




Genotypic diversity and population structure of the apricot landraces of the Campania region (Southern Italy) based on fluorescent SSRs

Boris Basile · Rosa Rao ·
Giandomenico Corrado 

Received: 2 March 2022 / Accepted: 20 May 2022 / Published online: 9 June 2022
© The Author(s) 2022

Abstract Apricot (*Prunus armeniaca* L.) is a cosmopolitan, major stone fruit species cultivated for its delicious drupes. Strong concerns about loss of diversity in apricot have been recently put forward mainly because of the enduring transition from local to globalized varieties more suitable for intensive agriculture. Italy, one of the top producing countries in the world, has arguably one of the richest apricot germplasms, with the Campania region being a traditional and leading area of cultivation since Roman times. This work aimed to study the molecular diversity of an ex-situ collection of local apricot varieties. They once flourished in conventional cropping systems, but are nowadays neglected, surviving in family farms, and sold only in local markets. We investigated the level of genetic and genotypic diversity, assessed clonality in the population, and verified the presence of genetic structure and differentiation in 73 apricot varieties by using fluorescent SSR markers resolved in capillary electrophoresis (SSR-CE). The analysis indicated a high level of genetic diversity (with little locus-specific differences) and of genotypic assortment, with few possible synonyms. Moreover, the

investigation of relatedness among apricot landraces and the negligible within-landraces differentiation are consistent with a common origin of the germplasm, a likely result of an adaptive diversification that is mainly due to similar environmental and agricultural factors. This work highlights the value of the traditional germplasm where crops thrived. Moreover, it provides support for studying the possible mechanisms of evolution and adaptation and for implementing measures for the exploitation, regeneration, and on-farm promotion of the apricot landraces.

Keywords Differentiation · DNA markers · Population genetics · *Prunus armeniaca*

Introduction

Apricot (*Prunus armeniaca* L.) is a deciduous stone fruit tree that is globally cultivated in temperate climate zones, especially around the Mediterranean basin and in the Middle East (Zhebentyayeva et al. 2012). Nowadays fruits are most often eaten fresh, and the home-made production of dried, canned, or jammed apricots has strongly declined in Western countries (Khursheed et al. 2020). Main industrial uses are for juices, puree, jam, drinks, and as flavoring agent (e.g., confectionary, ice creams, and dairy) (Roussos et al. 2016). Despite this gastronomic versatility, the sweet, savory, and vividly colored fruits have limited shelf life and are largely seasonal, two

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10722-022-01412-8>.

B. Basile · R. Rao · G. Corrado (✉)
Dipartimento di Agraria, Università degli Studi di Napoli
Federico II, Portici, Naples, Italy
e-mail: giandomenico.corrado@unina.it

main factors that negatively impact consumer choice and food industry, respectively.

Italy is one of the largest producers of apricots in the world (Khursheed et al. 2020), with the Campania region (Southern Italy) providing approximately one-third of the national production (Istituto Nazionale di Statistica; <https://www.istat.it>). In the last decades, the introduction of cultivars of foreign origin characterized a significant varietal renewal, further promoting the abandonment of traditional varieties (Biscotti et al. 2022; Corrado et al. 2021). The main drivers of such change were the market appreciation of strongly (over-)colored fruits, the preference for varieties specifically bred for fresh consumption, as well as resistance to stress and a ripening time that allows growers to extend the harvest season (Mennone 2016). Recent studies highlighted that germplasm erosion and loss of genetic diversity in the contemporary apricot germplasm are worldwide concerns and prompted for action to preserve traditional varieties and local landraces also for possible applied exploitation (Rossarolla et al. 2020; Herrera et al. 2021; Hagen et al. 2002). For instance, Italy has one of the largest availability of apricot landraces (Ledbetter 2008) and, despite the known limitations related to the fruit perishability, they should be considered a source of genetic variation for adaptive traits, bearing in mind the specific needs of the Italian production systems (Corrado and Rao 2017; Hormaza et al. 2007; Bartolini et al. 2020; Mennone 2016). Moreover, it is debated whether the quality standards of early blooming and maturing contemporary cultivars (mainly focused on the fruit shape, size, symmetry, and color) may have inadvertently caused an additional reason for the commercial decline of this crop, since Italian consumers were more accustomed to sweeter and less sour fruits (Sansavini 2019). Finally, the promotion of this germplasm is also a means to sustain local economy and to improve the management of peri-urban and rural landscape (Zimmerer et al. 2015; Corrado et al. 2021).

To protect the Campanian apricot germplasm, an ex-situ repository has been established based on folk taxonomy and local knowledge yet, this material has been little exploited (Banca Regionale del Germoplasma, Regione Campania; <http://www.agricoltura.regione.campania.it/>). The molecular characterization of the diversity is essential for managing germplasm (Sheikh et al. 2021; Gürcan et al. 2019), assessing

clonal relationships (Fossati et al. 2005), and identifying duplicated accessions and mislabeling (e.g., synonyms, homonyms) (Potts et al. 2012). Moreover, research on the population structure is important for the definition of conservation units, essential in the more demanding management of landraces that cannot be propagated by seeds. Under these perspectives, microsatellites, commonly known as Simple Sequence Repeats (SSRs) or Short Tandem Repeats (STRs), are the one of the markers of choice in plant science because they are codominant, multiallelic and usually, highly polymorphic (Varshney et al. 2005). SSRs are also appropriate to detect recent demographic events and population-specific alleles, thus being particularly suitable to reveal population structure in local populations (Tsykun et al. 2017). The informativeness and utility of SSRs have been considerably enhanced by the diffusion of capillary electrophoretic (CE) systems (Butler 2009). For germplasm management, and more generally forensics, CE has not only increased throughput and multiplexing ability, but greatly limited inter-laboratory variability, thus allowing to retrieve highly reproducible, robust, and transferable information, crucial to build databases of molecular profiles (Baric et al. 2008; Ordidge et al. 2021).

In this work, we analyzed an ex-situ germplasm collection of seventy-three apricot landrace varieties, collected in the Campania region and cultivated also in Southern Italy. By using fluorescent SSR markers resolved in capillary electrophoresis (SSR-CE), we aimed to investigate the level of molecular diversity and verify the presence of a possible genetic structure. Specifically, we provide an in-depth description of the relatedness among apricot landraces as the first step to facilitate their agronomic characterization, ex-situ management, and possibly, promotion in local markets.

Material and methods

Plant material

This work was carried out on seventy-three landrace varieties of *Prunus armeniaca* L., namely (know synonyms are in square brackets; abbreviations used in this article in round brackets): 'Abate' (ABA), 'Abatone' (ABT), 'Ananassa' (ANA), 'Antonaniello'

(ANT), 'Aronzo' (ARO), 'Baracca' (BAR), 'Boccuccia' (BOC), 'Boccuccia bianca' (BCA), 'Boccuccia di Eboli' (BCE), 'Boccuccia grossa' (BCG), 'Boccuccia liscia ii' (BCL), 'Cafona' (CAF), 'Campana' (CAM), 'Cardinale' (CAD), 'Carpona' (CAR), 'Casino' (CAS), 'Cerasiello' (CEI), 'Cerasona' (CEO), 'Cristiana' (CRI), 'Diavola' (DIA), 'Don Aniello' (DON), 'Ebolitana' (EBO), 'Fracasso' (FRA), 'Fronne fresche' (FRO), 'Grangicana' (GRA), 'Lisandrina' (LIS), 'Macona' (MAC), 'Magnalona' (MAG), 'Mammana' (MAM), 'Montodoro' (MND), 'Monteruscello' (MNR), 'Nonno' (NON), 'Ottavianese' (OTT), 'Palumella' (PAL), 'Palumella ii' (PAM), 'Panzona' (PAN), 'Paolona' (PAO), 'Pazza' (PAZ), 'Pelese Correale [Pelese]' (PEC), 'Pelese di Giovanniello' (PEG), 'Persechella' (PER), 'Piciona' (PIC), 'Portuallara' (POR), 'Presidente' (PRE), 'Prevetone' (PRV), 'Puscia' (PUS), 'Puzo' (PUZ), 'Resina' (RES), 'S. Francesco' (SAF), 'S. Giorgio' (SAG), 'Sant'Antonio' (SAN), 'Scassulillo' (SCA), 'Scassulillo grande' (SCG), 'Secquagliella II' (SCE), 'Schiavona' (SCH), 'Scialo'' (SCI), 'Secondina' (SEC), 'Setacciara' (SET), 'Signora' (SIG), 'Silvana' (SIL), 'Sonacampana' (SON), 'Sorrentino' (SOR), 'Stella' (STE), 'Stradona' (STR), 'Taviello' (TAV), 'Tre [Tre Palle]' (TRE), 'Vicario' (VCA), 'Vicenzo [Vicenzo 'e Marial]' (VCI), 'Zeppa [Zeppa 'e sisco]' (ZEP), 'Zeppona' (ZPO), 'Zi' Francesco' (ZIF), 'Zi' Luisa' (ZIL), and 'Zi' Ramunno' (ZIR). Adult trees belong to the collection of the Azienda Agricola Sperimentale Regionale 'Improsta' (Centro per la Ricerca Applicata in Agricoltura, Regione Campania), located in Eboli (SA, Italy).

DNA isolation and fluorescent SSR-capillary electrophoresis (SSR-CE)

Five young, healthy looking leaves per plant were harvested and immediately frozen in liquid nitrogen. We analyzed two trees per landrace. Leaves were stored at -80°C until analysis. Leaves were finely ground in liquid nitrogen and DNA was isolated as previously described (Corrado et al. 2021). DNA fingerprinting was performed using eight highly polymorphic apricot SSR loci (AMPA095, AMPA112, UDAp401, UDAp410, UDAp 414, UDAp415, UDAp420, and UDAp446), selected from the literature (Hagen et al. 2004) (Messina et al. 2004; Rao et al. 2010). Primer sequences and main features are reported in

Supplementary Table 1. Reactions were assembled in a final volume of 25 μL using as template 100 ng of genomic DNA as estimated in an agarose gel electrophoresis (Sambrook et al. 1989). The thermal profile of the PCR and the primer specific annealing temperatures are reported in Supplementary Table 1. The success of the amplification was first checked by agarose gel-electrophoresis (Sambrook et al. 1989), while allelic discrimination was carried out by fluorescence-based capillary electrophoresis on an ABI PRISM 3130 Avant Genetic Analyzer (Thermo Fisher Scientific, Milan, Italy) as already described (Verdone et al. 2018). Automated fragment data analysis was carried out with the GeneScan 4.3 software (Thermo Fisher Scientific) on the basis of the sizeable peaks of the Gene Scan 500 Liz-dye internal standard (Thermo Fisher Scientific). Manual binning was independently performed on each SSR locus to minimize the mean offset of allelic sizes within the instrument resolution (± 1 bp).

Data analysis

For the analysis of locus-based indices of genotypic diversity we calculated per each SSR locus: the allelic size range (ASR) in bp; the number of alleles (N_a); the number of MultiLocus Genotypes (MLG); the Effective number of alleles (N_e) as $1/(\sum p_i^2)$; the Shannon Index of Diversity (I) as $-1 \times \sum (p_i \times \ln(p_i))$; the Evenness (E) as $((1/I)-1)/((e^1) - 1)$; the Observed Heterozygosity; the Polymorphic Index Content (PIC, also known as gene diversity) as $1 - \sum p_i^2$; and the Wright's Fixation Index (F) as $(H_e - H_o)/H_e$, where for each locus, p_i is the frequency of the i -th allele, $\sum p_i^2$ is the sum of the squared population allele frequencies, and $1/I$ is the Stoddart and Taylor's index. The significance of bivariate correlations was assessed using the Pearson correlation coefficient. The index of association (Ia) and the modified scaled (ranging from 0 to 1) measure r_d were calculated to detect signatures of sexual reproduction as described (Brown et al. 1980; Agapow and Burt 2001). Missing data were ignored and data resampling for statistical testing were performed with permutations over alleles ($n=999$). Pairwise resemblance between varieties were calculated with the Prevosti's absolute genetic distance and the dendrogram was built using the unweighted pair group method with arithmetic mean (UPGMA) algorithm (Prevosti et al. 1975). These

calculations were carried out with Genalex and poppr (Smouse and Peakall 2012; Kamvar et al. 2014).

The occurrence of a structured population was evaluated utilizing the model-based Bayesian procedure implemented in the software Structure v2.3 (Pritchard et al. 2000). The analysis was performed using a burning period of 50,000 iterations and a run length of 250,000 MCMC replications. We tested a continuous series of K_s , from 1 to 11 in ten independent runs, without introducing prior knowledge about the population, and assuming correlated allele frequencies and admixture (Falush et al. 2003). The most informative K was identified using the so-called Evanno's method (DeltaK), based on the rate of change in the log probability of data between successive K values (Evanno et al. 2005). The estimated cluster membership coefficient matrices of the ten runs were permuted so that all replicates have the closest match possible and then averaged across replicates using the Greedy algorithm of the software CLUMMP with 9999 permutations (Jakobsson and Rosenberg 2007). To statistically validate the estimated populations, we calculated pairwise F_{st} and Nei's standard genetic distance (D_{st}) between populations using MSA (Dieringer and Schlotterer 2003). The reference distribution for p value calculation of the F_{st} analysis was based on 9999 permutations.

Result

Genetic and genotypic diversity in the apricot landraces

The SSR fingerprinting of the germplasm collection indicated that all the loci were polymorphic, the maximum ploidy for each sample was two, and that the allelic size range was consistent with the values reported in the literature (Table 1) (Hagen et al. 2004; Messina et al. 2004; Rao et al. 2010). The SSR output of the two trees of the same landrace was always identical, and therefore a single profile per variety was considered for subsequent analyses. The number of alleles greatly varied among loci, with a coefficient of variation (CV) of 27.5%. Specifically, the number of alleles ranged from twelve (UDAp446) to five (AMPA095). The most diverse locus according to the Effective number of alleles (i.e., the number of alleles weighted for their frequencies) was UDAp401 (with eight alleles), while the UDAp446 ranked in the bottom half. UDAp401 was also the most informative locus considering the PIC, although differences among loci were limited (CV: 12.1%). On the other hand, the number of alleles had a large positive correlation ($r=0.79$; $p=0.02$) with the number of detected multilocus genotypes (MLG). This index ranged from 18 for AMPA112, AMPA095 and UDAp415, to eight for UDAp446 UDAp414, and it was the most variable (CV: 31.6%) among the calculated indices of genetic diversity. The Simpson's index of diversity (1-D) also largely varied across loci (CV = 18.4%) yet it was higher than one for every locus. A slightly lower variation across loci was present for

Table 1 Main indices of molecular diversity of the apricot collection

Locus	ASR	Na	MLG	Ne	E	I	Ho	PIC	F
AMPA095	77–99	5	11	2.72	0.79	1.15	0.72	0.63	– 0.14
AMPA112	194–236	10	18	4.66	0.76	1.76	0.63	0.79	0.20
UDAp401	169–207	8	16	4.89	0.83	1.74	0.81	0.80	– 0.01
UDAp410	116–152	8	17	4.08	0.72	1.67	0.82	0.75	– 0.09
UDAp414	146–174	7	8	2.85	0.76	1.23	0.88	0.65	– 0.35
UDAp415	152–174	6	8	2.25	0.63	1.09	0.66	0.56	– 0.19
UDAp420	164–184	9	18	3.78	0.69	1.61	0.67	0.74	0.09
UDAp446	130–168	12	18	2.91	0.52	1.54	0.67	0.66	– 0.02
Mean		8.1	14.3	3.52	0.71	1.47	0.73	0.70	– 0.06
SE		0.79	1.59	0.34	0.04	0.10	0.03	0.03	0.06

the Evenness (CV: 13.9), a measure related to the ratio between the more abundant and the rarer genotypes. This index was on average high (0.71 ± 0.04) and less variable also than the number of alleles per locus. The moderate negative linear correlation ($r = -0.52$; Pearson's Correlation) between the number of alleles and their Evenness was not significant ($p = 0.18$). The observed heterozygosity (H_o) was high (0.73 ± 0.03) and little varied among loci (CV = 12.3%). As for the Evenness, H_o did not significantly associate with the number of alleles ($r = -0.34$; $p = 0.40$) and neither with the number of MLG ($r = -0.31$; $p = 0.45$). Overall, the indices of genetic diversity are consistent with the detection, by highly polymorphic loci, of a non-adaptive genetic diversity, possibly distributed through clonal selection. Consistently, the Fixation Index was close to zero for all loci (mean \pm s.e.: -0.07 ± 0.06), except for UDAp414 (-0.35), suggesting a limited heterotic selection or negative assortative mating in our population, despite the high H_o values.

Considering that the high number of genotypes was collected in a relatively small area, and the possible on-farm clonal propagation and exchange of fruit tree landraces, we tested if the population was partially or predominantly clonal, that is with a considerable disequilibrium among loci due to linkage. To this aim we calculated the index of association (I_a) and the r_d , a related index weighted for the number of loci, which are used to detect signatures of sexual reproduction. The analysis indicated a I_a of 0.58 and a p value lower than 0.01 for a r_d of 0.084. The latter falls well outside of the calculated

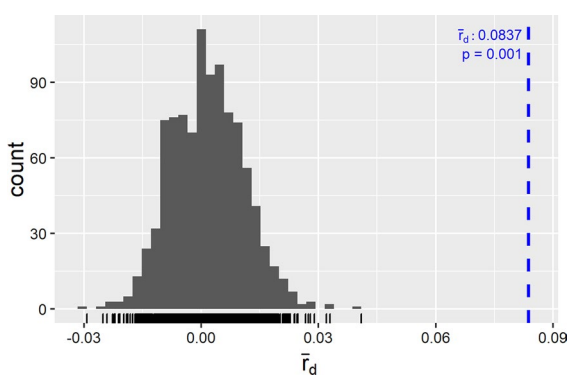


Fig. 1 Histograms of resampled r_d values ($n = 999$). The overlaid dashed vertical blue line indicates the observed r_d . The ticks at the bottom represent individual observations

distribution that is expected under no linkage (Fig. 1), indicating that the population under investigation is predominantly of sexual origin.

The molecular profiles were used to calculate pairwise genetic distances, which were then used to build an UPGMA dendrogram (Fig. 2).

This analysis illustrated that some varieties had the same genetic profile. The identities were (in parenthesis is reported the cumulative product of the genotype probability calculated over all the SSR loci; RMP): DIA and PAL (2.12×10^{-10}), BCA and

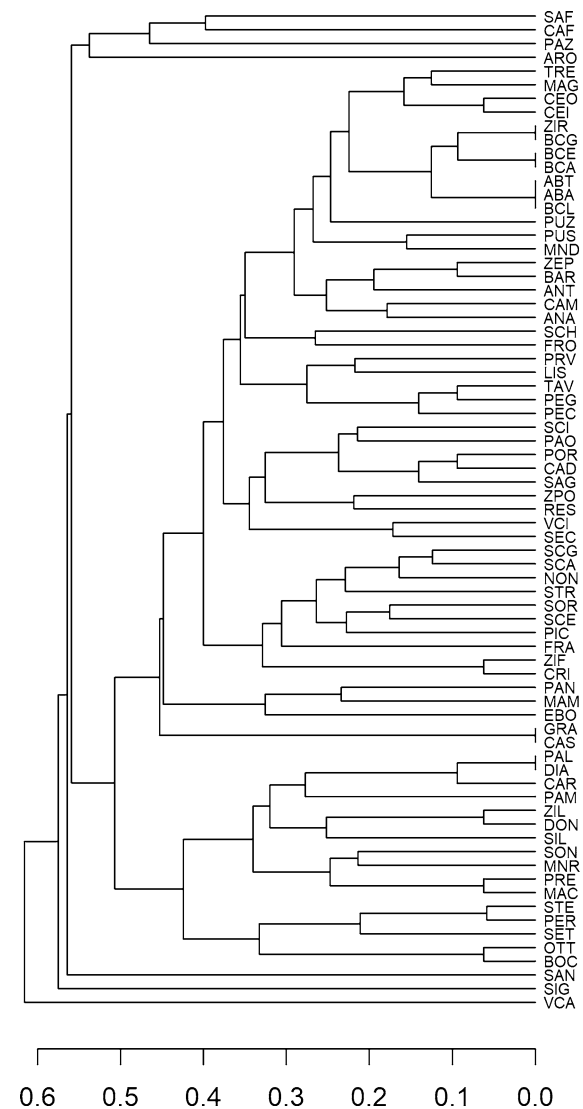


Fig. 2 Dendrogram of the apricot landrace varieties. Pairwise genetic distances were calculated with the Provetti's coefficient. The agglomerative clustering method was UPGMA

BCE (6.86×10^{-7}), BCG and ZIR (3.66×10^{-6}), CAS and GRA (2.44×10^{-11}), and the three samples ABA, ABT and BCL (1.23×10^{-06}). In most of the cases, the vernacular classification supports the presence of derived names from a possible landrace group, such as for the cluster that comprises the identical profiles ‘Boccuccia di Eboli’ (BCE) and ‘Boccuccia’ (BCA), and the ‘Abate’ (ABA) e ‘Abatone’ (ABT). In the other cases, it may be possible the presence of mislabeling or erroneous denominations (e.g., synonyms), also considering the low value of the genotype probability.

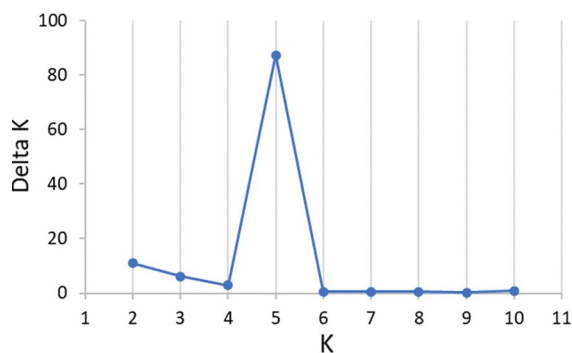


Fig. 3 Estimation of the optimum number of clusters in the apricot germplasm based on the second order rate of change of the conditional posterior probability of the simulation, also known as the Evanno’s test. The graph displays the Delta K for each of the K value tested

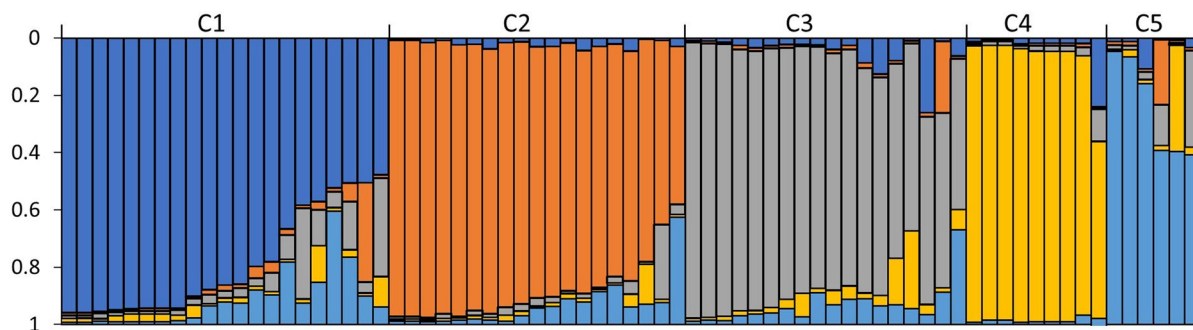


Fig. 4 Estimated population structure of the apricot varieties. Each variety is represented by a vertical line, which is partitioned into colored segments that represent the estimated

Analysis of the population structure and differentiation

In absence of an a priori classification, the identification of genetically similar groups of apricot landraces was performed using a widely employed admixture model-based clustering method that also allows proportional assignment to multiple populations. The most informative number of subpopulations (K) was five according to the second order rate of change of the estimate of the conditional posterior probability of the simulation (Fig. 3).

The inferred population structure for K=5 is presented in Fig. 4 and the CLUMPP generated Q-matrix is reported in Supplementary Table 2.

A high proportion of genotypes (68%) had a membership coefficient higher than 0.8 and overall, most varieties were strongly assigned to subpopulations (Supplementary Fig. 1), suggesting also a reduced genetic admixture.

We tested if the groups inferred by the population structure analysis represent statistically significant subpopulations considering pairwise measures of two widely used estimates of differentiation, F_{st} and the Nei’ standard genetic distance (D_{st}). The analysis indicated that the genetic divergence between the identified sub-populations was low and for instance, the maximum value was below 0.1 (Table 2). Only in two cases (C1 vs C4; C2 vs C4) the genetic differentiation based on the F_{st} calculation was statistically significant. Similarly, the genetic distances between sub-populations were on average very low (Table 2).

membership fractions in the five clusters (C). See Fig. 3 for the determination of the optimal number of cluster and Supplementary Table 2 for the Q-matrix

Table 2 Estimation of pairwise genetic differentiation and genetic distance among the five sup-populations (from C1 to C5, see Supplementary Table 2 for the list of varieties in each sub-population) as inferred by the Bayesian analysis implemented in the structure software

	C1	C2	C3	C4	C5
C1		0.0244	0.0099	0.0720	<0.0001
C2	0.0050		0.0132	0.1345	0.0507
C3	0.0017	<0.0001		0.0686	<0.0001
C4	0.0496*	0.0754***	0.0534		0.0570
C5	<0.0001	0.0198	<0.0001	0.0431	

Below the diagonal is the pairwise estimate of F_{st} , while D_{st} appears above the diagonal

The p value for the estimated F_{st} values (below diagonal) was calculated using 9999 permutations (***: $p < 0.01$; *: $p < 0.05$; absence of asterisk: not significant, $p \geq 0.05$)

Discussion

The ever-expanding consensus to preserve the genetic diversity of food plant varieties needs to be supported and guided by its characterization, especially in areas where landraces have been produced and survived in agriculture. Although Italy has a rich apricot germplasm (Ledbetter 2008), reports on genetic resources have rarely considered local realities in this country. The molecular analysis of the ex-situ collection indicated a high level of diversity considering both the total number of alleles in the population and the number of alleles per locus. Although these estimates are biased by the sample size, these values were within or above the range of other reports (Wang et al. 2011; Hagen et al. 2004; Messina et al. 2004; Li et al. 2018), which also included landraces (Sheikh et al. 2021; Junhuan et al. 2012; Lamia et al. 2010). Microsatellites were all highly polymorphic and informative (e.g., PIC values higher than 0.5), confirming the features of the selected SSRs (Hagen et al. 2004; Messina et al. 2004). The level of the observed heterozygosity was also high, a likely consequence of the agamic propagation of the germplasm. On the other hand, we could not distinguish all the varieties. Specifically, the number of multilocus genotypes was slightly lower than the number of landraces, as predicted from the analysis of a curated ex-situ collection. For example, in a study of a natural population of *P. avium*, the percentage of MLG was lower (i.e., 30%) (Jarni et al. 2015). In some instances, it could be

proposed the presence of synonyms or derived clones taking into consideration the folk names. Moreover, the observed genetic similarity could suggest the presence of a landrace group (e.g., for the ‘Boccuccia’ types) consisting of genetically similar types (Zeven 1998). These hypotheses should be tested by a detailed morphological analysis, also considering the limitations of the vernacular names (Wilkie and Saridan 1999). Nonetheless, our previous experience on apricot indicated that the number of unique profiles identified with microsatellites is larger than those obtained from qualitative morphological characters (Corrado et al. 2021). Moreover, although the very low values of the RMP, we cannot exclude that more in-depth DNA investigations may reveal adaptive or morphologically significant polymorphisms among landrace groups. However, for other varieties, the data favored the presence of erroneous denominations or sampling. It should also be added that other characterizations of traditional germplasm have also revealed cases of synonymy and/or duplicated accessions (Zhebentyayeva et al. 2003; Ispizua et al. 2007; Queiroz et al. 2015).

At least for apricots of the Campania region, grey literature reports that part of the germplasm is likely to derive from the on-farm selection of open-pollinated seedlings, and that the vegetative spread of plants between farms may favor the selection of possible phenotypic variants, contributing thus to the creation of derived accessions and/or synonyms (Pugliano et al. 1980; Nunziata and Petriccione 2019). For these reasons, we attempted to infer the level of clonality in our population. According to the statistical evaluation of two indices of association, the level of clonality was very low and not significant, indicating that it was not meaningful trying to identify possible multilocus lineages by exploiting, for instance, other information as well (e.g., names, sites of collection, etc.). Similarly, the evaluation of the genetic distance and related dendrogram revealed normally distributed genetic distances. Specifically, they did not make evident a distance threshold (e.g., by peaks or asymmetry in the histogram) at which varieties would be possibly considered deriving from clonal reproduction and recent divergence (Arnaud-Haond et al. 2007), excluding the above-mentioned identical genotypes. This also indirectly suggests that the identified clones did not significantly affect the estimation

of the level of genetic diversity within our collection. We, therefore, inferred a possible genetic structure using a model-based clustering method. The analysis clearly suggested the presence of five sub-populations, whose members were in general well assigned. However, the pairwise genetic distance and differentiation between those clusters were low and most often not significant, not only for the small groups (i.e., fewer than ten members). Although F_{st} is not fully suited to assess population structure (Meirmans and Hedrick 2011), it is a very popular index also to describe the evolutionary history of derived populations based on the level of heterozygosity. The very limited genetic differentiation implies a rather uniform genetic basis of the accessions (with little foreign introduction), supporting a true local origin of the samples. Moreover, the limited differentiation may be also justified by the presence of rather homogenous geographic, ecological, and for human selected plants, agronomic forces, consistent with standardized cultivation practices and commercial uses of the apricots (Aradhya et al. 2003). Local (fine-scale) genetic differentiation has been verified in several instances in plants (Savolainen et al. 2007; Linhart and Grant 1996) as well as in landraces (Santos et al. 2019; Corrado and Rao 2017), although comparisons are not easy because the scale of the differentiation is usually defined by means of dispersal, typically pollen. On the other hand, to explain the limited genetic distances among sub-populations, it should be also considered that, for asexually propagated plants, somatic mutations are expected to be the main drivers of adaptive evolution to new environments (Miller and Gross 2011).

In conclusion, our work highlighted the high level of genotypic diversity present in an ex-situ collection of traditional varieties of apricot. As expected, the molecular analysis revealed possible homonymy and spurious classifications, which should be confirmed or solved by implementing a thorough morphological classification. Moreover, the very low level of clonality and genetic differentiation among the sub-populations identified by Bayesian analysis can be indicative of a possible common origin of the germplasm, and of an adaptive diversification that is mainly due to similar environmental and human-driven factors. These specificities should not only be considered

important backers for the conservation of neglected apricot resources, but also prompt actions to identify and exploit agronomically useful traits (e.g., adaptive or fruit quality-related) behind conservation initiatives.

Acknowledgements We thank Rosanna Manco and Claudio Capuozzo for the assistance during DNA isolation.

Author contributions BB: data analysis, manuscript review; RR: conceptualization and manuscript review; GC: design, data analysis, interpretation, and manuscript writing and review. All authors read and approved the final manuscript.

Funding Open access funding provided by Università degli Studi di Napoli Federico II within the CRUI-CARE Agreement. This work was supported by the SALVE project (Salvaguardia della Biodiversità Vegetale della Campania-Regione Campania, PSR 2007–2013, misura 214 az. f2).

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. *Mol Ecol Notes* 1(1–2):101–102
- Aradhya MK, Dangl GS, Prins BH, Boursiquot J-M, Walker MA, Meredith CP, Simon CJ (2003) Genetic structure and differentiation in cultivated grape *Vitis vinifera* L. *Genet Res* 81(3):179–192
- Arnaud-Haond S, Duarte CM, Alberto F, Serrao EA (2007) Standardizing methods to address clonality in population studies. *Mol Ecol* 16(24):5115–5139

- Baric S, Monschein S, Hofer M, Grill D, Dalla Via J (2008) Comparability of genotyping data obtained by different procedures—an inter-laboratory survey. *J Hortic Sci Biotechnol* 83(2):183–190
- Bartolini S, Massai R, Viti R (2020) The influence of autumn-winter temperatures on endodormancy release and blooming performance of apricot (*Prunus armeniaca* L.) in central Italy based on long-term observations. *J Hortic Sci Biotechnol* 95(6):794–803
- Biscotti N, Bonsanto D, Laghetti G (2022) Ethnobotanical study on traditional use of local fruit varieties in Gargano promontory (Apulia, Italy). *Genet Resour Crop Evol* 69:871–925
- Brown A, Feldman M, Nevo E (1980) Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics* 96(2):523–536
- Butler JM (2009) Fundamentals of forensic DNA typing. Academic Press, Cambridge, pp 1–500
- Corrado G, Rao R (2017) Towards the genomic basis of local adaptation in landraces. *Diversity* 9(4):51
- Corrado G, Forlani M, Rao R, Basile B (2021) Diversity and relationships among neglected apricot (*Prunus armeniaca* L.) landraces using morphological traits and SSR markers: Implications for agro-biodiversity conservation. *Plants* 10(7):1341
- Dieringer D, Schlötterer C (2003) MicroSatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes* 3(1):167–169
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14(8):2611–2620
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164(4):1567–1587
- Fossati T, Zapelli I, Bisoffi S, Micheletti A, Vietto L, Sala F, Castiglione S (2005) Genetic relationships and clonal identity in a collection of commercially relevant popular cultivars assessed by AFLP and SSR. *Tree Genet Genomes* 1(1):11–20
- Gürçan K, Çetinsığ N, Pınar H, Macit T (2019) Molecular and biological assessment reveals sources of resistance to Plum pox virus-Turkey strain in Turkish apricot (*Prunus armeniaca*) germplasm. *Sci Hortic* 252:348–353
- Hagen L, Khadari B, Lambert P, Audergon J-M (2002) Genetic diversity in apricot revealed by AFLP markers: species and cultivar comparisons. *Theor Appl Genet* 105(2–3):298–305
- Hagen L, Chaïb J, Fady B, Decroocq V, Bouchet J, Lambert P, Audergon J (2004) Genomic and cDNA microsatellites from apricot (*Prunus armeniaca* L.). *Mol Ecol Notes* 4(4):742–745
- Herrera S, Hormaza JI, Lora J, Ylla G, Rodrigo J (2021) Molecular characterization of genetic diversity in apricot cultivars: current situation and future perspectives. *Agronomy* 11(9):1714
- Hormaza J, Yamane H, Rodrigo J (2007) Apricot. In: Chittaranjan K (ed) Fruits and nuts. Springer, Heidelberg (DE), pp 171–187
- Ispizua VN, Guma IR, Feingold S, Clausen AM (2007) Genetic diversity of potato landraces from northwestern Argentina assessed with Simple Sequence Repeats (SSRs). *Genet Resour Crop Evol* 54(8):1833–1848
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23(14):1801–1806
- Jarni K, Jakše J, Brus R (2015) Vegetative propagation: linear barriers and somatic mutation affect the genetic structure of a *Prunus avium* L. stand. *For Int J Forest Res* 88(5):612–621
- Junhuan Z, Xuan C, Haoyuan S, Li Y (2012) Molecular fingerprinting and relative relationship of apricot cultivars in China by Simple Sequence Repeat (SSR) markers. *Afr J Biotechnol* 11(11):2631–2641
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281
- Khursheed A, Rasool A, Nazir R, Rather MA, Shalla AH (2020) Apricot. In: Antioxidants in fruits: properties and health benefits. Springer, Dordrecht, pp 523–545
- Lamia K, Hedia B, Jean-Marc A, Neila T-F (2010) Comparative analysis of genetic diversity in Tunisian apricot germplasm using AFLP and SSR markers. *Sci Hortic* 127(1):54–63
- Ledbetter CA (2008) Apricots. In: Hancock JF (ed) Temperate fruit crop breeding: germplasm to genomics. Springer, Dordrecht, pp 39–82
- Li M, Zheng P, Ni B, Hu X, Miao X, Zhao Z (2018) Genetic diversity analysis of apricot cultivars grown in China based on SSR markers. *Europ J Hortic Sci* 83(1):18–27
- Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. *Annu Rev Ecol Syst* 27(1):237–277
- Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Mol Ecol Resour* 11(1):5–18
- Mennone C (2016) Effect of climate on apricot productivity. *Informatore Agrario* 72(31):47–50
- Messina R, Lain O, Marrazzo M, Cipriani G, Testolin R (2004) New set of microsatellite loci isolated in apricot. *Mol Ecol Notes* 4(3):432–434
- Miller AJ, Gross BL (2011) From forest to field: perennial fruit crop domestication. *Am J Bot* 98(9):1389–1414
- Nunziata A, Petriccione M (2019) Storia della frutticoltura in Campania. In: Guidi S, Ercole S, Forconi V (eds) Frutti dimenticati e biodiversità recuperata. Quaderni Natura e Biodiversità. ISPRA, Rome, pp 8–82
- Ordidge M, Litthauer S, Venison E, Blouin-Delmas M, Fernandez-Fernandez F, Höfer M, Kägi C, Kellerhals M, Marchese A, Mariette S (2021) Towards a joint international database: alignment of SSR marker data for european collections of cherry germplasm. *Plants* 10(6):1243
- Potts SM, Han Y, Khan MA, Kushad MM, Rayburn AL, Korban SS (2012) Genetic diversity and characterization of a core collection of *Malus* germplasm using simple sequence repeats (SSRs). *Plant Mol Biol Rep* 30(4):827–837
- Prevosti A, Ocana J, Alonso G (1975) Distances between populations of *Drosophila subobscura*, based on

- chromosome arrangement frequencies. *Theor Appl Genet* 45(6):231–241
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959
- Pugliano G, Forlani M, Giofrè D, Pasquarella C, Rotundo A, Sonnino G (1980) Individuazione di cv idonee alla trasformazione industriale. *Riv Agr e Ricerca* 14(3):3–57
- Queiroz A, Assunção A, Ramadas I, Viegas W, Veloso MM (2015) Molecular characterization of Portuguese pear landraces (*Pyrus communis* L.) using SSR markers. *Sci Hortic* 183:72–76
- Rao R, Bencivenni M, Mura L, Araujo-Burgos T, Corrado G (2010) Molecular characterisation of Vesuvian apricot cultivars: implications for the certification and authentication of protected plant material. *J Hortic Sci Biotechnol* 85(1):42–47
- Rossarolla MD, Tomazetti TC, Vieira LN, Guerra MP, Klabunde GHF, Scherer RF, Pescador R, Nodari RO (2020) Identification and characterization of SSR markers of *Guadua chacoensis* (Rojas) Londoño & P.M. Peterson and transferability to other bamboo species. *3 Biotechnology* 10(6):273
- Roussos PA, Denaxa N-K, Tsafouros A, Efstathios N, Intidhar B (2016) Apricot (*Prunus armeniaca* L.). In: Simmonds M, Preedy V (eds) *Nutritional composition of fruit cultivars*. Elsevier, Amsterdam, pp 19–48
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, vol 2. Cold Spring Harbor Laboratory Press, New York
- Sansavini S (2019) Continua il rinnovo culturale, resta d'instabilità del mercato. *Rivista Di Frutticoltura e Di Ortofloricoltura* 83(5):46–47
- Santos JD, Chebotarov D, McNally KL, Bartholomé J, Droc G, Billot C, Glaszmann JC (2019) Fine scale genomic signals of admixture and alien introgression among Asian rice landraces. *Genome Biol Evol* 11(5):1358–1373
- Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. *Annu Rev Ecol Evol Syst* 38:595–619
- Sheikh ZN, Sharma V, Shah RA, Raina S, Aljabri M, Mir JI, AlKenani N, Hakeem KR (2021) Elucidating genetic diversity in apricot (*Prunus armeniaca* L.) cultivated in the North-Western Himalayan Provinces of India using SSR markers. *Plants* 10(12):2668
- Smouse RPP, Peakall R (2012) GenALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19):2537–2539
- Tsykun T, Rellstab C, Dutech C, Sipos G, Prospero S (2017) Comparative assessment of SSR and SNP markers for inferring the population genetic structure of the common fungus *Armillaria cepistipes*. *Heredity* 119(5):371–380
- Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. *Trends Biotechnol* 23(1):48–55
- Verdone M, Rao R, Coppola M, Corrado G (2018) Identification of zucchini varieties in commercial food products by DNA typing. *Food Control* 84:197–204
- Wang Y, Zhang J, Sun H, Ning N, Yang L (2011) Construction and evaluation of a primary core collection of apricot germplasm in China. *Sci Hortic* 128(3):311–319
- Wilkie P, Saridan A (1999) The limitations of vernacular names in an inventory study, Central Kalimantan, Indonesia. *Biodivers Conserv* 8(11):1457–1467
- Zeven AC (1998) Landraces: a review of definitions and classifications. *Euphytica* 104(2):127–139
- Zhebentyayeva T, Reighard G, Gorina V, Abbott A (2003) Simple sequence repeat (SSR) analysis for assessment of genetic variability in apricot germplasm. *Theor Appl Genet* 106(3):435–444
- Zhebentyayeva T, Ledbetter C, Burgos L, Llácer G (2012) Apricot. In: Badenes ML, Byrne DH (eds) *Fruit breeding*. Springer, New York (USA), pp 415–458
- Zimmerer KS, Carney JA, Vanek SJ (2015) Sustainable smallholder intensification in global change? Pivotal spatial interactions, gendered livelihoods, and agrobiodiversity. *Curr Opin Environ Sustain* 14:49–60

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.