Genet Resour Crop Evol
DOI 10.1007/s10722-006-0034-0

RESEARCH PAPER

Identification of 'Sorrento' walnut using simple sequence repeats (SSRs)

I. Foroni · K. Woeste · L. M. Monti · R. Rao

Received: 21 December 2005 / Accepted: 13 June 2006

Abstract Among the Italian walnut (*Juglans regia* L.) landraces, 'Sorrento' is the most famous. It originated in the Sorrento peninsula, but today it is grown in the entire Campania region. 'Sorrento' is a mixture of genetic entities and is often heterogeneous for important commercial traits such as fruit size and yield. Consequently, the morphological description of 'Sorrento' is extremely difficult. To evaluate the genetic diversity in 'Sorrento' walnut, we analyzed 16 'Sorrento' plants grown in Caserta (10 originated from seeds and six from grafts), and 26 grafted 'Sorrento' clones grown in the Sorrento peninsula. We compared their genotypes along with six other walnut cultivars using 12 microsatellite (SSR) markers. A total of 66 putative alleles were detected, 16 of which were unique to one individual. Two loci, WGA9 and WGA71, were particularly useful for distinguishing Caserta samples from Sorrento peninsula clones. The phylogenetic and STRUCTURE analysis highlighted the genetic distance between the Sorrento peninsula and Caserta groups, assigning the samples to two different clusters (or populations) corresponding closely, but not perfectly, to each sample's geographic origin. The name 'Sorrento' walnut should not be assigned to samples falling outside the two genotypic clusters, even if they have phenotypes traditionally associated with the 'Sorrento' walnut.

Keywords Genetic diversity · Genetic structure · *Juglans regia* · Landrace · Microsatellites · Campania

I. Foroni · L. M. Monti · R. Rao (✉)
Dipartimento di Scienze del Suolo della Pianta e dell'Ambiente, Università degli studi di Napoli ''Federico II'', via Universita' 100, 80045 Portici, Italy
e-mail: rao@unina.it

K. Woeste
USDA Forest Service, Hardwood Tree Improvement and Regeneration Center, Department of Forestry and Natural Resources, Purdue University, 715 West State Street, West Lafayette, IN 47907-2061, USA

Introduction

Juglans regia (walnut) seems to have originated in a large area among the mountains of central Asia (Nekrassowa 1927; Komarow 1936; Berg 1937; Browicz 1976). Thousands of years ago, Greeks introduced wild *J. regia* germplasm to the Balkan Peninsula, Turkey and southeastern Europe. From there, the species spread to western Europe and northern Africa via trade within the Roman empire (Beug 1975; Davis 1982; Huntley and Birks 1983). In the 18th century, walnut began to be cultivated in South America by Colonial Spanish who exported *J. regia* genotypes from Europe (Bazzanella et al. 2001). Finally, in the 19th century, walnuts were imported to North America

separately by two men, Felix Gillet, who imported *J. regia* plants from France to Northern California, and Joseph Sexton, who imported *J. regia* plants from China to southern California (Beede and Hasey 1998; Tuleke and McGranahan 1994).

Italy produced more walnuts (*J. regia* L.) than any other country in Europe until the beginning of the 20th century. Italy lost its primacy, becoming an importer of in-shell walnuts, as a consequence of the high cost of traditional farming practices and mixed quality standards that damage the commercial competitiveness of Italian producers (Marangoni et al. 2004). In fact, the high geoclimatic and environmental variability of Italy is reflected in its crops that, like walnut, are sometimes genetically heterogeneous (Ducci 2001). While variability within a crop has long-term benefits, it can have negative consequences commercially.

Walnut cultivation is carried out over all of Italy, from Piedmont to Sicily, but it is concentrated in Campania, which produces 75% of the Italian crop (Bazzanella et al. 2001). Italian varieties are made up of genotypes that are similar but not identical (i.e. without well-defined phenotypes), and the varieties are called by the name of their region of origin (e.g. 'Sorrento', 'Blegiana', 'Cerreto', 'Feltrino' and 'Benevento'). These ''cultivars'' are best understood as landraces [...geographically or ecologically distinctive crop populations, adapted but variable...in equilibrium with both environment and pathogens and genetically dynamic...the result of millennia of natural and artificial selection...lacking formal definition and organization (Pingali and Smale 2001; Harlan 1975)].

Among the Italian walnut landraces, 'Sorrento' is the most famous. It originated in the Sorrento peninsula, but today it is grown in the entire Campania Region (Southern Italy). Most 'Sorrento' plantations were established, historically, for both fruit and timber production. Often these trees originated by grafting, (although scions were not always from local genotypes), but sometimes the scions were seedlings of the 'Sorrento' landrace. In both cases the plants were labeled as 'Sorrento' (Regione Campania—Assessorato all'Agricoltura 2004). Over time, these cultural habits brought together a mixture of genetic entities that originated from 'Sorrento' but that include a large range of genetic variability.

The evaluation of the molecular diversity of plants labeled 'Sorrento' growing throughout Campania is an important pre-requisite for the selection of closely related plants producing fruits of standard quality that can compete with those produced in France or the USA. Knowing and understanding the genetic structure of landrace populations is also useful for developing an optimal strategy for in situ conservation (Gomez et al. 2004), for breeding programs (Portis et al. 2004), for Denomination of Protected Origin certification (Gemas et al. 2004), or to understand how the landraces adapt to different ecological and environmental stresses (Farid et al. 2000). In the last decade, SSR markers have become an important tool to understand the molecular diversity and the genetic relationships within crops such as tomato (He et al. 2003), pear (Yamamoto et al. 2002), coconut (Meerow et al. 2003), and peach (Dirlewanger et al. 2002).

Previous studies of walnut genetic diversity were carried out with isozymes (Arulsekar et al. 1986; Solar et al. 1994), RFLP (Fjellstrom et al. 1994), RAPD (Nicese et al. 1998), and ISSR (Potter et al. 2002). A few years ago, we started using AFLP to study the molecular diversity of 'Sorrento' walnuts in Campania that originated from seed and from vegetative propagation (Andreakis et al. 2002). The data indicated a high level of variability among the seed propagated plants, and clearly indicated that the plants that grew from seed had different origins than the vegetatively propagated ones. The vegetatively propagated plants, although not identical, were similar to one another, indicating that they originated from genetically related mother plants. In a following study, we analyzed the same genotypes using SSRs, and found that three 'Sorrento' samples were genetically distant from most of the other 'Sorrento' genotypes. We also described evidence of gene flow between 'Sorrento' and the American cultivars 'Serr' and 'Hartley' (Foroni et al. 2005). In the present study, our aim was to extend the number of samples analyzed and the number of SSR loci observed to characterize the 'Sorrento' plants grown in different parts of Campania and in particular to compare the genetic diversity obtained in the previously

analyzed sample set with plants grown in the Sorrento peninsula, possibly the region containing the largest genetic variability of the original 'Sorrento' landrace.

Materials and methods

Plant material

All the samples originated from grafted 'Sorrento' or seed grown for several years in Campania (Fig. 1). Seedling plants (S0–S9) were arbitrarily selected from a pool of genotypes that originated from 'Sorrento' seeds belonging to the germplasm collection of the Istituto Sperimentale per la Frutticoltura (ISF-CE) (Table 1). Two of the seedling plants had traits associated with ancient origins; one had rounded fruits (S9), and the other one (S0), had elongated fruits, a phenotype considered the most widespread in Campania in the past (Bazzanella et al. 2001). Grafted plants ($n = 32$) were from walnut plantations; six from the area around Nola (G) plus 26 from the Sorrento peninsula (P). Six walnut cultivars from Europe ['Blegiana' (B), 'Malizia'(M), 'Parisienne' (P), 'Franquette'(F)], or the US ['Hartley'(H), 'Serr'(SR)] were used as controls.

Fig. 1 Map of Campania region with the walnut growing areas (ringed) near Sorrento (MAPS: http://www.knowital.com)

Table 1 Origin, name and code of plants sampled

Group/location	Name[a]	Code	Genotypes[b]
'Sorrento' seedlings – *Caserta* –	MU 6	S1	A
	MA 1	S2	B
	SORRENTO XI	S3	C
	CC1 PT1	S4	D
	CC1 PT2	S5	E
	VR3 PT2	S6	F
	VR3 PT3	S7	G
	SORRENTO IV	S8	H
	CASOLLA	S9	I
	SORRENTO 30	S0	J
'Sorrento' Clones – *Caserta* –	NOLA 6	G1	K
	NOLA 13	G2	L
	NOLA 16	G3	M
	NOLA 19	G4	N
	NOLA 26	G5	O
	NOLA 36	G7	P
'Sorrento' Clones – *Sorrento Peninsula**–	BUONOCORE 2	P1	Q
	BUONOCORE 4	P2	Q
	CIOFFI 1	P3	Q
	CIOFFI 2	P4	Q
	DI PAOLA 2	P5	R
	DI PAOLA 4	P6	R
	DI PAOLA M. 1	P7	R
	DI PAOLA M. 3	P8	R
	DI PAOLA M. 4	P9	S
	MINIERI 1	P10	R
	MINIERI 2	P11	R
	MINIERI 3	P12	R
	MINIERI 8	P13	Q
	MONTUORI 1	P14	R
	MONTUORI 2	P15	Q
	MONTUORI 3	P16	Q
	MONTUORI 4	P17	Q
	MONTUORI 8	P18	R
	MONTUORI 10	P19	T
	VANACORE G.1	P20	U
	VANACORE G.2	P21	V
	VANACORE G.4	P22	W
	VANACORE L. 2	P23	X
	VANACORE L.3	P24	Y
	SCARAMELLINO 1	P25	R
	SCARAMELLINO 3	P26	R
Cultivars	'BLEGIANA'	B	Z
	'FRANQUETTE'	F	A1
	'HARTLEY'	H	B1
	'MALIZIA'	M	C1
	'PARISIENNE'	P	D1
	'SERR'	SR	E1

[a] Names refer to cultivars, plantations or germplasm

[b] Samples sharing the same letter or symbol were indistinguishable based on the microsatellites described here

[*] Samples never previously analyzed

DNA extraction

Young leaves were ground in liquid nitrogen, and the DNA extracted and stored at − 80 °C. The extraction protocol of Doyle and Doyle (1987) was modified as follows: 1 g of frozen, ground leaves was added to 10 ml of preheated (50 °C) 2 × CTAB buffer (2% CTAB, 50 mM DTT, 0.3% β-mercaptoethanol, 1.4 M NaCl, 100 mM Tris, 20 mM EDTA, pH = 8.0) and incubated at 65 °C for 30 min. The aqueous solution was extracted with 20 ml of (24:1) chloroform–octanol,

centrifuged 15 min at 13,000 rpm, and the aqueous layer retained. Two volumes of 100% cold ethanol were added to precipitate the nucleic acids. The precipitate was spooled, washed with 0.2 M ammonium acetate in 75% ethanol and air-dried for 5 min. The pellet was then resuspended in 500 µl of 65 °C preheated H_2O and treated with 125 µl (10 mg/ml) RNase at 37 °C for 30 min. The DNA was precipitated, washed, dried, resuspended in 250 µl H_2O and quantified in a 1% agarose gel stained with ethidium bromide (10 mg/ml) against DNA lambda standards.

SSR primers and amplification

Eleven primers designed from the sequences of clones isolated from an enriched (GA/CT)n library of *J. nigra* (Woeste et al. 2002) were used to amplify genomic DNA of *J. regia* cultivars and 'Sorrento' samples. Amplification products from these reactions were separated on 2% agarose gel (not shown). Six primers combinations that produced clear and polymorphic products without artifactual bands (WGA69, 89, 118, 202, 276, and 321) were selected (Table 2)

PCR reactions contained 3.3 mM $MgCl_2$, 0.10 mM dNTPs, 0.5 unit *Taq* polymerase (New England Biolab, Beverly, MA), 0.4 µM unlabeled primer (reverse), 0.4 µM forward primer labeled with Hex, Ned or 6 Fam consistent with matrix D on an ABI 377 automated sequencer (Applied Biosystems, Foster City, CA), 0.4% BSA, and 15 ng DNA template in a final volume of 10 µl. Thermal cycling conditions were as follows: denaturation 5 min at 95 °C; 35 cycles of 1 min at 94 °C, 1 min at 55 °C and 2 min at 72 °C; and final extension 10 min at 72 °C. Each amplification included a negative control reaction without DNA template.

Table 2 Properties of the microsatellite loci used to characterize 48 walnut genotypes

Locus	Primers sequences 3′-5′	T.m. (°C)	Size range (bp)	Total alleles	Gel set	Private alleles[1]
WGA1*	ATTGGAAGGGAAGGGAAATG CGCGCACATACGTAAATCAC	48	181–195	3	a	
WGA5*	CAGTTTGTCCCACACCTCCT AACCCATGGTGAGAGTGAGC	47	240–266	7	a	252 bp: S2 256 bp: F 266 bp: B
WGA9*	CATCAAAGCAAGCAATGGG CCATTGCTCTGTGATTGGG	48	236–248	6	b	236 bp: G4
WGA27*	AACCCTACAACGCCTTGATG TGCTCAGGCTCCACTTCC	51	192–210	5	a	192 bp: F
WGA32*	CTCGGTAAGCCACACCAATT ACGGGCAGTGTATGCATGTA	47	120–196	4	b	120 bp: S1
WGA69	TTAGTTAGCAAACCCACCCG AGATGCACAGACCAACCCTC	56	162–182	6	c	164 bp: G5 176 bp: P20
WGA71*	ACCCGAGAGATTTCTGGGAT GGACCCAGCTCCTCTTCTCT	45	136–210	8	b	136 bp: S1
WGA89	ACCCATCTTTCACGTGTGTG TGCCTAATTAGCAATTTCCA	53	212–222	5	d	212 bp: P19 220 bp: SR
WGA118	TGTGCTCTGATCTGCCTCC GGGTGGGTGAAAAGTAGCAA	62	197–201	3	d	
WGA202	CCCATCTACCGTTGCACTTT GCTGGTGGTTCTATCATGGG	56	260–276	4	c	
WGA276	CTCACTTTCTCGGCTCTTCC GGTCTTATGTGGGCAGTCGT	55	168–194	6	d	168 bp: P24 194 bp: P19
WGA321	TCCAATCGAAACTCCAAAGG TGTCCAAAGACGATGATGGA	50	227–265	7	c	225 bp: SR 249 bp: S1 265 bp: S1
Total number of alleles				66		16
Mean				5.5		

[1] Size of a specific allele that is present only in one individual

* Primers previously published (Foroni et al. 2005)

To prepare samples for loading, every amplification was diluted up to 1:4 in water and then 1.0 μl PCR product was mixed with 0.5 μl blue dextran 2.5 ml formamide and 0.5 μl Tamra 500-labeled molecular size standard. The PCR products were combined into four gel sets (Table 2). Products within a gel set could be multiloaded to run in a single lane using a CAL 96 paper comb (The Gel Company, San Francisco, CA). The samples were denatured at 94 °C for 1 min before loading onto an ABI 377 sequencer (Applied Biosystems, Foster City, CA) and electrophoresed in the 6% polyacrylamide gels under denaturing conditions [7 M urea, 1 × TBE buffer (90 mM Tris borate, pH 8.3, and 2 mM EDTA)]. DNA fragment lengths were analyzed with the software GENESCAN (Applied Biosystems, Foster City, CA). Three replicate experiments were carried out for each SSR primer pair-genotype combination.

Evaluation of polymorphisms and data analysis

The SSR allele composition of each plant was determined and the observed heterozygosity (H_o) value of each microsatellite was calculated according to the formula (Nei 1987):

$$H_o = 1 - \sum_{j=1}^{n} x^2 ij$$

where x_{ij} is the frequency of the jth allele for each SSR marker i and summation extends over n alleles.

The CONVERT (Glaubitz 2003, http://www.agriculture.purdue.edu/fnr/htlm/faculty/rhodes/student%20and%20Staff/glaubitz/software.htm), MICROSAT (Minch et al. 1997) and PHYLIP (Felsenstein 1989) programs were used to analyze the data and to construct a neighbor-joining dendrogram based on allele frequency data calculated by Nei's genetic distance (Nei 1972). The program STRUCTURE (Pritchard et al. 2000) was used to investigate population structure, assign individuals to populations, and identify intermediate genotypes.

The software CONVERT facilitates the conversion of diploid genotypic data files into formats that can be directly read by a number of commonly used population genetic computer programs. It also can calculate a variety of genetic distance estimates, many of which are appropriate solely for microsatellites. MICROSAT can be used to generate numerous 'bootstrap' replicate distance matrices by re-sampling loci with replacement. We used PHYLIP to produce dendrograms based on allele frequency data (Glaubitz 2003).

The data obtained here were analyzed together with the data of the previous work (Foroni et al. 2005).

Results

Sixty six different amplicons with sizes from 120 to 276 bp were obtained at the 12 loci for the 48 genotypes analyzed. These products were all considered as different alleles. The total number of alleles per locus varied from 3 to 8, with a mean of 5.5 (Table 2). Each genotype had 1 or 2 alleles per locus, and the 12 loci analyzed were all polymorphic. Sixteen alleles were private to single individuals (Table 2). The allele present with the highest frequency (81.4)% was the 196 bp allele at the WGA32 locus, while the alleles with lowest frequency (1.1%) were the 164 and 176 bp alleles at locus WGA69. The most polymorphic locus was WGA71 with eight alleles.

Although a unique SSR genotype was generated for most of the samples, a few genotypes were quite similar. For instance genotypes S4 and S5 differed only at locus WGA321, while genotypes S0 and S9, G2 and G7 differed only at locus WGA9.

The Caserta group of genotypes (S0–S9, G1–G7) was more heterogeneous than the Sorrento peninsula clones (P1–P26) as measured by the average number of alleles per locus (Table 3). Thirty alleles were private to either the Caserta (23 alleles) or Sorrento peninsula (7 alleles) provenances. Eleven of the 30 private alleles were found in one individual only of the Caserta provenance (Table 3). The 242 and 250 bp alleles at locus WGA5 were found at a high frequency in Caserta genotypes (69%) while the 248 bp allele at the same locus was found in all Sorrento peninsula clones. A similar result was found at

Table 3 Genotypic diversity within and between Caserta and Sorrento walnut samples

Provenance	Samples	Unique genotypes	Alleles per locus (mean)	Private alleles (provenance) [a]	Private alleles (individual)[b]
Caserta	16	16	4.2	23	7
Sorrento	26	13	2.8	7	4

[a] Allele present in only Caserta group or Sorrento peninsula group

[b] Allele present in only one individual

locus WGA9, where 81% of Caserta samples had alleles 240 and 248, while every Sorrento peninsula clone had alleles 238 and 246 (Table 4).

Caserta, Sorrento peninsula and other cultivars groups had different levels of per-locus observed heterozygosity as well as total numbers of alleles (Table 5). An average heterozygosity of 0.46 ± 0.21 and 33 alleles was observed for the Sorrento peninsula genotypes, much lower then Caserta (0.59 ± 0.16 with 48 alleles) and other cultivars (0.68 ± 0.12 with 49 alleles). There was no significant difference among the three groups in the number of loci that were heterozygous and homozygous as determined by chi-square (data not shown).

The Nei's genetic distance among samples was calculated on the basis of allele frequency data, and clusters were generated using the neighbor-joining algorithm (Fig. 2). The structure of the resulting tree was tested by 1,000 bootstraps. The unrooted tree organized most of the samples into five main clusters. Bootstrap values clearly supported the relative positions of cluster I, II and V while, due to the low bootstrap value between cluster III and IV, branch collapsing cannot be ruled out. Cluster I includes the Italian cultivar 'Malizia' (M) and most of seedlings and clones from Caserta. The inclusion of the cultivar 'Malizia' in this cluster was not unexpected as this cultivar originated from an open pollination of 'Sorrento'. Cluster II included only samples from the Sorrento peninsula. Cluster III included the Caserta seedling S2, the 'Sorrento' peninsula clone P21, the American cultivar 'Hartley' (H), and the French cultivars ['Franquette' (F) and 'Parisienne' (P)]. The presence of the cultivars 'Franquette', 'Parisienne' and 'Hartley' in the same cluster is consistent with previous results indicating that 'Hartley' is derived from crosses involving French cultivars (Potter et al. 2002). The 'Hartley' and 'Parisienne' branches are supported by 60% bootstrap re-sampling. The presence of S2 and P21 in this cluster indicates that these two genotypes are distinct from other 'Sorrento' genotypes, and raises the possibility that there has been recent gene flow between 'Sorrento' and other non-Sorrento genotypes. In fact, S2 has a unique allele at WGA5 locus (252 bp), and the allelic composition of P21 suggests possible gene flow from 'Franquette'. Cluster IV included two Caserta seedling samples (S4 and S5) and the American cultivar 'Serr' (SR), confirming previous results (Foroni et al. 2005). However, Caserta seedling samples S0 and S9 and the Italian cultivar 'Blegiana' (B) were placed together in cluster V in contrast with previous work (Foroni et al. 2005). The association of S0 and S9 (seedling plants with nut traits associated with ancient origins) with 'Blegiana' may indicate that the two groups share some common ancestry. The same may be true for S4, S5, and 'Serr'. Four of the Sorrento peninsula clones (P9, P19, P23 and P24) remain outside the clusters.

When the Sorrento peninsula clones and Caserta samples (clones and seedlings) were analyzed without the six cultivars (B, F, H, M, P, and S) there were three clusters: cluster A included all Sorrento peninsula clones except P9, P19, P21, P23 and P24; cluster B included all Caserta genotypes except S2, S4, S5, S9, and S0; and cluster C included three Sorrento peninsula clones (P19, P21, P23) and five Caserta samples (S2, S4, S5, S9, S0) (Fig. 3). P9 and P24 samples remained outside all three groups. The organization of the three clusters underlines both the differentiation between the plants of the two locations (Caserta and Sorrento peninsula) and the genetic divergence of cluster C plants from the other two groups of 'Sorrento' genotypes, raising the possibility of a flow of genetic

Table 4 Allelic profiles of 48 samples tested using 12 SSR loci

Code	SSR loci											
	WGA1*	WGA5*	WGA9*	WGA27*	WGA32*	WGA69	WGA71*	WGA89	WGA118	WGA202	WGA276	WGA321
S1	181/195	–[a]	240/248	210	120/196	180/182	136/210	222	197/199	260/266	176/180	249/265
S2	193	252	238/244	206	196	162/178	210	214	197	264/266	192	227/243
S3	181/195	242/250	240	196/210	166/196	180/182	200/210	222	197/199	260/266	176/180	227/243
S4	195	242	240/244	206/210	166/196	178/180	198/208	218	199	260/266	174/192	239
S5	195	242	240/244	206/210	166/196	178/180	198/208	218	199	260/266	174/192	243/245
S6	181/195	242/250	240/248	196/210	196	180/182	200/210	222	197/199	260/266	176/180	243/245
S7	181/195	242/250	240/248	210	196	180/182	202/210	222	197	260/266	176/180	243/245
S8	181/195	242/250	238/248	196/210	196	180/182	202/210	222	197/199	260/266	176/180	243/245
S9	195	240/250	238	206/210	196	162	202/210	222	197	266/276	180	243/245
SO	195	240/250	240	206/210	196	162	202/210	222	197	266/276	180	243/245
G1	181/195	242/250	238	210	196	180/182	206	222	197/199	260/266	176/180	243/245
G2	181	242/250	238	196/210	196	180/182	198/208	222	199	260/266	176/180	227/243
G3	181/195	242/250	–	210	196	180/182	198/208	222	197/199	260/266	176/180	243/245
G4	181	240/250	236/240	196/210	196	180/182	198/206	222	199	260/266	176/180	227/243
G5	181	242	238/240	196/210	196	164	202/210	222	199	260/266	180	227/243
G7	181	242/250	238/244	196/210	196	180/182	198/208	222	199	260/266	176/180	227/243
B	195	266	240/246	206/210	170	178	200/210	222	197	276	176/180	243/245
F	193	242/256	240/246	192/206	170	162/178	198/208	218/222	197/199	264	192	243
H	193	240	238/244	204	196	162	200/210	218	201	260/266	180/192	227/243
M	181/195	240/248	238	210	196	180/182	200/208	214	197/199	260/266	176/180	227/243
P	195	240/250	238/244	204/206	170/196	162/182	200/210	218	201	260/264	192	227/243
SR	181/195	242	240/246	204	166/196	180/182	198/208	218/220	199	260/266	174/192	225/239
P14	181/193	240/248	238/246	206	196	180/182	204/206	222	197/199	260/266	176/180	243/245
P15	181/193	240/248	238/246	206	196	180/182	206/208	222	197/199	260/266	176/180	243/245
P16	181/193	240/248	238/246	206	196	180/182	206/208	222	197/199	260/266	176/180	243/245
P17	181/193	240/248	238/246	206	196	180/182	206/208	222	197/199	260/266	176/180	243/245
P18	181/193	240/248	238/246	206	196	180/182	204/206	222	197/199	260/266	176/180	243/245
P19	193	240/248	238/246	206	196	178	206/208	212/222	197/199	260/276	180/194	227/245
P10	181/193	240/248	238/246	206	196	180/182	204/206	222	197/199	260/266	176/180	243/245
P11	181/193	240/248	238/246	206	196	180/182	204/206	222	197/199	260/266	176/180	243/245
P12	181/193	240/248	238/246	206	196	180/182	204/206	222	197/199	260/266	176/180	243/245
P13	181/193	240/248	238/246	206	196	180/182	206/208	222	197/199	260/266	176/180	243/245
P1	181/193	240/248	238/246	206	196	180/182	206/208	222	197/199	260/266	176/180	243/245
P2	181/193	240/248	238/246	206	196	180/182	206/208	222	197/199	260/266	176/180	243/245
P5	181/193	240/248	238/246	206	196	180/182	204/206	222	197/199	260/266	176/180	243/245
P6	181/193	240/248	238/246	206	196	180/182	204/206	222	197/199	260/266	176/180	243/245
P3	181/193	240/248	238/246	206	196	180/182	206/208	222	197/199	260/266	176/180	243/245
P4	181/193	240/248	238/246	206	196	180/182	206/208	222	197/199	260/266	176/180	243/245

Table 4 continued

Code	SSR loci											
	WGA1*	WGA5*	WGA9*	WGA27*	WGA32*	WGA69	WGA71*	WGA89	WGA118	WGA202	WGA276	WGA321
P20	–	–	–	–	–	176/180	–	222	197/199	260/266	176/180	243/245
P21	–	–	–	–	–	178/182	–	218/222	197/199	260/264	180/192	243/245
P22	–	–	–	–	–	180/182	–	222	197/199	260/266	176/180	243/245
P7	–	–	–	–	–	180/182	–	222	197/199	260/266	176/180	243/245
P8	–	–	–	–	–	180/182	–	222	197/199	260/266	176/180	243/245
P9	–	–	–	–	–	180/182	–	222	197/199	260/266	176/192	243
P23	–	–	–	–	–	162	–	222	197/199	260/264	176/180	243/245
P24	–	–	–	–	–	180/182	–	222	197/199	260/264	168/176	227/243
P25	–	–	–	–	–	180/182	–	222	197/199	260/266	176/180	243/245
P26	–	–	–	–	–	180/182	–	222	197/199	260/266	176/180	243/245

[a] Missing sample

* Primers previously published (Foroni et al. 2005)

information between Caserta and Sorrento peninsula plants. Inside cluster B it was possible to identify two sub-clusters; sub-cluster I included five seedling plants (S1, S3, S6, S7 and S8) and two clonal plants (G1 and G3). Sub-cluster II included only clonal plants from Nola (G2, G4, G5 and G7). Samples S9/S0, and S4/S5 were very similar to each other (supported by 92% and 99% bootstrap re-sampling) but different from cluster A and B.

The samples were further analyzed with the software STRUCTURE to assign individuals to populations (Pritchard et al. 2000). A simple model of two distinct populations was imposed in keeping with the results of phylogenetic analysis and the geographic origins of the samples. The result showed that nearly all the samples could be assigned to one of the groups (Fig. 4), and nearly all the samples were assigned the provenance they originated from. S0 and S9, the Sorrento seedlings with nut shapes associated with traditional 'Sorrento' types, were both assigned to Caserta. Samples P21, P23 and P24, in spite being of Sorrento peninsula provenance, were grouped with a population mostly containing Caserta samples. Caserta samples G1, G4 and S2 and Sorrento peninsula sample P19 were placed in an intermediate position indicating uncertain affinities.

Discussion

Microsatellite markers are useful for genetic studies at varietal, species and genus level, due to the high conservation of the flanking regions (Hamza et al. 2004). In our work, interspecific transferability of walnut microsatellites, originally selected to study the genetic diversity in *J. nigra* (Woeste et al. 2002), has been tested on *J. regia* varieties and closely related members of a landrace. The average number of alleles per locus in this study was 5.5, much higher than 1.3 and 3.9 detected in *J. regia* with RAPDs (Nicese et al. 1998) and ISSRs (Potter et al. 2002), respectively. The high level of variability observed at 12 SSR loci we used is consistent with results from studies carried out on different species (Carriero et al. 2002).

Table 5 Heterozygosity index and total number of alleles value calculated to characterize 48 walnut samples

	Caserta genotypes		Other cultivars genotypes		Sorrento peninsula genotypes	
	H_o^a	N_a^b	H_o^c	N_a^d	H_o^e	N_a^f
Locus						
WGA1*	0.56	3	0.63	3	0.50	2
WGA5*	0.80	4	0.83	6	0.52	2
WGA9*	0.71	4	0.76	4	0.65	2
WGA27*	0.58	3	0.77	5	0.48	1
WGA32*	0.33	3	0.58	3	0	1
WGA69	0.72	5	0.77	4	0.55	5
WGA71*	0.80	7	0.74	4	0.62	3
WGA89	0.31	3	0.37	4	0.08	3
WGA118	0.51	2	0.71	3	0.50	2
WGA202	0.59	4	0.70	4	0.57	4
WGA276	0.54	4	0.70	4	0.57	5
WGA321	0.69	6	0.70	5	0.54	3
Mean	0.59 ± 0.16		0.68 ± 0.12		0.46 ± 0.21	
Total		48		49		33

[a, b] Number of alleles (N_a) and observed heterozygosity (H_o) calculated only for Caserta genotypes

[c, d] Number of alleles (N_a) and observed heterozygosity (H_o) calculated only Other cultivars genotypes

[e, f] Number of alleles (N_a) and observed heterozygosity (H_o) calculated only for Sorrento peninsula genotypes

* Primers previously published (Foroni et al. 2005)

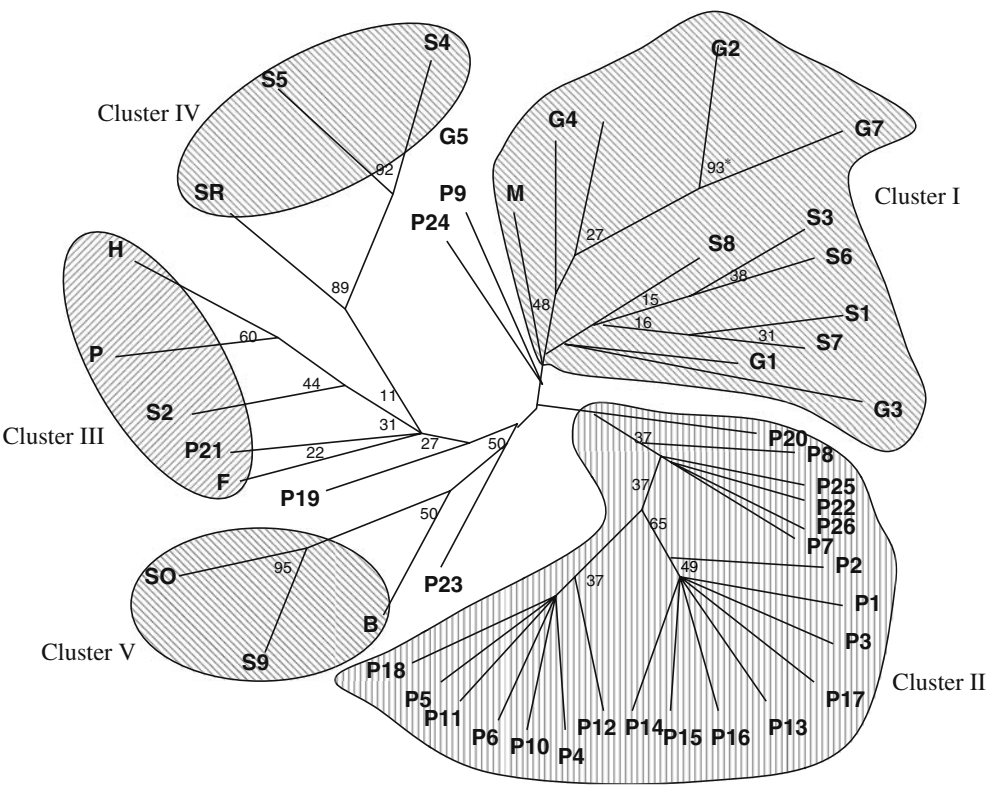

Fig. 2 Unrooted tree of 'Sorrento' landrace samples and *Juglans regia* varieties obtained using neighbor-joining algorithm. *Bootstraspping value

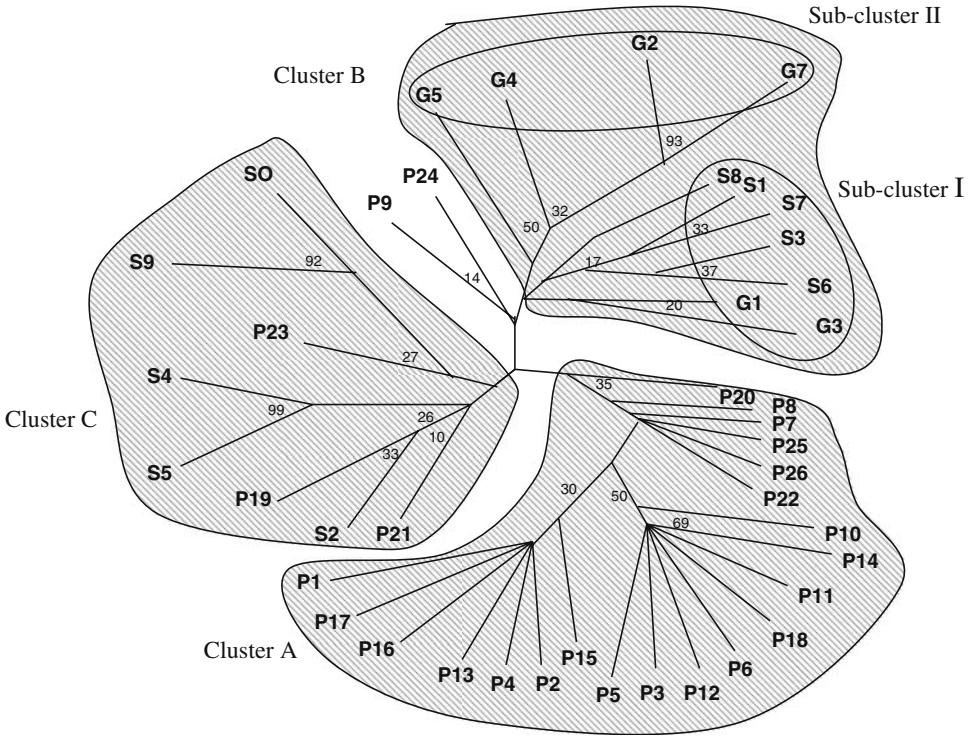

Fig. 3 Unrooted tree representing relationships among 'Sorrento' landrace samples and Caserta ones without the cvs B, F, H, M, P and S, obtained using 12 polymorphic microsatellites

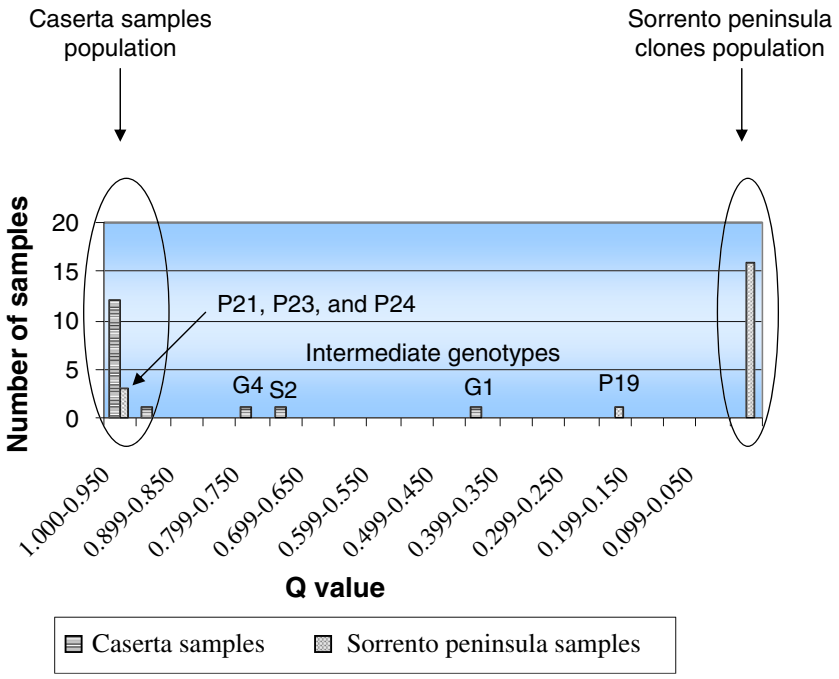

Fig. 4 Assignment of Campania walnut samples to populations based on 12 microsatellite loci using STRUCTURE. See also methods for details. Q value: the probability that the sample is correctly assigned to the Caserta population

In previous studies, we analyzed genetic variation among the Caserta genotypes originated by graft (G1–G7) and by seed (S0–S9) through AFLP and SSR markers (Andreakis et al. 2002; Foroni et al. 2005). We observed considerable variability among the genotypes, above all among the seed propagated plants, with some samples very genetically distant from most of the other 'Sorrento' genotypes. In this study, we added 26 samples (P1–P26), and analyzed all the samples using six new SSR loci (WGA69, WGA89, WGA118, WGA202, WGA276, and WGA321). Our objective was to estimate the genetic diversity present in the 'Sorrento' landrace and to tentatively identify plants that could be considered representative of the 'Sorrento' landrace.

A unique SSR genotype was found for 25 out of the 42 (60%) samples analyzed (Table 3). The genotypes P5–P8, P10–P12, P14, P18, P25, and P26 were identical, as were P1–P4, P13, and P15–P17. This result was not surprising as these trees were clonally propagated. Thirty alleles were present only in Caserta or only in Sorrento peninsula samples, permitting easy distinction of the two groups (Table 3). Loci WGA9 and WGA71 were the most informative ones with respect to their ability to distinguish the two provenances. The locus WGA71 showed five alleles present only among Caserta genotypes (alleles 136,198, 200, 202, and 210) and one allele found only among Sorrento peninsula clones (allele 204); locus WGA9 showed four alleles present only among Caserta genotypes (alleles 236, 240, 244, and 248) and one allele found only among Sorrento peninsula clones (allele 246). WGA9 and WGA71 are good candidate loci for additional screening of Campania samples.

The phylogenetic analysis highlighted the genetic distance between the Sorrento peninsula and Caserta groups, both commercially labeled as 'Sorrento' walnut. In fact, the neighbor-joining tree grouped almost all the Sorrento peninsula clones in cluster A and all the Caserta samples except S2, S4, S5, S9, and S0 in cluster B, indicating differentiation between the provenances. (Fig. 3).The level of similarity found among S9, S0, and 'Blegiana' either indicates gene flow among these genotypes or homoplasy (Fig. 2). The same is true for S4, S5, and SR, and for S2, P21, 'Hartley', 'Parisienne' and 'Franquette'. Gene flow into 'Sorrento' may be partially explained by the high level of outcrossing that occurs in all *Juglans* species as a result of heterodichogamy (McGranahan and Leslie 1990).

The STRUCTURE analysis assigned the Caserta samples and Sorrento peninsula clones to different populations corresponding closely, but not perfectly, with the samples' geographic origins. Samples P21, P23, and P24, in spite of their origin in the Sorrento peninsula, were assigned to Caserta. The placement of P24 into the Caserta population is not surprising, as that confirms the result of the cluster analysis, which showed P24 to be similar to 'Malizia'. 'Malizia' is believed to have originated from an open pollination of 'Sorrento'. It is possible that P21 and P23 are products of crosses between 'Sorrento' and non-Sorrento cultivars. Genotypes P19, S2, G1, and G4 could not be assigned and stayed outside the two populations. P19, G1, and G4 could be 'Malizia' offspring, as they were identical to the Italian cultivar at 10 (P19) and 9 (G1, G4) loci, while S2 was probably originated from cross between 'Sorrento' and non-Sorrento genotypes, as already discussed.

Both cluster analysis and STRUCTURE indicated that the Sorrento peninsula clones and Caserta samples are genetically different. Whether they will remain distinct will depend on how the forces of selection and gene flow interact. In the past, some farmers on the Sorrento peninsula appear to have performed a strong selection in favor of a small group of plants to improve yield and quality. However, the ancient types continue to be cultivated in Caserta.

We believe that the name 'Sorrento' walnut should not be assigned to samples outside cluster A and B (Fig. 3) even if they have phenotypes, like S0 and S9, traditionally associated with the 'Sorrento' walnut, as these phenotypes are no longer characteristic of 'Sorrento' genotypes.

References

Andreakis N, Piccirillo P, Santangelo I, Di Vaio C, Monti LM, Rao R (2002) Diversità molecolare di biotipi di noce provenienti da semenzali e innesti della cv. 'Sorrento'. Frutticoltura LXIV, Gennaio: 71–74

Arulsekar S, McGranahan GH, Parfitt DE (1986) Inheritance of phosphoglucomutase and esterase isozymes in Persian walnut. J Hered 77:220–221

Bazzanella G, Gilli G, Montedoro M (2001) Castagno e noce nella provincia di Trento. In: Artemisia Casa Editrice Trento (ed), pp 68–78

Beede RH, Hasey JK (1998) History of the walnut in California. In: Ramos DE (ed) Walnut production manual. Univ. of Calif. Div., of Agr. natural Resources, Oakland, pp 8–15

Berg LS (1937) Natural regions of the U.S.S.R. (Translated by O.A. Titelbaum). In: Macmillan Co. (ed) N.Y. 1950

Beug HJ (1975) Man as a factor in the vegetational history of the Balkan peninsula. In: Publishing House of the Bulgarian Academy of Sciences (ed) Problems of Balkan flora and vegetation. Proc. First International Symposium on Balkan Flora and Vegetation. Varna, Bulgaria. pp 72–78

Browicz K (1976) Juglandaceae. In: Rechinger KH (ed) Flora Iranica, Akademische Druck-u Verlagsanstalt, vol 121, pp 1–5

Carriero F, Fontanazza G, Cellini F, Giorio G (2002) Identification of simple sequenze repeats (SSRs) in olive (*Olea europaea* L.). Theor Appl Genet 104:301–307

Davis PH (1982) Flora of Turkey and the Aegean Islands, 654 p

Dirlewanger E, Cosson P, Tavaud M, Aranzana MJ, Poizat C, Zanetto A, Arus P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor Appl Genet 105:127–138

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Ducci F (2001) Conservazione ed utilizzazione delle risorse genetiche forestali in Italia. http://www.forgen.net/main/forum.asp

Farid A, O'Reilly E, Dollard C, Kelsey CR Jr (2000) Genetic analysis of ten sheep breeds using microsatellite markers. Can J Anim Sci 80 (1):9–17

Felsenstein J (1989) PHYLIP—Phylogeny Inference Package (Version 3.2). Cladistics 5:164–166

Fjellstrom RG, Parfitt DE, McGranahan GH (1994) Genetic relationship and characterization of Persian walnut (*Juglans regia* L.) cultivars using restriction fragment length polymorphisms (RFLPs). J Am Soc Hortic Sci 119(4):833–839

Foroni I, Rao R, Woeste K, Gallitelli M (2005) Characterization of *Juglans regia* L. through SSR markers and evaluation of genetic relationships among cultivars and the 'Sorrento' landrace. J Hortic Sci Biotech 80:49–53

Gemas VJV, Almadanim MC, Tenreiro R, Martins A, Fevereiro P (2004) Genetic diversity in the olive tree (*Olea europea* L. subsp. *europea*) cultivated in Portugal revealed by RAPD and ISSR markers. Genet Resour Crop Evol 51:501–511

Glaubitz JC. (2003). Convert: (version 1.2): A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Access date: May 2005 http://www.agriculture.purdue.edu/fnr/htlm/faculty/rhodes/student%20and%20Staff/glaubitz/software.htm

Gomez OJ, Blair MW, Frankow-Lindberg BE, Gullberg U (2004) Molecular and phenotypic diversity of common bean landraces from Nicaragua. Crop Sci 44:1412–1418

Hamza S, Hamida WB, Rebai A, Harrabi M (2004) SSR-based genetic diversity among Tunisian winter barley and relationship with morphological traits. Euphytica 135:107–118

Harlan JR (1975) Our vanishing genetic resources. Science 188:618–621

He C, Poysa V, Yu K (2003) Development and characterization of simple sequenze repeat (SSR) markers and their use in determining reletionships among *Lycopersicon esculentum* cultivars. Theor Appl Genet 106:363–373

Huntley B, Birks HJB (1983) An atlas of past and present pollen maps for Europe: 0–13,000 years ago. Cambridge University Press, NY

Komarow VL (1936). Flora of the US.S.R. Moscow. In: Komarow VL edition (ed) Botanical Institute of the Academy of Science of the U.S.S.R (Translated by Israel Program for Scientific Translations), vol 5, pp 197–199

Marangoni B, Malaguti D, Toselli M. and Lucchi A (2004) Primi risultati della ricerca sul noce da frutto in Emilia Romagna. Relazione—Macfruit 2004

McGranahan G, Leslie C (1990) Walnuts (Juglans) In: Moore JN, Ballington JR (eds) Genetic resources of temperate fruit and nut crops. Acta Hortic 290:907–974

Meerow AW, Wisser RJ, Brown JS, Kuhn DN, Schnell RJ, Broschat TK (2003) Analysis of genetic diversity and population structure within Florida coconut (*Cocos nucifera* L.) germplasm using microsatellite DNA, with special emphasis on the Fiji Dwarf cultivar. Theor Appl Genet 106:715–726

Minch E, Ruitz-Linares A, Goldstein D, Feldman M, Kidd JR, Cavalli-Sforza LL (1997) Microsat 1.5: a computer program for calculating various statistics on microsatellite allele data. Access date: May 2005 http://www.hpgl.stanford.edu/projects/microsat/.

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nekrassowa VL (1927) The genus *Juglans* in Turkestan. Bull App Bot Plant Breed 18:303–360

Nicese FP, Hormaza JI, McGranahan GH (1998) Molecular characterization and genetic relatedness among walnut (*Juglans regia* L.) genotypes based on RAPD markers. Euphytica 101:199–206

Pingali P, Smale M (2001) Agriculture, industrialized. In: Levin SA (ed) Encyclopedia of Biodiversity. Academic Press, San Diego, pp 85–97

Portis E, Acquadro A, Comino C, Lanteri S (2004) Effect of farmers' seed selection on genetic variation of a landrace population of pepper (*Caspicum annum* L.), grown in North-West Italy. Genet Resour Crop Evol 51:581–590

Potter D, Fangyou G, Aiello G, Leslie G, McGranahan G (2002) Intersimple sequences repeat markers for fingerprinting and determining genetic relationships of walnut (*Juglans regia*) cultivars. J Am Soc Hortic Sci 127(1):75–81

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

Regione Campania—Assessorato all'Agricoltura. Access date: August 2004 http://www.sito.regione.Campania.it/agricoltura/home.htm

Solar A, Smole J, Stampar F, Virscek-Marn M (1994). Characterization of isozyme variation in walnut (*Juglans regia* L.). Euphytica 77:105–112

Tuleke W and McGranahan GH (1994). The walnut germplasm collection at the University of California, Davis: a description of the collection and a history of the breeding program of Eugene''. F. Serr and HI, Forde. Rpt. 13 Univ. Calif. Genetic Resources Conservation Program, Davis

Woeste K, Burns R, Rhodes O, Michler C (2002) Thirty polymorphic nuclear microsatellite loci from black walnut. J Hered 93:58–60

Yamamoto T, Kimura T, Sawamura Y, Manabe T, Kotobuki K, Hayashi T, Ban Y, Matsuta N (2002) Simple sequence repeats for genetic analysis in pear. Euphytica 124:129–137