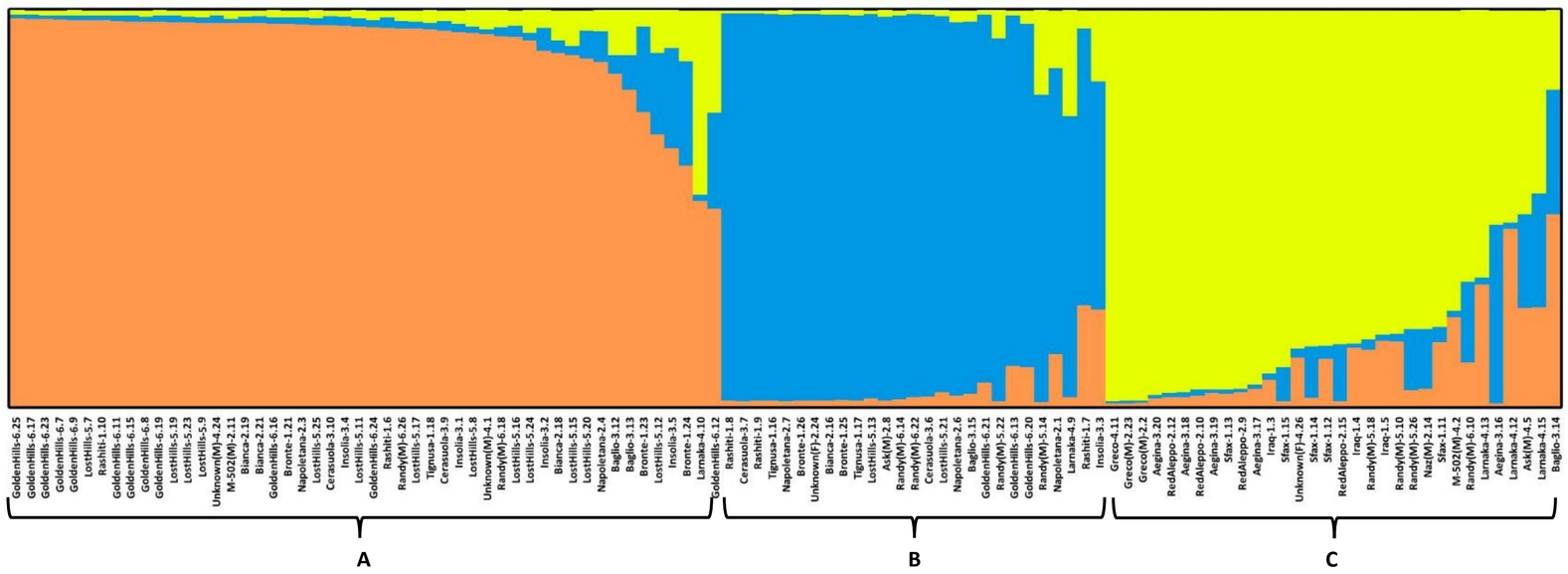
8.23 - PISTACIA GERMPLASM COLLECTION OF THE NATIONAL FRUIT GERMPLASM CENTRE (NFGC) MANAGEMENT THROUGH MICROSATELLITE MARKERS

SANTILLO N.[1], MUSKAJ A. [1], SBROCCA I. [1], VERDE I. [1], MICALI S. [1], VENDRAMIN E. [1]
[1] CREA - Research centre for Olive, Fruit and Citrus Crops

BACKGROUND

The Research Centre for Olive, Fruit and Citrus Crop in Rome hosts the National Fruit Germplasm Centre (NFGC), which preserves the largest Italian fruit germplasm collection, comprising over 5,000 accessions from 40 fruit species and their wild relatives. The Pistacia genus (Anacardiaceae family) includes at least eleven dioecious species (Zohary, 1952). Among them, only *P. vera* L. has commercial importance for its edible nuts, while others are mainly used as rootstocks (*P. integerrima*, *P. mutica*, *P. terebinthus*, and *P. atlantica*) or as pollinators (*P. terebinthus*). The NFGC Pistacia collection consists of 120 accessions: 24 *P. vera*, 13 *P. integerrima* and four *P. terebinthus*. This study aims to assess genetic diversity within the collection and to clarify synonymies and homonymies. A set of 27 polymorphic EST-SSR markers, transferable across *Pistacia* species, was used to characterize *P. vera* and related taxa, including *P. integerrima*, *P. terebinthus* and interspecific hybrids.

Figure 2. Population genetic structure of *Pistacia* accessions at **K = 3**. Each bar represents one individual and colors indicate cluster membership.



MATERIALS AND METHODS

Genomic DNA of 140 accessions was extracted from young apical shoots using the GENEzol DNA Reagent Plant (Geneaid) with a modified protocol. DNA quality and integrity were assessed by gel electrophoresis and spectrophotometer.

Ninety nine EST-SSR loci from the EPV series were amplified using PCR conditions as described in Vendramin et al. (2010). PCR products ranging from 100 nt to 350 nt were visualized using capillary electrophoresis (CEQ 8800, Beckman). Larger fragments (up to 900 bp) were assessed on high-resolution MetaPhor agarose gel (Lonza).

After filtering for samples with missing data > 20%, the phylogenetic tree was built using Past 5.0.2 (Hammer, Ø. Et al., 2001). A similarity matrix was constructed,, according to Dice index. Dendrograms were obtained by the NJ method.

The number of alleles per locus (Na), the number of rare (frequency <0.05) and private alleles (specific to a genotype), the observed and expected heterozygosity (Ho and He), the polymorphic information content (PIC) and the discrimination power (DP) of each marker were estimated by iMEC (Amiryousefi et al., 2018).

Population structure analysis, performed on 109 *P. vera* genotypes and based on Bayesian statistics using 27 EST-SSR markers, was obtained using Structure 2.3.4 (Pritchard et al., 2000), considering a K number from 1 to 10, with 10 iterations for each value of K. The settings for burning-in and MCMC (Markov Chain Monte Carlo) were 10,000 and 100,000, respectively. CLUMPAK (Kopelman et al., 2015) was used to find the optimal alignment of the ten independent replications. DISTRUCT 1.1 (Rosenberg, 2004) was used to graphically display the population structure.

RESULTS

Twenty-seven out of 50 EST-SSR markers showed a polymorphic profile (Table 1). In total 91 alleles were amplified and the allele numbers ranged from 2 to 6. The 28.2% is classified as rare alleles and 24 as private. The most part of private alleles are species-specific.

The average number per locus is equal to 3.4, higher than reported in Vendramin et al. (2010). The observed heterozygosity (Ho) varied between 0.01 and 0.91 for EPVM033 and EPVM063, respectively, with an average of 0.36 per locus. The most informative markers were EPVM002, EPVM022, EPVM049, EPVM050 and EPVM054 with a PIC value equal or above 0.50, while the less informative were EPVF013, EPVF018, EPVF030 and EPVM016 (PIC = 0.43). The low level of information catched by the markers was expected in since they derived from the transcribed regions known to be less variable. The discrimination power (DP) showed an average of 0.33. EPVM056 marker showed the highest discrimination power (PD = 0.89).

Table 1. Information about the 27 EST-SSR primer pairs

Marker	Range	Ho	N. Alleles	He	PIC	D	MAF <0,05
EPVF010	208-216	0,03	3	0,46	0,46	0,03	2
EPVF013	570-670	0,64	3	0,5	0,43	0,7	1
EPVF018	700-720	0,06	2	0,5	0,43	0,75	--
EPVF021	219-228	0,1	3	0,56	0,48	0,23	1
EPVF023a	191-205	0,17	4	0,41	0,48	0,37	--
EPVF023b	217-245	0,11	4	0,42	0,48	0,08	2
EPVF030	600-610	0,04	2	0,52	0,43	0,04	1
EPVF032	610-680	0,88	3	0,47	0,44	0,61	1
EPVM002	635-645	0,79	2	0,21	0,53	0,19	--
EPVM016	490-500	0,56	3	0,5	0,43	0,72	--
EPVM017	245-249	0,04	2	0,64	0,47	0,33	--
EPVM022	238-250	0,09	5	0,37	0,5	0,09	2
EPVM024	227-248	0,03	5	0,53	0,49	0,34	4
EPVM032	220-245	0,75	5	0,52	0,47	0,31	2
EPVM033	302-304	0,01	2	0,67	0,48	0,45	1
EPVM035	140-143	0,19	2	0,53	0,46	0,26	--
EPVM040	107-116	0,07	2	0,61	0,47	0,3	--
EPVM041	307-330	0,07	4	0,44	0,49	0,11	3
EPVM043	940-950	0,39	2	0,49	0,49	0,34	--
EPVM049	249-291	0,29	6	0,38	0,5	0,13	3
EPVM050	920-940	0,72	2	0,24	0,53	0,21	--
EPVM051	110-149	0,76	4	0,56	0,46	0,34	1
EPVM054	900-950	0,58	2	0,38	0,51	0,32	--
EPVM056	405-425	0,59	5	0,44	0,45	0,89	1
EPVM058	245-275	0,05	4	0,49	0,49	0,19	2
EPVM059	220-245	0,66	5	0,57	0,48	0,39	2
EPVM063	157-190	0,91	5	0,67	0,49	0,19	2

The NJ tree (Figure 1) clearly separates all the 140 accessions, with the more genetically different being the *P. integerrima* and *P. terebinthus* genotypes, as expected. The accessions with the same name don't belong to the same cluster, this is probably due to some labelling or propagating errors common for germplasm collections. Also, the history of the collection, planted 30 years ago, can explain this result in fact, some genotypes were introduced by seed instead of propagation. Furthermore, 'Bianca/Napoletana' cultivar is defined as cultivar population and for this reason internal genetic variability is expected.

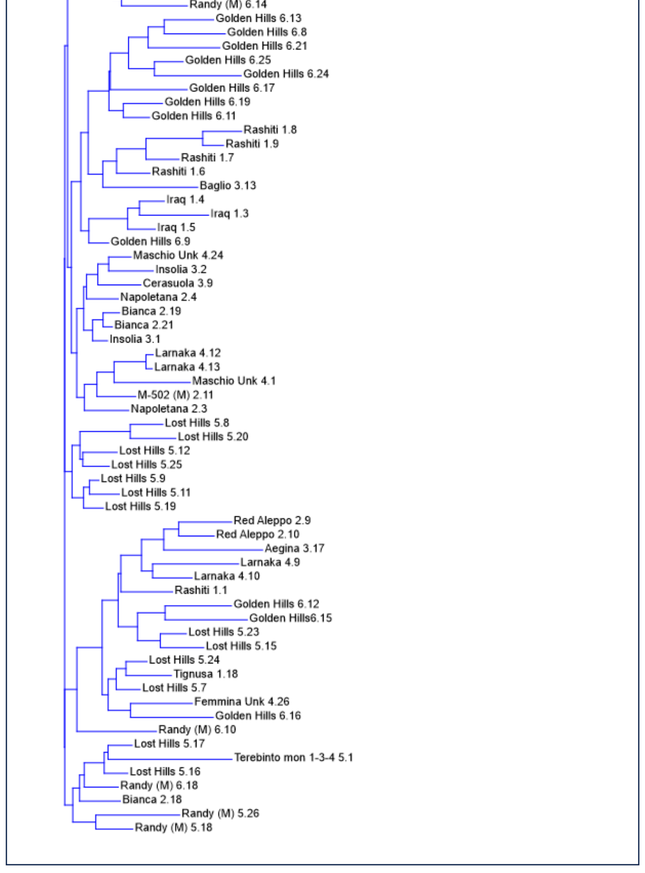


Figure 1. Phylogenetic tree of the 140 *Pistacia* accessions

The *P. vera* population structure analysis revealed three subpopulations (K = 3) (Figure 2). Considering the membership coefficient Q ≥ 0.80, the 109 samples are clustered into three main subpopulations (PA, PB, PC). The PA group (orange) is composed mainly of American cultivars ('Golden Hills', 'Lost Hills', 'Randy') with some Italian traditional genotypes ('Bianca/Napoletana', 'Baglio', 'Bronte', and 'Insolia'). This is in accordance with the Pistacia breeding history, in fact in the 1960 some Italian cultivars were introduced to north California and used for several breeding programs enhancing overall quality. The sPB (blue) grouped traditional Italian cultivars ('Bianca/Napoletana', 'Insolia', 'Bronte', 'Tignusa', 'Baglio', 'Cerasuola') all known to be indehiscent, with small fruits and the typical deep green kernels. The sPC (yellow) was mainly composed of Mediterranean and Middle Eastern cultivars ('Greco', 'Red Aleppo', 'Aegina', 'Sfax', 'Iraq', 'Larnaka', and 'Ask').

SUMMARY AND PROSPECTS

The molecular genetic characterization of 140 *Pistacia* accessions allowed to estimate the genetic diversity and to build a fingerprint database of the NGFC CREA collection. The genetic characterization clearly identified different species belonging to Pistacia genus and underlined ambiguous situations that require further study to preserve the higher level of genetic diversity and to rationalize the collection.

Acknowledgments: We thank S. Vona for his valuable contribution in field tree management and sample collection.

Progetto "Azioni di VALOrizzazione e REcupero per le filiere Italiane di Nocciolo, CAstagno, Mandorlo,Pistacchio e carrubO - VALO.RE I.N. CA.M.P.O" - Finanziato dal MASAF D.D. N. 667521 del 30.12.2022