RESEARCH ARTICLE

Genetic diversity of fig (*Ficus carica* L.) germplasm from the Mediterranean basin as revealed by SSR markers

Athanasios Sclavounos · Petros Roussos · Sotiria Milla · Panagiotis Kostas · Yiannis Samaras · Carlo Pozzi · Johiruddin Molla · Annapurna Chitikineni · Rajeev K. Varshney · Andreas Voloudakis

Abstract Fig (*Ficus carica* L.) tree is cultivated worldwide and is highly appreciated for its fruit, which is consumed fresh or dried, having high nutritional and pharmaceutical value and for these reasons there is an increasing interest for its cultivation. In the present study, an ex situ collection of 60 fig accessions (41 indigenous Greek and 19 from other Mediterranean countries) was established and its diversity was analyzed using eight simple sequence repeat (SSR) loci. Greek fig genotypes showed relatively low allelic variation (the average number of SSR alleles per locus was 3.75), an excess of heterozygosity (mean $H_e = 0.489$ and $H_o = 0.557$), and extensive outbreeding (mean F index -0.151). Cluster analysis showed that the established fig population exhibited weak genetic structure, with most of the genetic variation (89%) being present within individual members of the clusters. Both cluster and principal coordinate analysis confirmed that there is little correlation between genetic makeup and geographical origin of the fig accessions. Polymorphism information content with an average of 0.421 was reasonably informative. An identification key scheme for fig cultivars that

A. Sclavounos · A. Voloudakis (✉)
Laboratory of Plant Breeding and Biometry, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece
e-mail: avoloud@aua.gr

Present Address:
A. Sclavounos
Directorate of Decentralized Agencies of Peloponnese, Western Greece and the Ionian Sea, Office of Rural Development of Kephalonia, Ministry of Rural Development and Food, 28100 Argostoli Kephalonia, Greece

P. Roussos
Laboratory of Pomology, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece

S. Milla · Y. Samaras
Department of Food Technology, Ionian University, Bergotis Ave., 28100 Argostoli Kephalonia, Greece

P. Kostas
Directorate of Plant Genetic Resources and Propagating Material of Cultivated Plant Species, Ministry of Rural Development and Food, Kapnokoptiriou 6, 10433 Athens, Greece

C. Pozzi
Department of Agricultural and Environmental Sciences DISAA, University of Milano, Via Celoria 2, 20133 Milan, Italy

J. Molla · A. Chitikineni · R. K. Varshney
Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad 502324, India

Present Address:
A. Chitikineni · R. K. Varshney
Murdoch's Centre for Crop and Food Innovation, State Agricultural Biotechnology Centre, Food Futures Institute, Murdoch University, Murdoch, Australia

will be useful in cultivar discrimination and intellectual property protection was developed. This work will contribute to a sustainable fig production regionally and worldwide, through the establishment and conservation of a reference fig collection, providing germplasm for future breeding efforts.

Keywords *Ficus carica* · Genetic diversity · Genetic resources · Microsatellite · Molecular marker · Genotyping

Introduction

Fig (*Ficus carica* L.) (2n=26) belongs to the Moraceae family, known worldwide for its fruit and the presence of latex in all plant parts. Figs are eaten fresh or dried, are rich in phenolic antioxidants and nutrients with high fiber content (Vinson et al. 2005), being ideal for the human diet. According to Vavilov (1951), fig's origin is Transcaucasian, whereas other authors place it in southern Arabia or the eastern Mediterranean basin (Stover 2007). From its center of origin, fig spread to the Mediterranean basin, and in the 16th century to the east (Southeast Asia, China, Japan) as well as the New World. Nowadays, the cultivation of fig is worldwide, including Australia and South Africa.

In Greece, fig is cultivated almost in the entire country, with three main production centers, namely Attica (Mesogaia region), Evia (Kimi and Taxiarches regions) and Messinia (Peloponnese prefecture). Fig cultivation declined in Greece in the last decades; in 1961 fig was cultivated on 250,000 acres, the production was 169,259 tons, with a yield of 700 kg/acre, whereas in 2011 only 38,000 acres were cultivated, with a total production of 9400 tons and a yield of 250 kg/acre (FAO stat 2012, www.fao.org).

Fig is threatened by genetic erosion, mainly due to urbanization, monoculture, and introduction of plant material from remote regions. Today, because of the increasing interest in fresh and dried figs, fig cultivation assumes new perspectives. Since fig is not subject to intensive breeding (Flaishman et al. 2008) the improvement of fig cultivars should be based on the rich genetic diversity present in established fig populations. The ambiguity in the description of fig cultivars, the lack of official cultivar names and the existence of variants within cultivars hinder proper

fig cultivar identification (Perez-Jimenez et al. 2012). Condit (1955) listed more than 700 fig cultivars, with a great deal of confusion in cultivar identification and their relationships still existing.

Plant germplasm characterization, aiming at its conservation, is traditionally carried out using morphological or agronomical traits. These criteria are often variable across years and locations since these phenotypic traits are influenced by the genotype-environment interactions. In contrast, DNA-based data are stable, reliable, and detectable in all tissues regardless of developmental and differentiation stage and are not confounded by environmental, pleiotropic, and epistatic effects (Mondini et al. 2009). Molecular markers such as microsatellites (simple sequence repeats, SSRs), RAPDs, ISSRs, RFLPs and others have been used in fingerprinting and assessing genetic diversity in various fig collections (Papadopoulou et al. 2002; Chatti et al. 2010; Perez-Jimenez et al. 2012; Ganopoulos et al. 2015; Boudchicha et al. 2018; Rodolfi et al. 2018; Ergül et al. 2021). Microsatellites provide adequate resolution of germplasm differences due to their high polymorphism and codominance, are simple, quick, relatively inexpensive, and exhibit high reproducibility among laboratories, and as a result are still used for fingerprinting in plant species (Gupta and Varshney 2000; Mondini et al. 2009).

In the present study, the genetic diversity in fig germplasm was evaluated for an established population in an ex situ collection (60 fig accessions, from four Mediterranean countries and the USA) and an identification key for fig cultivars was developed using SSRs. The work was also aimed at assigning each accession to a group based on genetic diversity, thus evaluating the structure of the fig population.

Materials and methods

Plant material

In the present study the fig population comprised of 60 accessions categorized based on their geographical origin, with 41 accessions from Greece (central-eastern, central-western, northern Greece, as well as Crete, Peloponnesus, Lesvos, Syros) and 19 from other Mediterranean countries—Italy (12), Cyprus (6), Turkey (1), Spain (1); Suppl.

Table S1). In particular, 33 of them were selected from different regions of Greece, based on their special agronomical, morphological, and fruit quality characteristics, based on information obtained from farmers and authors' personal field inspection. The name of the accession usually represents a geographical qualifier or sometimes it derives from a morphological or fruit quality characteristic. Plants were vegetatively propagated by cuttings and planted in the experimental orchard of the Agricultural University of Athens (AUA). The remaining 27 fig accessions were acquired as trees and planted in the same experimental orchard and used as reference material. Eight of the 27 accessions were Greek fig cultivars analyzed previously using RAPD markers (Papadopoulou et al. 2002). The remaining 19 accessions were internationally well-known fig varieties from other countries (Suppl. Table S1).

DNA isolation

Plant DNA was isolated from fig leaves of all 60 accessions using the CTAB method (Murray and Thompson 1980). The DNA concentration was estimated spectrophotometrically and its integrity was evaluated by electrophoresis on 0.8% agarose gel followed by ethidium bromide staining. DNA suitability as PCR template was checked by PCR reaction using primers for the ITS (Internal Transcribed Spacer) locus, following the methodology described by Roy et al. (2010).

SSR analysis

Eight SSR markers, namely MFC1 to MFC8, developed by Khadari et al. (2001) for fig, were used in this study. DNA amplification reactions were carried out in a total volume of 25 μl containing 0.5 mM of each PCR primer, 200 mM of each deoxynucleotide triphosphate, 1.5 mM of $MgCl_2$, 1 U of *Taq* DNA polymerase (New England Biolabs, USA) and 50 ng of template DNA (Khadari et al. 2001), using a PTC-200/A100 thermocycler (BioRad, USA). No DNA negative control reactions were performed.

PCR products were resolved using a 12% non-denaturing PAGE in a 20×20 gel (Biorad Protean II, USA) at 60 V for 20 min, followed by 180 V for 5 h. Gels were subsequently stained with ethidium bromide and photographed under UV, with photos digitized for further analysis. A DNA ladder (50 bp GeneRuler, Thomas Scientific, USA) was loaded in three wells in each gel, in asymmetric locations, to avoid gel orientation problems when scoring and assist in allele size determination (Hoffman and Amos 2005).

The amplified bands per SSR were scored for each fig accession, using GelAnalyzer (2010a) (http://www.GelAnalyzer.com). Only gels/lanes with unambiguous band patterns, after background subtraction using GelAnalyzer, were considered for allele assignment. A band was accepted, when the corresponding fluorescence intensity value, from the digitized photos, was > 10 fluorescent units. Bands with a fluorescence intensity < 35% (i.e. stutter bands) of the main fluorescence intensity value were filtered out following previous recommendations (Ewen et al. 2000; UPOV/INF/17/1 2010). The microsatellite alleles were sized using a standard curve generated for each gel, employing the known molecular size DNAs of the DNA ladder. Since PCR products of a fig accession per microsatellite were electrophoresed two to four times (each time in a different gel), the size of an allele was estimated several times, with the mean value recorded as the allele size. The allele sizes, for each microsatellite and accession, were recorded in an Excel spreadsheet (Microsoft Inc., Redmond, USA) producing thus a data matrix for data storage and further processing. SSR allele size data were binned using Flexibin (Amos et al. 2007), following the methodology described by Ghosh et al. (1997). A final correction of the allele's size was done by visual gel inspection as previously reported (Pompanon et al. 2005; Hoffman and Amos 2005).

Estimation of genotyping error

MicroChecker v.2.2.3 (Van Oosterhout et al. 2004) was used to statistically estimate the percentage of null (nonamplified alleles due to nucleotidic changes in flanking sequences of the SSR) alleles per SSR, which is the main non-technical contributor to the genotyping error.

In order to estimate the genotyping error, a subset of 20% randomly selected genotypes of the fig population was reanalyzed (Pompanon et al. 2005) following the same methodology, except that a different PCR thermocycler (Eppendorf Mastercycler

Gradient 5341, USA), DNA polymerase (Phusion® High-Fidelity, NEB, USA), and a new fig DNA preparation were used. In addition, to strengthen the reliability of the obtained results: a) independent random PCRs, for each of the eight SSRs, were conducted again for approximately 20% of the fig accessions and were re-genotyped, b) for each PCR, PCR samples were re-electrophorized two to four times in different gels, and c) the allele sizes were scored twice by two different persons (Hoffman and Amos 2005).

Data and cluster analyses

Based on the SSR allele size data, genetic variability parameters (Allele per locus (Na), effective allele per locus (Ne), observed (Ho) and expected heterozygosity (He), Fixation index (F), χ^2 test for deviation from Hardy-Weinberg equilibrium (HWE) (Table 1) per locus, private alleles summary (*PAS*) per country, and the number of genotypes for all SSR loci) were computed for all the 60 fig accessions originating from Greece and other Mediterranean countries using GenAlEx v.6.5 (Peakall and Smouse 2012).

In order to depict the genetic relationships among accessions of the established fig population the DARwin v6 (Perrier and Jacquemoud-Collet 2006) was employed. Missing allelic data were handled choosing the pairwise allele deletion option, at a threshold of 70%. A DARwin file with extension ".DIS" stores the dissimilarity lower semi-matrix (without the diagonal) as computed by the software. Dissimilarity re-sampling done with 10,000 bootstrap and each semi-matrix is successively recorded at the end of the file. Dissimilarity based cluster analysis was performed and dendrogram was done following the Weighted Neighbor-Joining (WNJ) method with 10,000 bootstraps. In the dendrogram, the scale defined the edge length. In order to determine genetic relationship of fig accessions per country of origin MEGA11 (Tamura et al. 2021) was employed using Nei's distance (Nei 1972). Analysis of molecular variance (AMOVA) was carried out by GenAlEx, with the analysis based on groups as revealed by the above CA analysis for all SSR loci.

Population structure

The population structure was investigated using non-Bayesian procedure, the Discriminant Analysis of Principal Components (DAPC) (Jombart et al. 2010) in the *adegenet* package for R software (R Development Core Team 2011), where variance in the sample is partitioned into a between-group and within-group component, without making assumptions on panmixia. The number of clusters was assessed using the *find.clusters* function, which runs successive *K-means* clustering with increasing number of clusters (k). The optimal number of clusters was selected using the Bayesian Information Criterion (BIC) for assessing the best supported model, and therefore the number and nature of clusters.

Establishment of an identification key for fig

The polymorphism information content (PIC) value per locus was estimated using PICCalc software (Nagy et al. 2012). To establish an identification (Id) key for fig, the methodology of Tessier et al. (1999) was followed. According to this method two parameters were estimated; the confusion probability C*j*, and the discriminating power D*j* ($Dj = 1 - Cj$). These parameters were computed by the frequencies of allelic pattern revealed per SSR locus in the established fig population.

To determine the sequence and the number of SSRs that are sufficient to establish an identification key for fig: (a) an alphabetic character was given to each SSR allele, with fig accessions exhibiting the same SSR allele pattern obtaining the same alphabetic character designation, (b) SSR loci were sorted by descending order of D*j*, (c) fig accessions with the same alphabetic designation per SSR were pooled together in the same group, and (d) the total number of expected and the observed non-differentiated fig pairs of comparisons were determined after the completion of grouping.

Table 1 Genetic parameters per SSR marker for the fig accessions from Mediterranean basin Italy, Greece, Turkey, Cyprus, Spain

Microsatellite	N	Allele range	Allele size	Genotypes per locus	Na	Ne	Ho	He	F	Hardy–Weinberg equilibrium χ^2-test/significant level a=0.05		
										DF	χ^2	P-value
MFC1	59[a]	170–202	170[b], 186, <u>192</u>, <u>198</u>, 202	8	5	2.177	0.424	0.541	0.251	10	39.759	<.00001*
MFC2	60	158–194	<u>158</u>, 164, 174, 186, <u>194</u>	7	5	2.459	0.800	0.593	−0.348	10	20.038	.028895*
MFC3	60	126–154	<u>126</u>, 128, 130, 132, 138, <u>154</u>	9	6	2.419	0.533	0.587	0.091	15	40.462	.000386*
MFC4	60	206–230	206, **230**	2	2	1.946	0.833	0.486	−0.714	1	30.612	<.00001*
MFC5	60	134–142	134, **140**	2	2	1.260	0.233	0.206	−0.132	1	1.047	.306199
MFC6	48*	284–328	284, <u>290</u>, 304, 316, 328	8	5	2.869	0.646	0.651	−0.009	10	6.771	.746873
MFC7	60	152–168	152, 160, 168	6	3	2.515	0.700	0.602	−0.162	3	39.244	<.00001*
MFC8	60	177–205	**177**, 205	2	2	1.321	0.283	0.243	−0.165	1	1.634	.201151
Average					3.75	2.121	0.557	0.489	−0.151			

Bold allele size indicates monomorphic allele

Underlined allele size indicates rare allele

N Number of fig accessions; *Na* Alleles per locus; *Ne* Effective alleles per locus; *Ho* Observed heterozygosity; *He* Expected heterozygosity; *F* Fixation Index; *DF* Degree of freedom; χ^2 *est* Chi squeare estimation per locus; *not signif* Not significant

[a]For MFC1 and MFC6, 1 and 12 out of 60 fig accessions respectively, did not reveal any banding pattern. Genetic parameter estimated only for the fig accessions that revealed banding pattern

[b]In reference fig accession from foreign countries, one more different allele were revealed, in MFC1, 170 bp

*Indicates significance at a=0.05

Results

Genetic diversity among fig genotypes revealed by SSR markers

In the present study, 60 fig accessions were examined using eight microsatellites (SSRs), designated MFC1, MFC2, MFC3, MFC4, MFC5, MFC6, MFC7 and MFC8 (Khadari et al. 2001). For each fig accession showing a banding pattern, one allele (homozygous individual) or two alleles (heterozygous individual) were identified, and their size was estimated. All figs revealed PCR banding pattern for seven out of the eight SSRs (Suppl. Fig. S1). For MFC6, no banding pattern (no PCR products) was observed for 12 out of 60 fig accessions, namely seven from Greece (Acc. No 118, 124, 131, 139, 140, 142, 146), three from Italy (Acc. No 251, 252, 253), one from Cyprus (Acc. No 236) (Suppl. Fig. S1), and one from Turkey (Acc. No 230). PCR reactions for the 12 above-mentioned fig accessions were repeated, confirming the initial observation.

MicroCheker analysis did not reveal statistically significant indications of the stuttering, scoring, large allele dropout and null allele presence, except the case of MFC1 (only for the presence of null alleles). To validate the above analysis, the genotyping procedure was repeated for approximately, randomly selected, 20% of the fig accessions, and no new alleles appeared for any of the SSRs.

In total, the population of the 60 fig accessions under study resulted in 58 different genotypes for the eight SSR loci. The eight SSRs resulted in 30 alleles with a mean value of 3.75 alleles per SSR. The observed heterozygosity Ho (0.557) was higher than He (0.489). The Fixation Index (F) was negative − 0.152. In particular, six (MFC 2, MFC4, MFC5, MFC6, MFC7 and MFC8) out of the eight SSR have negative F value, while the remaining two (MFC1 and MFC3) have positive F value. Based on χ^2-test, three SSRs (MFC5, MFC6, and MFC8) follow the HWE, while the remaining five deviated from the HWE, at a significance level $\alpha = 0.05$ (Table 1).

We detected private alleles in Greek fig accessions as follows: for MFC1 the 192 bp (in genotype 109); for MFC2 the 158 bp (in genotype 145), and 186 bp (in genotypes 101,112,121,123,128,132,155,157, and 162); for MFC3 the 126 bp (in genotypes 118 and 131) and 154 bp (in genotypes 101,155); for MFC6 the 284 bp (in genotype 127) and 290 bp (in genotype

161). In Cyprus the private allele was a 170 bp MFC1 allele (in genotypes 220, 233, 236, and 249).

Lastly, a comparison was made for the genetic parameters found in the literature where the MFC SSRs were utilized (Suppl. Table S2). The values of the genetic parameters obtained in the present study agree with the previously published work.

Genetic relationship among fig genotypes—population structure of fig germplasm collection

From the dendrogram generated (Fig. 1) it appears that the fig population studied could be divided in three large groups, which were named Cluster I, II, and III, with each cluster subdivided into two subgroups, 1 and 2 (subgroup I-1, subgroup I-2 etc.), resulting in a total of 6 subgroups. Figs from different countries could be found in the same subgroup, except for subgroup III-2, which contains figs only from Greece. The AMOVA analysis for the 6 subgroups showed that 89% of the total variability within the population is due to genetic differences between the individual fig accessions of each subgroup and only 11% is accounted for by differences between the subgroups. Fst value of 0.132 indicates a small to negligible difference between the six subgroups. When the country of origin for the fig accessions was considered, based on the genetic distance of Nei, two major groups are created, one including the figs from Greece, Italy and Spain, and the second one figs from Cyprus and Turkey (Suppl. Fig. S2).

The genetic structure of the fig tree cultivars was investigated by non-Bayesian population assignment analysis. For DAPC analysis, 20 PCA axes and three discriminant functions were retained. The DAPC analysis indicates a partial sub-structuring of fig accession groups. The Cypriot and Italian accessions were grouped separately, while the remaining accessions were clustering together, most likely due to potential gene flows occurring between these populations (Fig. 2). The data were consistent when the analysis was conducted with clone correction (data not shown).

Allele patterns per SSR locus analysis, in fig accessions, revealed different genotypes, ranging from two genotypes for MFC4, MFC5, and MFC8, to nine genotypes for MFC3, as shown in Table 1. Combining the data for all eight SSR markers
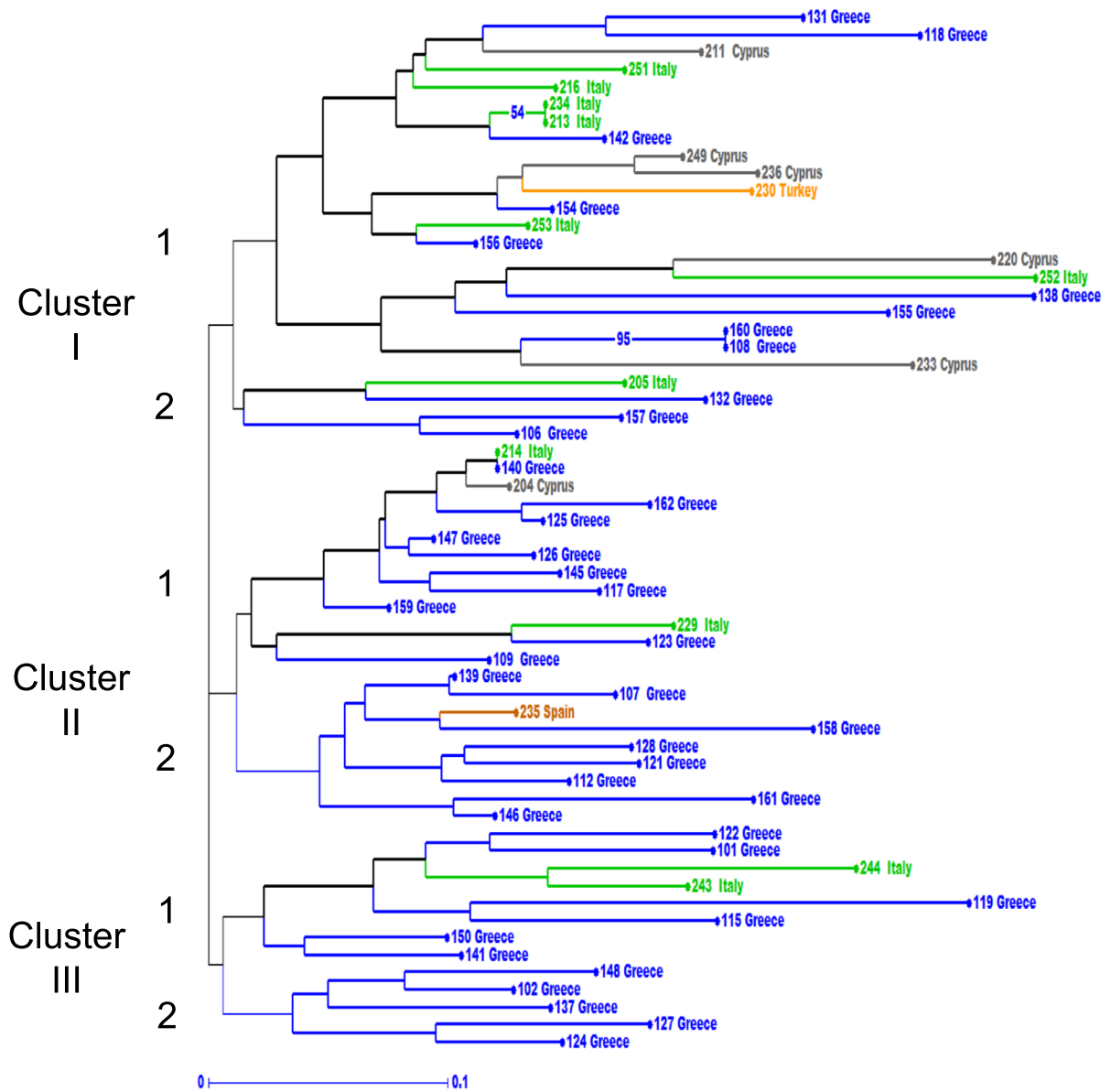
Fig. 1 Dendrogram showing the genetic relationships of 60 Fig accessions based on SSR data. The dendrogram was constructed in DARwin v 6.0.15 using the WNJ method (confirmed by 10,000 bootstraps). The scale indicates the edge length

indicated the presence of 58 different genotypes for the 60 fig accessions studied (Suppl. Table S2).

Identification key for figs

In this study, an identification (Id) key was generated based on the discriminating power (Dj) of each SSR. The SSR loci were hierarchically ordered according to their Dj values. MFC6 ranked first, as it is the most

discriminative among the eight SSRs. As a result, the Id key produced was: MFC6-MFC3-MFC1-MFC2-MFC7-MFC8-MFC5-MFC4. MFC4 and MFC5, having the smallest Dj, were not included in the Id key as they lacked discriminative power (Table 2). The sixty fig accessions in this study produced 1770 pairs of comparisons. Using the above Id key, 58 out of the 60 fig accessions analysed were differentiated.
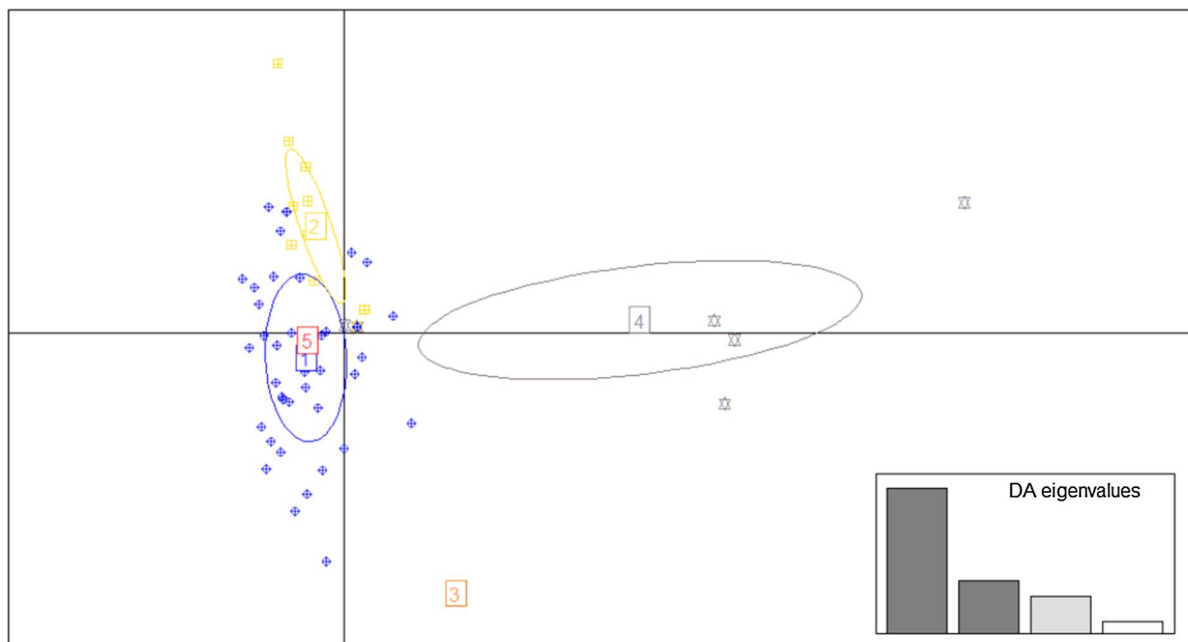
Fig. 2 Scatterplot of individuals on the two principal components of DAPC. The graph represents the individuals as symbols and the groups as inertia ellipses. Eigenvalues of the analysis are displayed in the inset. The numbers refer to the country of origin: 1(blue) = Greece; 2 (yellow) = Italy; 3 (dark range) = Turkey; 4 (grey) = Cyprus; 5 (red) = Spain. (Color figure online)

Discussion

Genetic erosion threatens the local fig varieties and landraces and intensifies the need to protect them. For this reason, an ex situ fig collection was established in the experimental orchard of the Agricultural University of Athens. This comprised mainly of uncharacterized Greek fig accessions, along with few partly characterized Greek varieties (Papadopoulou et al. 2002) along with some well-known foreign cultivars. Ex situ collection where landraces are preserved without the accompanying information on their characteristics and registration of their region of origin would be of limited value. The information describing plant germplasm could be used for identification purposes, recognition of deficiencies of the collection and planning future efforts to strategically enrich it with new plant material. Such information includes morphological, agronomical features, biochemical and molecular data. In the present study, the genetic characterization of the established fig population was described using SSR markers. Such investigations with plant genetic resources are a prerequisite for breeding crops in order to face new challenges, including climate change. Data obtained were also used to propose an identification key scheme.

In a number of genetic studies using microsatellites, genotyping errors that are due to null alleles (nonamplified alleles), DNA degradation and low DNA concentrations are increasingly recognized as important factors that could render the conclusions doubtful (Hoffman and Amos 2005). Especially null alleles that usually result from changes in flanking region sequence of the SSR could alter the estimation of the genetic parameters of the population under study. In the present study, the frequency of null alleles was statistically estimated per SSR, and it ranged from 0 to 3.53%. This range is considered non-significant since frequencies of 5–8% introduce only a small bias in the genetic parameters investigated (Chapuis and Estoup 2007). Moreover, re-genotyping 20% of the members of the population studied reinforced the reliability of the results.

The dendrogram (Fig. 1) shows that in the sub-group I-1 two Italian varieties, namely Dottato

Table 2 Genetic Information indexes

SSR locus*	No Gen.**	PIC***	Cj	Dj	C_L	D_L	Steps	Locus combination	Different multilocus genotypes produced per step	Total pairs of genotypes: 1770 Pairs of genotypes that do not differentiate	
										Expected	Observed
MFC6	8	0.584	0.15	0.85	0.17	0.83	1	MFC6	8	265.50	269
MFC3	9	0.546	0.21	0.79	0.22	0.78	2	MFC6+MFC3	30	55.76	59
MFC1	8	0.435	0.29	0.71	0.31	0.69	3	MFC6+MFC3+MFC1	43	16.17	23
MFC2	7	0.511	0.37	0.61	0.38	0.62	4	MFC6+MFC3+MFC1+MFC2	53	5.98	8
MFC7	6	0.525	0.4	0.6	0.41	0.59	5	MFC6+MFC3+MFC1+MFC2+MFC7	56	2.39	5
MFC8	2	0.214	0.59	0.41	0.59	0.41	6	MFC6+MFC3+MFC1+MFC2+MFC7+MFC8	58	1.41	2
MFC5	2	0.185	0.64	0.36	0.64	0.36	7	MFC6+MFC3+MFC1+MFC2+MFC7+MFC8+MFC5	–	0.90	
MFC4	2	0.368	0.74	0.26	0.74	0.26	8	MFC6+MFC3+MFC1+MFC2+MFC7+MFC8+MFC5+MFC4	–	0.67	

*Genetic parameters based on genetic fingerprinting/genotype frequencies per SSR locus: Cj: confusion propability, CL confusion probability when sample is very large ($N \rightarrow \infty$). Dj: discriminating power, DL discriminating power when sample is very large ($N \rightarrow \infty$). Mathematic relationship is Cj+Dj=1

**No Gen: Number of genotypes per Locus references in all 60 fig tree accessions not only to that with origin from Greece

***PIC: polymorphism information content based on allele frequencies per SSR locus

(accession number 234) and San Pietro (accession number 213), are genetically close, something that is known in the literature (Papadopoulou et al. 2002). However, these two varieties differ phenotypically in leaf shape and cavity size within the fruit (Ntanos et al. 2015). In addition, in the same subgroup two figs from Greece, namely Maurosykia (accession number 108) and Zakynthos (accession number 160) appear genetically and morphologically close, even though they were collected in very distant regions of Greece. Minor differences between Maurosykia and Zakynthos need to be further investigated. Finally, in subgroup II-1 the genotypes Vasilika Mellisi (140) from Greece and Rosso Dendro (214) from Italy are similar. However, these two varieties differ phenotypically in the size and shape of the leaf and fruit stalk length is significantly longer in the second one (Ntanos et al. unpublished data).

In the fig population studied in the present work, the parameters that quantify the genetic variability were calculated using 8 microsatellites. Our results are consistent with published data for (Suppl. Table S1) (Khadari et al. 2003; Giraldo et al. 2008, 2005; Saddoud et al. 2007; Achtak et al. 2009; Aradhya et al. 2010; Caliskan et al. 2012; Perez-Jimenez et al. 2012; Ganopoulos et al. 2015; Boudchicha et al. 2018; Rodolfi et al. 2018; Ergül et al. 2021). The observed differences in genetic parameters observed in published works may be due to the selection of different SSRs, the genotypes analyzed, the differences in the methodology followed for the examination of the samples. Our finding that observed heterozygosity is higher than expected, resulting in a negative F value, agrees with previous works (Suppl. Table S2). Negative F values are the result of a negative assortative mating (Lachance 2016), due to fig's entomophilous pollination, or to heterotic selection by man. Klekowski (1988) pointed out that perennial species tend to exhibit high heterozygosity as a mechanism to overcome the harmful effects of residual mutations. Despite the high observed heterozygosity in our fig population we observed a limited genetic grouping, as previously published (Aradhya et al. 2010; Caliskan et al. 2012; Perez-Jimenez et al. 2012).

From the present study it appears that the geographical origin could not be the main criterion for classification, as described also elsewhere (Giraldo et al. 2008; Aradhya et al. 2010; Boudchicha et al. 2018). Only Ikegami et al. (2009), using a population of eight Japanese and 11 foreign imports, suggested that figs from Japan are not genetically related to figs from other countries. Many authors attribute the absence of grouping of figs to the propagation of the fig tree, which favors the exchange of plant material between different geographical areas (Giraldo et al. 2008; Aradhya et al. 2010). The same partial lack of clustering was observed in the DAPC analysis. We opted to use this approach, instead of a Bayesian one, because the latter assumes that markers are not linked and that the population is panmictic. The DAPC analysis is a convenient approach for clonal or partially clonal populations (Jombart et al. 2010), where it aims at maximizing the discriminatory capacity of the between-groups variance.

The development of a reliable tool for the identification of the variety is necessary to ensure the identity of the plant material for the registration, the protection of the variety and the management of the propagating material. In the present study, we proposed a key identification scheme using six microsatellites that could distinguish 58 of the 60 fig genotypes in the ex situ collection (98% resolution). A study on the morphological characteristics showed that fig accessions that did not differ by molecular analysis could differ in the color of the fruit peal and of the fruit flesh (Ntanos et al. 2015). The latter suggests that a combined approach involving molecular and morphological analyses may be necessary to increase the resolution power.

Declarations

References

Achtak H, Oukabli A, Ater M, Santoni S, Khadari B (2009) Microsatellite markers as reliable tools for fig cultivar identification. J Am Soc Hortic Sci 134:624–631

Amos WJ, Hoffman I, Frodsham A, Zhang L, Best S, Hill AV (2007) Automated binning of microsatellite alleles: problems and solutions. Mol Ecol Notes 7:10–14. https://doi.org/10.1111/j.1471-8286.2006.01560.x

Aradhya MK, Stover ED, Velasco D, Koehmstedt A (2010) Genetic structure and differentiation in cultivated fig (*Ficus carica* L.). Genetica 138:681–694. https://doi.org/10.1007/s10709-010-9442-3

Boudchicha RH, Hormaza JI, Benbouza H (2018) Diversity analysis and genetic relationships among local Algerian fig cultivars (*Ficus carica* L.) using SSR markers. S Afr J Bot 116:207–215. https://doi.org/10.1016/j.sajb.2018.03.015

Caliskan O, Polat AA, Celikkol P, Bakir M (2012) Molecular characterization of autochthonous Turkish fig accessions. Span J Agric Res 10:130–140. https://doi.org/10.5424/sjar/2012101-094-11

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol 24:621–631. https://doi.org/10.1093/molbev/msl191

Chatti K, Baraket G, Ben Abdelkrim A, Saddoud O, Mars M, Trifi M, Salhi Hannachi AC (2010) Development of molecular tools for characterization and genetic diversity analysis in Tunisian fig (*Ficus carica*) cultivars. Biochem Genet 48:789–806. https://doi.org/10.1007/s10528-010-9360-1

Condit IJ (1955) Fig varieties: a monograph. Hilgardia 23:323–538

Ergül A, Büyük BP, Hazrati N, Yılmaz F, Kazan K, Arslan N, Özmen CY, Aydin SS, Bakir M, Tan N, Kösoğlu I, Çobanoğlu F (2021) Genetic characterisation and population structure analysis of Anatolian figs (*Ficus carica* L.) by SSR markers. Folia Hortic 33:49–78. https://doi.org/10.2478/fhort-2021-0005

Ewen KR, Bahlo M, Treloar SA, Levinson DF, Mowry B, Barlow JW, Foote SJ (2000) Identification and analysis of error types in high-throughput genotyping. Am J Hum Genet 67:727–736. https://doi.org/10.1086/303048

Flaishman M, Radov V, Stover ED (2008) The fig: botany, horticulture, and breeding. In: Jules J (ed) Horticultural reviews, vol 34. John Wiley & Sons Ed., Inc., New Jersey, pp 113–1967

Ganopoulos I, Xanthopoulou A, Molassiotis A, Karagiannis E, Mousiadis T, Katsaris P, Aravanopoulos F, Tsaftaris A, Kalivas A, Madesis P (2015) Mediterranean basin *Ficus carica* L.: from genetic diversity and structure to authentication of a protected designation of origin cultivar using microsatellite markers. Trees 29:1959–1971. https://doi.org/10.1007/s00468-015-1276-2

Ghosh S, Karanjawala ZE, Hauser ER, Ally D, Knapp JI, Rayaman JB, Musick A, Tannenbaum J, Te C, Shapiro S, Eldridge W, Musick T, Martin C, Smith JR, Carpten JD, Brownstein MJ, Powell JI, Whiten R, Chines P, Nylund SJ, Magnuson VL, Boehnke M, Collins FS (1997) Methods for precise sizing, automated binning of alleles, and reduction of error rates in large-scale genotyping using fluorescently labeled dinucleotide markers. FUSION (Finland–US Investigation of NIDDM Genetics) Study Group. Genome Res 7:165–178. https://doi.org/10.1101/gr.7.2.165

Giraldo E, Viruel MA, Lopez-Corrales M, Hormaza JI (2005) Characterisation and cross-species transferability of microsatellites in the common fig tree (*Ficus carica* L.). J Horticult Sci Biotechnol 80:217–224. https://doi.org/10.1080/14620316.2005.11511920

Giraldo E, Lopez-Corrales M, Hormaza JI (2008) Optimization of the management of an ex-situ germplasm bank in common fig with SSRs. J Am Soc Horticult Sci 133:69–77. https://doi.org/10.21273/JASHS.133.1.69

Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163–185. https://doi.org/10.1023/A:1003910819967

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Mol Ecol 14:599–612. https://doi.org/10.1111/j.1365-294X.2004.02419.x

Ikegami H, Nogata H, Hirashima K, Awamura M, Nakahara T (2009) Analysis of genetic diversity among European and Asian Fig varieties (*Ficus carica L.*) using ISSR, RAPD and SSR markers. Genet Resour Crop Evol 56:201–209. https://doi.org/10.1007/s10722-008-9355-5

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94. https://doi.org/10.1186/1471-2156-11-94

Khadari B, Hochu I, Santoni S, Kjellberg F (2001) Identification and characterization of microsatellite loci in the common fig (*Ficus carica* L.) and representative species of the genus *Ficus*. Mol Ecol Notes 1:191–193. https://doi.org/10.1046/j.1471-8278.2001.00072.x

Khadari B, Hochu I, Bouzid L, Roger JP, Kjellberg F (2003) The use of microsatellite markers for identification and genetic diversity evaluation of the fig collection in CBNMP. Acta Horticult 605:77–86. https://doi.org/10.17660/ActaHortic.2003.605.10

Klekowski EJ (1988) Genetic load and its causes in long-lived plants. Trees 2:195–203

Lachance J (2016) Hardy–Weinberg equilibrium and random mating. In: Encyclopedia of evolutionary biology, pp. 208–211. ISBN 978-0-12-800426-5

Mondini L, Noorani A, Pagnotta MA (2009) Assessing plant genetic diversity by molecular tool. Diversity 1:19–35. https://doi.org/10.3390/d1010019

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325. https://doi.org/10.1093/nar/8.19.4321

Nagy S, Poczai P, Cernak I, Gorji AM, Hegedus G, Taller J (2012) PICcalc: An online program to calculate polymorphic information content for molecular genetic studies. Biochem Genet 50:670–672. https://doi.org/10.1007/s10528-012-9509-1

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292. https://doi.org/10.1086/282771

Ntanos EK, Roussos PA, Sklavounos A, Voloudakis A, Assimakoroulou A (2015) Phenotypic characteristics of

various fig genotypes. In: 27th conference of the Greek Society for horticultural science. Volos

Papadopoulou K, Ehaliotis C, Tourna M, Kastanis P, Karydis I, Zervakis G (2002) Genetic relatedness among dioecious *Ficus carica* L. cultivars by random amplified polymorphic DNA analysis, and evaluation of agronomic and morphological characters. Genetica 114:183–194. https://doi.org/10.1023/A:1015126319534

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Perez-Jiminez M, Lopez B, Dorado G, Pujadas-Salva A, Guzman G, Hernandez P (2012) Analysis of genetic diversity of southern Spain fig tree (*Ficus carica* L.) and reference materials as a tool for breeding and conservation. Hereditas 149:108–113. https://doi.org/10.1111/j.1601-5223.2012.02154.x

Perrier X, Jacquemoud-Collet JP (2006) DARwin Software. http://darwin.cirad.fr/darwin

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev Genet 6:847–859. https://doi.org/10.1038/nrg1707

R Development Core Team (2011) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rodolfi M, Ganino T, Chiancone B, Petrucelli R (2018) Identification and characterization of Italian common figs (*Ficus carica*) using nuclear microsatellite markers. Genet Resour Crop Evol 65:1337–1348. https://doi.org/10.1007/s10722-018-0617-6

Roy S, Tyagi A, Shukla V, Kumar A, Singh UM, Chaudhary LB, Datt B, Bag SK, Singh PK, Nair NK, Husain NT, Tuli R (2010) Universal plant DNA barcode loci may not work in complex groups: a case study with Indian Berberis species. PLoS ONE 5:e13674. https://doi.org/10.1371/journal.pone.0013674

Saddoud O, Chatti K, Salhi-Hannachi A, Mars M, Rhouma A, Marrakchim M, Trifi M (2007) Genetic diversity of Tunisian figs (*Ficus carica* L.) as revealed by nuclear microsatellites. Hereditas 144:149–157. https://doi.org/10.1111/j.2007.0018-0661.01967.x

Stover E, Aradhya M, Ferguson L, Crisosto CH (2007) The fig: overview of an ancient fruit. HortScience 42:1083–1087

Tamura K, Stecher G, Kumar S (2021) Molecular evolutionary genetics analysis version 11. Mol Biol Evol 38:3022–3027. https://doi.org/10.1093/molbev/msab120

Tessier C, David J, This P, Boursiquot JM, Charrier A (1999) Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. Theor Appl Genet 98:171–177. https://doi.org/10.1007/s001220051054

UPOV/INF/17/1, 2010. Guidelines for DNA-profiling: molecular marker selection and database construction ("BMT guidelines")

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Vavilov NI (1951) The origin, variation, immunity and breeding of cultivated plants. The Ronald Press Company, NY

Vinson JA, Zubik L, Bose P, Samman N, Proch J (2005) Dried fruits: excellent in vitro and in vivo antioxidants. J Am Coll Nutr 24:44–50. https://doi.org/10.1080/07315724.2005.10719442