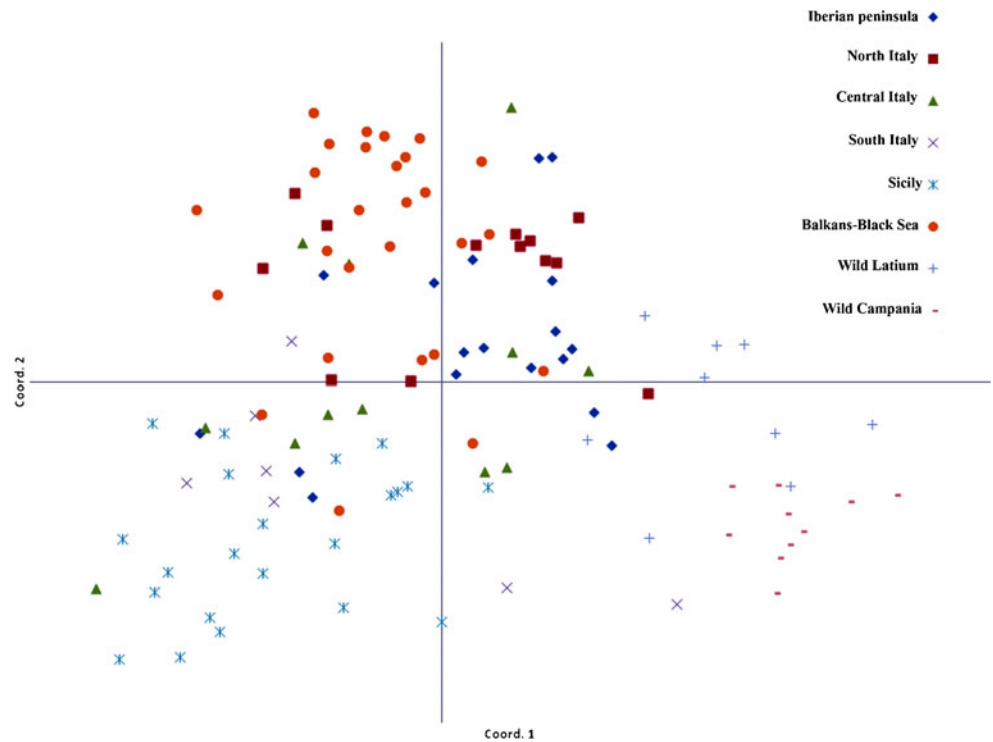


Fig. 3 Two-dimensional plot obtained from PCoA for 118 hazelnut genotypes classified in eight geographical groups and analyzed at ten SSR loci



and SIC17 from Sicily; (c) ‘Cosford.’ The PCA did not separate into distinct groups the landraces from the Iberian Peninsula, Slovenia, and Greece.

Genetic relationships and population structure analysis

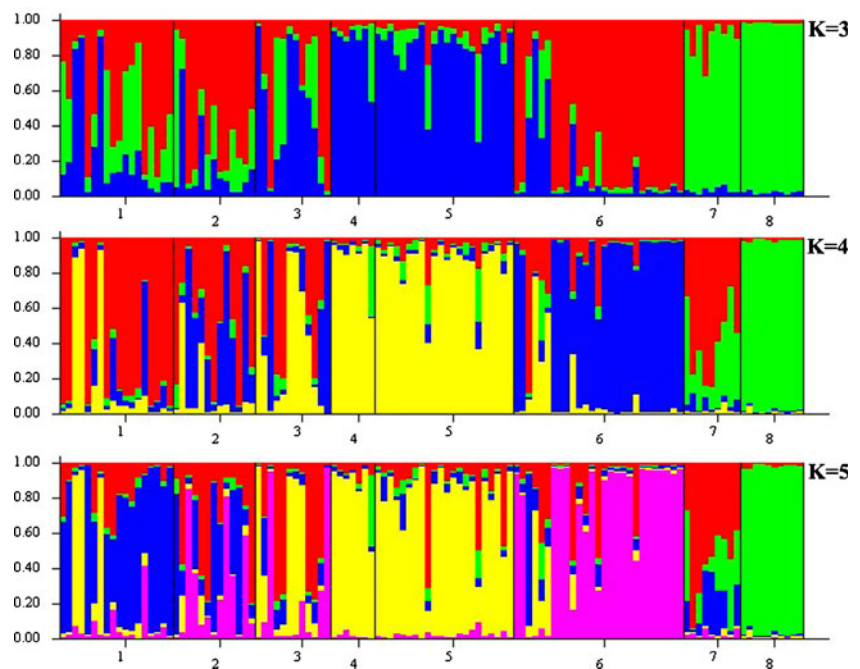
A dendrogram was constructed depicting the genetic relationships among the 42 unique landrace genotypes, 57 cultivars, and 19 wild individuals. Accessions were grouped into eight clusters (Fig. 2). Group A included cultivars from the Black Sea region (Turkey and Greece) as well as three accessions surveyed in Italy (‘San Vicino Vittori’ from Marche, ‘Lunghera’ and ‘Seigretta’ from Liguria) and one surveyed in Greece (‘Philio’). Group B was a mixture of cultivars and landraces from different geographical areas. ‘T/10’ surveyed in Slovenia appeared adjacent to the cultivars ‘TGL’ (Italy), ‘Trenet,’ and ‘Morell’ (Spain), while the landraces ‘Ciasetta’ from Liguria and ‘Nocciola Benedetta,’ ‘Nocciola lunga,’ and ‘Allungata’ from Latium were adjacent to the Turkish cultivars ‘Yassi Badem’ and ‘Yuvarlak Badem.’ Group C contained cultivars from the Iberian Peninsula and landraces surveyed in Asturias (‘Allande 3,’ ‘Robriguedo 2,’ ‘Las Cuevas 1,’ and ‘Priero 1’) and Liguria (‘Noscello,’ ‘Menoia,’ and ‘Bardina’). The landraces from northern Spain constituted a sub-group with ‘Casina’ and ‘Noscello.’ Most Italian cultivars and landraces were placed in the main cluster D, which includes germplasm from central Italy (Latium), south Italy (Campania), and Sicily. In this group were also placed the accessions ‘T/0’ and ‘T/16’ from Slovenia and ‘Quinta Vila Nova Do Rego’ from Portugal.

Group E was composed of only four landraces surveyed in Liguria (northwestern Italy). Finally, the last three clusters (F, G, and H) were comprised of wild genotypes and a few cultivars. Clusters F and G included both wild and landraces from Latium, while almost all wild individuals from Latium and Campania were included in the larger group H with ‘Tonda rossa’ and ‘Tonda bianca.’ These two cultivars are grown only in Avellino province (Campania, South Italy), are distinct from the other cultivars in Campania, and are morphologically similar.

In the PCoA, the first two PCs explained 48.7 % of the total variation. The first coordinate explained 26.1 % of the variation and the second coordinate an additional 22.6 %. The projection of 118 hazelnut accessions on a two-dimensional plane defined by the first two PCs (Fig. 3) showed a tendency to separate the cultivated accessions from the wild genotypes. Considering the geographical origin of the cultivars and landraces analyzed, the scatter-plot showed a tendency of the central–southern Italian accessions to cluster together in the lower half of the graph. Accessions from the Black Sea were preferentially placed in the upper left and those from the Iberian Peninsula in the upper right. Among the northern Italian accessions, three (‘Lunghera,’ ‘Seigretta,’ and ‘Trietta’) clustered with those from the Black Sea, seven (‘Bardina,’ ‘Del Rosso,’ ‘Dell’Orto,’ ‘Gianchetta,’ ‘Menoia,’ ‘Noscello,’ and ‘Tapparona’) clustered with the Iberian accessions, and three (‘Catainetto,’ ‘Ciasetta,’ and ‘TGL’) were placed in an intermediate position along the X-axis of the graph.

The 118 hazelnut genotypes were further evaluated for population stratification using the STRUCTURE software. SSR

Fig. 4 Hierarchical organization of genetic relatedness of 118 unique hazelnut genotypes based on ten SSR markers and analyzed by the STRUCTURE program, with three, four, and five populations ($K=3$, $K=4$, and $K=5$). Legend geographical groups: 1 Iberian Peninsula, 2 northwestern Italy, 3 central Italy, 4 southern Italy, 5 Sicily, 6 Balkans–Black Sea, 7 wild Latium, 8 wild Campania



data were analyzed increasing the number of subpopulations (*K*) from 1 to 10. The estimation of ΔK revealed the highest value for $K=3$ ($\Delta K=48.1$), indicating the existence of three groups, composed mainly of Turkish, wild, and central–southern Italian accessions, respectively (Fig. 4). Several genotypes were not clearly placed in separate groups, such as those from Spain or Liguria that clustered both with the Turkish and wild accessions. At $K=4$ ($\Delta K=21.2$) and $K=5$ ($\Delta K=12.0$), these three initially identified groups remained almost constant, whereas several Spanish accessions showed the tendency to constitute a separate group, Ligurian accessions were placed both with the Turkish and Spanish accessions, and some cultivated forms collected in Latium showed introgression with the local wild germplasm (Fig. 4). Comparing these results with the UPGMA dendrogram and the PCA scatterplot, there was general agreement about the population subdivisions and the genetic relationships among genotypes.

Differentiation among geographical gene pools

On the basis of their geographical area of origin, the 118 unique genotypes were divided in eight gene pools (Table 4 and 5). The observed (N_a) and effective (N_e) number of alleles and the observed (H_o) and expected heterozygosity (H_e) were calculated to evaluate the level of genetic diversity within each gene pool (Table 4). N_a and N_e ranged from 4.2 to 7.8 (average 6.3) and from 3.1 to 4.7 (average 3.8), respectively. H_o (average 0.81) was generally higher than H_e (average 0.71) in each group, with the exception of the wild individuals from Latium. The level of genetic diversity observed was high, a consequence of the self-incompatibility mating system of *C. avellana* and wind pollination, and similar to that found by other authors (Boccacci and Botta 2010; Boccacci et al. 2006, 2008; Gökirmak et al. 2009; Gürcan et al. 2010a; Campa et al. 2011).

Table 4 Genetic diversity for hazelnut accessions classified in eight geographical groups

Gene pools	N_i	N_a	N_e	H_o	H_e
Iberian Peninsula	18	6.5	3.9	0.87	0.73
Northwestern Italy	13	7.1	4.1	0.78	0.72
Central Italy	12	6.6	4.2	0.89	0.75
Southern Italy	7	4.2	3.1	0.86	0.66
Sicily	22	7.2	3.5	0.75	0.67
Balkans–Black Sea	27	7.8	3.7	0.83	0.71
Wild Latium	9	6.4	4.7	0.73	0.76
Wild Campania	10	4.8	3.2	0.74	0.67

N_i number of individuals, N_a observed number of alleles, N_e effective number of alleles, H_o observed heterozygosity, H_e expected heterozygosity

Table 5 Genetic identity (G_i), genetic distances (G_d), gene flow (N_m), and genetic differentiation (F_{st}) among and between hazelnut gene pools analyzed with SSR markers

Gene pools	Iberian Peninsula	NW Italy	Central Italy	Southern Italy	Sicily	Balkans–Black Sea	Wild Latium	Wild Campania
G_i	–	0.918	0.857	0.735	0.747	0.812	0.674	0.475
Iberian Peninsula	–	0.086	0.0853	0.687	0.684	0.880	0.601	0.452
Northwestern Italy	0.086	–	0.159	0.827	0.866	0.891	0.614	0.445
Central Italy	0.154	0.159	–	0.827	0.866	0.891	0.614	0.445
Southern Italy	0.308	0.376	0.190	–	0.824	0.667	0.428	0.413
Sicily	0.292	0.380	0.144	0.194	–	0.732	0.490	0.436
Balkans–Black Sea	0.209	0.127	0.115	0.405	0.313	–	0.557	0.400
Wild Latium	0.395	0.509	0.488	0.848	0.713	0.586	–	0.694
Wild Campania	0.745	0.795	0.811	0.885	0.830	0.916	0.366	–
G_d								
N_m	–	16.07	7.19	2.80	2.61	4.03	2.99	1.33
Iberian Peninsula	–	–	7.90	2.43	2.09	7.39	2.54	1.30
Northwestern Italy	0.015	–	–	5.23	5.63	8.59	2.91	1.36
Central Italy	0.034	0.031	–	–	3.87	2.01	1.46	1.05
Southern Italy	0.082	0.093	0.046	–	–	2.32	1.42	1.04
Sicily	0.087	0.107	0.043	0.061	–	–	1.85	1.07
Balkans–Black Sea	0.058	0.033	0.028	0.111	0.098	–	–	–
Wild Latium	0.077	0.090	0.079	0.146	0.150	0.119	–	3.12
Wild Campania	0.159	0.162	0.155	0.192	0.194	0.189	0.074	–
F_{st}								

Genetic identity (G_i), genetic distance (G_d), fixation index (F_{st}), and gene flow (N_m) were calculated to investigate genetic differentiation among gene pools (Table 5). G_i was the highest among gene pools comprised of cultivated accessions, ranging from 0.667 to 0.918. On the contrary, G_i values were lower between wild and cultivated groups, ranging from 0.400 to 0.694. Correspondingly, a higher G_d was found between cultivated and wild groups, ranging from 0.086 (Iberian Peninsula vs. northwestern Italy) to 0.916 (wild Campania vs. Balkans–Black Sea). All pairwise comparisons yielded significant differentiation values, ranging from 0.015 (Iberian Peninsula vs. northwestern Italy) to 0.194 (wild Campania vs. Sicily) with $P_{(F_{st} \text{ not } > 0)} < 0.05$ for each F_{st} value, and an equal distribution of N_m values between gene pools was observed.

Discussion

Mislabeled and the existence of synonyms and homonyms are important challenges for germplasm conservation. In the past decade, SSR markers have become very valuable tools in the management of ex situ hazelnut collections. In their study, Boccacci et al. (2006) reported six sets of synonyms among 78 accessions from European collections, Gökirmak et al. (2009) found 72 duplicates among 270 accessions from USDA-ARS-NCGR and OSU germplasm repositories, and 6 Turkish accessions conserved in the US collection fields were found to be synonyms of cultivars from the HRI collection by Gürçan et al. (2010a).

Among the 77 landraces surveyed in five southern European countries (Portugal, Spain, Italy, Slovenia, and Greece), SSR profiles indicated that 42 were unique genotypes, while 35 accessions appeared to be synonyms. A total of ten sets of duplicates were found between landraces and some reference cultivars. Among them, landraces from Portugal, Slovenia, and Greece had profiles identical to those of six foreign cultivars: ‘Barcelona’ (syn. ‘Castanyera’) from Spain, ‘Fructo rubro’ from the Balkans, and ‘Karidaty’ (syns. ‘Imperiale de Trebizonde,’ ‘Kargalak’), ‘Yassi Badem,’ ‘Palaz,’ and ‘Tombul Ghiaghli’ from Turkey. Hazelnut is cultivated on only a few hectares in Portugal and Slovenia. Beginning in the 1980s, several introduced cultivars have been evaluated for growth and nut production in both countries alongside local cultivars (Solar and Štampar 2011). In Greece, hazelnut cultivation began when Greek immigrants came from the Pontus region in northern Turkey. With them they brought Turkish cultivars that are still cultivated today, such as ‘Extra Ghiaghli’ (a clone of ‘Tombul’), ‘Tombul Ghiaghli,’ and ‘Sivri Ghiaghli.’ Landraces surveyed in Spain and Italy showed identical SSR profiles and similar morphological traits with important local cultivars: ‘Casina’ in northern Spain, ‘TGL’ in Piedmont, ‘TGR’ in Latium, and ‘Nocchione’ (syn. ‘Siciliana’) in central and southern Italy. Hazelnut growing has a strong

tradition in Spain and Italy, where orchards have a limited number of cultivars. In Spain, the northeastern province of Tarragona (Catalonia) accounts for 88 % of the total area planted to hazelnut, and ‘Negret’ is the most widely planted cultivar. Minor hazelnut-growing areas include Asturias and adjacent regions in northern Spain, where cultivated forms are found in small orchards and gardens. In the past, hazelnut was an important crop in this region, and ‘Casina’ was the most commonly grown cultivar (Ferreira et al. 2010; Campa et al. 2011). In Italy, 98 % of the producing surface is located in four regions: Piedmont, Latium, Campania, and Sicily. Production is limited in Liguria, Sardinia, Emilia, Veneto, and Calabria. A wide and varied hazelnut germplasm base exists in the Italian Peninsula, but it has been little studied in Sicily and minor growing regions. In Campania, seven main varieties are cultivated (‘Camponica,’ ‘Mortarella,’ ‘Riccia di Talanico,’ ‘San Giovanni,’ ‘Tonda di Giffoni,’ ‘Tonda bianca,’ and ‘Tonda rossa’). In Piedmont (‘TGL’) and Latium (‘TGR’), production is based on a single cultivar, while ‘Nocchione’ (syn. ‘Siciliana’) is the main pollinizer of ‘TGR’ in Latium.

In their studies, Boccacci et al. (2006) and Gökirmak et al. (2009) reported that ‘Nocchione’ and a group of Sicilian cultivars, renamed ‘Siciliana’ by Alberghina (1982), had identical profiles at 24 and 21 SSR loci, respectively. It was an unexpected result, since these cultivars are grown in two distant Italian regions: Latium and Sicily (Boccacci et al. 2006). Our results confirm this synonymy and clarify the origin of ‘Nocchione.’ STRUCTURE analyses revealed that ‘Nocchione’ grouped (98 %) with the accessions from southern Italy and Sicily rather than with those from Latium (central Italy). The Bayesian clustering and admixture analysis can be considered a standard method to identify the ancestral populations from which cultivars originated and quantify genetic relationships with probabilities and proportions (Breton et al. 2008). Thus, our data indicate that ‘Nocchione’ originated in southern Italy, most likely in Sicily. Moreover, the molecular analysis of most Sicilian accessions surveyed in situ confirmed the existence of a single dominant cultivar in local orchards (Alberghina 1982). In fact, six ‘Caraffara’ accessions, known as ‘Nostrale’ in Enna province, showed the same SSR profile as ‘Siciliana,’ indicating that Sicily is very likely the origin of ‘Nocchione’ from which it spread to central and southern Italy. It is probable that ‘Nocchione’ was also introduced into Calabria during the second half of XIX century (Piccirillo et al. 2007), as indicated by the genetic identity between ‘Nocchione’ and ‘Tonda Calabrese.’

Morphological characterization revealed a wide diversity among the 42 unique landraces. The H index was high (average of 1.1), and most of the phenotypic classes were present for the evaluated descriptors (Table 2). These accessions should be considered original and valuable PGR and additional local genetic diversity which needs to be conserved in situ. In addition, some landraces showed morphological and

technological traits appreciated by the market (Table 3). Accessions ‘Robriguedo-2’ (Asturias), ‘Noscello’ (Liguria), ‘Barrettona,’ ‘Itavex,’ ‘Cappello del Prete,’ ‘Madonnella’ (Latium), and ‘Selvaggiola Tardiva SIC12’ (Sicily) were interesting for the food industry. Nuts with globular or ovoid shape, kernels with medium size, and a caliber ≥ 12 mm are the ideal traits for the industry processing (Garrone and Vacchetti 1994). On the contrary, ‘Selvaggiola SIC3,’ ‘Trichette’ (Sicily), ‘San Vicino Vittori’ (Latium), and ‘T/16’ (Slovenia) showed the large nut and kernel size desired by the in-shell market.

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 UPGMA tree (Fig. 2), the PCoA scatter plot (Fig. 3), and the STRUCTURE analyses (Fig. 4) revealed a high level of differentiation between wild and cultivated forms. The wild genotypes from Latium and Campania were closely related but were separated from cultivars and landraces. Nevertheless, an introgression and admixture of genotypes between wild accessions and some landraces from Campania (‘Tonda bianca’ and ‘Tonda rossa’) or from Latium (‘Nocciola centenaria,’ ‘Cappello del prete,’ and ‘Barrettona’) was identified. Similar results were also obtained by Campa et al. (2011) between 40 wild hazelnuts collected in northern Spain and 62 locally cultivated accessions, investigated at 13 SSR markers. The SSR data 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).

The cultivated forms showed the tendency to constitute two main groups located in the Mediterranean basin in the West (Spain–Italy) and Black Sea basin in the East (Turkey). They are two of the four major geographical gene pools described in the European hazelnut: English, central European, Spanish–Italian, and Black Sea (Gökirmak et al. 2009). A high level of genetic similarity between cultivars grown in the Iberian and Italian Peninsula has also been reported by other authors (Bocacci and Botta 2010; Bocacci et al. 2006; Gürcan et al. 2010a). In our study, almost all accessions from the Iberian Peninsula were separated from Italian ones. The cultivars from northeastern Spain (Tarragona) were more closely related to accessions surveyed in northern Spain (Asturias) than to cultivars in central and southern Italy; the landraces surveyed in Asturias showed a tendency to cluster into a separate ‘Casina’ group. These results agrees with those of Campa et al. (2011) and encourage their hypothesis that local hazelnuts belong to the northeastern Spanish germplasm, but constitute a separate group. A significant genetic differentiation between the Spanish and Italian gene pools was also observed by Bocacci and Botta (2010). Then, it is probable that Spain and Italy are two hazelnut diversification areas,

where a shared gene flow occurred between the western and central Mediterranean basin as a consequence of human migration and trade during and after the Roman civilization (Bocacci and Botta 2009, 2010). Among the accessions from Italy, most from central and southern areas comprised the largest gene pool, while some landraces surveyed in Liguria (‘Gianchetta,’ ‘Dell’Orto,’ ‘Tapparona,’ and ‘Del Rosso’) were arranged in a differentiate group. A congruent topology was reported in the PCA scatter plot obtained from morphological data (Fig. 1). These genetic and morphological data indicate less gene flow between northern and southern Italy, whereas exchange of plant material very likely occurred in southern Italy between Campania and Sicily. The existence of a main gene pool in southern Italy supported the hypothesis that it was an important centre of origin and diffusion of hazelnut cultivars, as suggested by Bocacci and Botta (2009) who analyzed cultivars from Spain, Italy, Turkey, and Iran at 13 chloroplast SSR (cpSSR) loci. The grouping of several Italian landraces did not fit their geographic origin. Some accessions from Latium (‘San Vicino Vittori,’ ‘Nocciola Benedetta,’ ‘Nocciola lunga,’ and ‘Allungata’) and Liguria (‘Trietta,’ ‘Lunghera,’ and ‘Seigretta’) were genetically closer to Turkish cultivars, while others from Liguria (‘Noscello,’ ‘Menoia,’ ‘Catainetto,’ and ‘Bardina’) showed a genetic similarity to Spanish accessions. It is likely that gene flow occurred from the western Mediterranean to northern Italy and from the Black Sea to northern and central Italy, most likely as a consequence of commercial exchanges. In fact, during the eleventh century, hazelnuts produced in Turkey were traded in Liguria on the Genoa market (Rosengarten 1984). These results supported the hypothesis that hazelnut cultivation and cultivars were not introduced from the eastern Mediterranean/Black Sea basin into Campania and Sicily by Greeks or by Arabs, as reported by Bocacci and Botta (2009).

The genetic diversity calculated between each geographical gene pool pair (Table 5) also supported the above mentioned considerations: (a) high genetic differentiation between northern and southern Italian groups, (b) low genetic diversity among central and southern Italian gene pools, (c) higher genetic similarity between Iberian and north–west Italy groups and between Balkans–Black Sea and northwestern and central Italian groups, and (d) low gene flow between southern Italy and Black Sea. Finally, these results also indicated that north-eastern Spain, southern Italy, and Black Sea could have been three important hazelnut domestication areas.

Archeological findings, historical documents, pollen data, and cpSSR analysis support the hypothesis that Campania (southern Italy) was an important centre of origin and diffusion of hazelnut cultivars (Bocacci and Botta 2009). It seems likely that this germplasm originated from the post-glacial refuge in southern Italy (Palmé and Vendramin 2002) and spread around the Mediterranean Sea beginning with the Roman civilization. Our data support this hypothesis,

indicating that a varietal sharing occurred among the regions Latium, Campania, and Sicily. Moreover, genetic relationships also showed that the Sicilian cultivars ‘Napoletana’ and ‘Napoletaneda’ were very close to those from Campania, indicating their introduction to Sicily from Campania. Thus, it can be hypothesized that a first gene flow occurred from Campania southward to Sicily and northward to Latium, while a second gene flow carried cultivars from Sicily to Latium and Calabria. These results confirm that hazelnut cultivation was not introduced to Sicily by Arabs but from Campania by the Romans. The Arabs dominated Sicily beginning in the second half of the ninth century, but hazelnuts were already being cultivated in Roman times (Bocacci and Botta 2009).

In summary, the molecular and morphological characterizations of surviving *on-farm* landraces were useful for identifying duplications and mistakes as well as the most interesting accessions and provided justification for their *in situ* preservation. These materials have been grafted and propagated on their own roots for planting in two hazelnut collections: IRTA in Reus (Spain) and in the country where the material originated, for evaluation *ex situ* and use in breeding. Findings about genetic relationships and population structure also raise an interesting question about the origin and diffusion of the hazelnut germplasm cultivated in southern Europe. According to several authors (Bocacci and Botta 2009, 2010; Gökirmak et al. 2009; Gürçan et al. 2010a), *C. avellana* seems to have been domesticated independently in six different areas: British Islands, central Europe, Spain, Italy, Black Sea, and Iran. Our results are in agreement with these conclusions, indicating the existence of three main germplasm groups in the Mediterranean basin which could correspond with three domestication areas: northwestern Spain and southern Italy in the West and the Black Sea region in the East. Moreover, the data indicate the existence of secondary gene pools in the Iberian (Asturias) and Italian (Liguria and Latium) Peninsulas, where local varieties have been domesticated in subsequent times from wild forms and/or from the introduction of ancient domesticate varieties, followed by a relatively local evolution that could include crosses among them and with local hazelnuts. The introduction of plant material from other areas influenced the local gene pool, but it is more likely that this was due to introgression of genes from foreign germplasm into local accessions followed by selection rather than to the direct adoption of introduced cultivars.

Acknowledgments Authors thank Prof. Shawn A. Mehlenbacher (Oregon State University, Corvallis, OR, USA) for reviewing the manuscript. The research was funded by the European Union project SAFENUT (AGRI GEN RES 068).

Data Archiving Statement Following the standard Tree Genetics and Genomes policy, we declare that the list of accessions surveyed (Table 1), phenotype data (Table 3), and SSR data (Supplementary Table 1) were deposited in the Dryad Repository (<http://datadryad.org/>).

References

- Alberghina O (1982) Indagine sulla corilicoltura siciliana. *Frutticoltura* 2:27–30
- Bacchetta L, Di Giovanni B, Avanzato D, Botta R, Bellon B, Bocacci P, Drogoudi P, Metzidakis I, Rovira M, Sarraquigne JP, Silva AP, Solar A, Spera D (2011) The AGRI GEN RES ‘SAFENUT’ action: a European strategy for the preservation and utilization of hazelnut and almond genetic resources. *Nucis-Newsletter* 15:30–33
- Bassil NV, Botta R, Mehlenbacher SA (2005a) Microsatellite markers in the hazelnut: isolation, characterization and cross-species amplification in *Corylus*. *J Am Soc Hort Sci* 130:543–549
- Bassil NV, Botta R, Mehlenbacher SA (2005b) Additional microsatellites of the European hazelnut. *Acta Hort* 686:105–110
- Bioversity, FAO, CIHEAM (2008) Descriptors for hazelnut (*Corylus avellana* L.). Bioversity International, Rome, Italy; Food and Agriculture Organization of the United Nations, Rome, Italy; International Centre for Advanced Mediterranean Agronomic Studies, Zaragoza, Spain
- Bocacci P, Botta R (2009) Investigating the origin of hazelnut (*Corylus avellana* L.) cultivars using chloroplast microsatellites. *Genet Resour Crop Evol* 56:851–859
- Bocacci P, Botta R (2010) Microsatellite variability and genetic structure in hazelnut (*Corylus avellana* L.) cultivars from different growing regions. *Sci Hortic* 124:128–133
- Bocacci P, Akkak A, Bassil NV, Mehlenbacher SA, Botta R (2005) Characterization and evaluation of microsatellite loci in European hazelnut (*Corylus avellana* L.) and their transferability to other *Corylus* species. *Mol Ecol Notes* 5:934–937
- Bocacci P, Akkak A, Botta R (2006) DNA-typing and genetic relationships among European hazelnut (*Corylus avellana* L.) cultivars using microsatellite markers. *Genome* 49:598–611
- Bocacci P, Rovira M, Botta R (2008) Genetic diversity of hazelnut (*Corylus avellana* L.) germplasm in northeastern Spain. *HortSci* 43:667–672
- Breton C, Pinatel C, Médail F, Bonhomme F, Bervillé A (2008) Comparison between classical and Bayesian methods to investigate the history of olive cultivars using SSR-polymorphisms. *Plant Sci* 175:524–532
- Campa A, Trabanco E, 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
- CBD (1992) Convention on biological diversity: text and annexes. Secretariat of the Convention on Biological Diversity, Montreal
- CBD (2002) Global strategy for plant conservation. Secretariat of the Convention on Biological Diversity, Montreal. Available from <http://www.biodiv.org/decisions/?lg=0&dec=VI/9>. Accessed 29 Oct 2012
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genet Resour* 4:359–361
- 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
- FAO (2001) International treaty on plant genetic resources for food and agriculture. Available from: <http://www.planttreaty.org/content/texts-treaty-official-versions>. Accessed 29 Oct 2012
- FAOstat (2011) Agriculture data. Available from: <http://faostat.fao.org/site/408/default.aspx>. Accessed 31 May 2011
- Ferreira JJ, Garcia-González C, Tous J, Rovira M (2010) Genetic diversity revealed by morphological traits and ISSR markers in hazelnut germplasm from northern Spain. *Plant Breed* 129:435–441

- Garrone W, Vacchetti M (1994) Hazelnut quality in relation with the requirements of confectionary industry. *Acta Hort* 351:641–648
- Gökirmak T, Mehlenbacher SA, Bassil NV (2009) Characterization of European hazelnut (*Corylus avellana* L.) cultivars using SSR markers. *Genet Resour Crop Evol* 56:147–172
- Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *J Hered* 86:485–486
- Gürçan K, Mehlenbacher SA (2010a) Transferability of microsatellite markers in the Betulaceae. *J Amer Soc Hort Sci* 135:159–173
- Gürçan K, Mehlenbacher SA (2010b) Development of microsatellite marker loci for European hazelnut (*Corylus avellana* L.) from ISSR fragments. *Mol Breed* 26:551–559
- Gürçan K, Mehlenbacher SA, Erdoğan V (2010a) Genetic diversity in hazelnut (*Corylus avellana* L.) cultivars from Black Sea countries assessed using SSR markers. *Plant Breed* 129:422–434
- Gürçan K, Mehlenbacher SA, Botta R, Boccacci P (2010b) Development, characterization, segregation, and mapping of microsatellite markers for European hazelnut (*Corylus avellana* L.) from enriched genomic libraries and usefulness in genetic diversity studies. *Tree Genetics & Genomes* 6:513–531
- Hammer Ø, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. *Palaeontol Electron* 4(1), 9pp
- Jarvis DI, Myer L, Klemick H, Guarino L, Smale M, Brown AHD, Sadiki M, Sthapit B, Hodgkin T (2000) A training guide for in situ conservation *on-farm*, Version 1. IPGRI, Rome
- Mehlenbacher SA, Miller AN (1989) 'Barcelona' hazelnut. *Fruit Varieties J* 43:90–95
- Mehlenbacher SA, Brown RN, Nouhra EN, 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
- Minch E (1997) MICROSAT. Available from: <http://genetics.stanford.edu/hppl/projects/microsat/>. Accessed 29 Oct 2012
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583–590
- Palmé AE, Vendramin GG (2002) Chloroplast DNA variation, postglacial recolonization and hybridization in hazel, *Corylus avellana*. *Mol Ecol* 11:1769–1779
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Piccirillo P, Petriccione M, Clasadonte A, Rosato T, De Luca A (2007) Problematiche e prospettive di rilancio per la corilicoltura calabrese. *Frutticoltura* 11:62–66
- Polegri L, Negri V (2010) Molecular markers for promoting agrobiodiversity conservation: a case study from Italy. How cowpea landraces were saved from extinction. *Genet Resour Crop Evol* 57:867–880
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Rosengarten F (1984) The book of edible nuts. Walker, New York
- Rovira M, Avanzato D, Bachetta L, Botta R, Drogoudi P, Ferreira JJ, Sarraquigne JP, Silva AP, Solar A (2011) European *Corylus avellana* germplasm collections. *Acta Hort* 918:871–876
- Sathuvalli VR, Chen H, Mehlenbacher SA, Smith DC (2011) DNA markers linked to eastern filbert blight resistance in "Ratoli" hazelnut (*Corylus avellana* L.). *Tree Genetics & Genomes* 7:337–345
- Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana
- Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43:1349–1368
- Solar A, Štampar F (2011) Characterization of selected hazelnut cultivars: phenology, growing and yielding capacity, market quality and nutraceutical value. *Jour Sci Food Agric* 91(7):1205–1212
- 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 Evol* 28:2731–2739
- Thomas MR, Matsumoto S, Cain P, Scott NS (1993) Repetitive DNA of grapevine: classes present and sequences suitable for cultivar identification. *Theor Appl Genet* 86:173–180
- Thompson MM, Romisondo P, Germain E, Vidal-Barraquer R, Tacias Valls J (1978) An evaluation system for filberts (*Corylus avellana* L.). *HortSci* 13(5):514–517
- 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
- UPOV (1979) Guidelines for the conduct of test for distinctness, homogeneity and stability (hazelnut). UPOV, Geneva. TG/71/3
- Wagner HW, Sefc KM (1999) IDENTITY 1.0. Available from: <http://www.uni-graz.at/~sefck/>. Accessed 29 Oct 2012
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370
- Yeh FY, Boyle R, Ye T, Mao Z (1997) POPGENE, the user-friendly shareware for population genetic analysis, version 1.31. Molecular Biology and Biotechnology Centre, University of Alberta, Alberta