IRIS - Archivio Istituzionale dell'Università degli Studi di Sassari

The genetic diversity of selections and wild populations of myrtle revealed by molecular geographic contexts.

Questa è la versione Post print del seguente articolo:

Original
The genetic diversity of selections and wild populations of myrtle revealed by molecular geographic contexts / Mele, Caterina; Corona, Lucia; Melito, Sara; Raggi, Lorenzo; Mulas, Maurizio. - In: INDUSTRIAL CROPS AND PRODUCTS. - ISSN 0926-6690. - 132:(2019), pp. 168-176. [10.1016/j.indcrop.2019.02.018]

Availability:
This version is available at: 11388/220099 since: 2022-05-26T07:16:04Z

Publisher:

Published
DOI:10.1016/j.indcrop.2019.02.018

Terms of use:

Chiunque può accedere liberamente al full text dei lavori resi disponibili come “Open Access”.

Publisher copyright

note finali coverpage

(Article begins on next page)

26 May 2026

The genetic diversity of selections and wild populations of myrtle revealed by molecular geographic contexts

Caterina Mele, Lucia Corona, Sara Melito* Lorenzo Raggi, Maurizio Mulas

Department of Agricultural Sciences, University of Sassari, Via Enrico De Nicola 9, I-07100 Sassari (Italy)
Dipartimento di Scienze Agrarie Alimentari e Ambientali, Università degli Studi di Perugia, Borgo XX Giugno 74, I-06121 Perugia (Italy)

***Corresponding author:** Melito Sara, email: smelito@uniss.it; Office telefone:+39 079229383

Abstract

Myrtus communis L. is an aromatic shrub widely diffused in the Mediterranean area. A growing interest for this species is due to its pharmacological and aromatic properties. This study explored the genetic diversity and population structure of 460 myrtle plants including 46 selections genotypes and 414 wild accessions from Sardinia using simple 14 SSR markers. Genetic diversity parameters showed that 10 markers were sufficiently polymorphic to move forward in the further analysis. The observed heterozygosity ranged from 0.21 (Porto Ferro) to 0.44 (Olia Speciosa), with an average value of 0.32. This parameter significantly varied considering the genetic origin of the plant (selections and natural), and the localization of the myrtle sampling sites (North, Center and South area) ($P < 0.0001$). The selections showed an observed heterozygosity's lower (0.40) compared to the wild group (0.45). Based on the geographic division of Sardinia, North region grouped myrtle accessions with the highest level of genetic diversity (0.45), followed by the Central area (0.43) and the South, that collect the lowest value (0.42). Bayesian clustering analysis divided the myrtle accessions in two principal groups, which are in overlap with the division in selections and natural populations. Significant different distribution of the two clusters was found (Pearson, $P = 0.003$): one cluster was mainly distributed in the South-Center of Sardinia, while the other cluster was predominant the North area. These results and the obtained genetic characterization are a description of an *in situ* germplasm collection, an important step for planning future *ex situ* collection to reinforce the genetic base of breeding programs, allowing to diversify the industrial potential of products derivatives from myrtle.

Keywords Natural populations; selections genotypes, *Myrtus communis*, genetic diversity, habitat distribution.

1. Introduction

Myrtus communis is a diploid (2n = 2x =22) and outcrossing specie with a partial self-pollinating mating system (Mulas and Fadda 2004; Lughadha and Proenca 1996) wild shrub of the Myrtaceae family that widely grows in the Mediterranean Basin (Mendes et al., 2001). Myrtle is considered as one of the of the predominant species of the Mediterranean shrub-land that typically characterises the specific vegetation form called “maquis” (Bruna et al., 2007; Albaladejo et al., 2010). Its value as ornamental and aromatic plant is known since ancient Greeks and Romans. In the last years, a renewed interest on myrtle was discovered because of it’s aromatic and medicinal uses for food and not food industries (Lange 2006). Young leaves as well as mature fruits are both used for the essential oil extraction and for local liquor production (Mulas and Melis 2011; Mulas 2012; Gardelli et al. 2008; Petretto et al., 2016). Both leaves and berries are used as raw material for the essential oil extraction as sources of antioxidants (Tuberoso et al., 2007; Maldini et al., 2016), and liquor production (Mulas and Cani 1999). The success of some food products (as liquor) and the increasing request by food and not food industries of raw material is causing an intensification of the harvesting of leaves and berry’s from the wild plants. The harvest of biomass, the grazing by livestock, the firewood and the urbanization are the principal factors causing the reduction of natural biodiversity of the species, especially in Sardinia (Italy), where the liquor represents an important consume product. The increasing reduction of myrtle populations numbers and sizes, have prompted two requests. Firstly, the creation of a domestication program with candidate cultivar selections for industrial purposes. Several myrtle phenotypic characters has been identified as important traits for food and liquor production, that range from the plant features (big plant size; high green biomass yield; no multiple flowering fluxes) to the berry morphology and production (high berry production; presence of pigmented berries; early ripening time; long peduncle) (Mulas and Cani, 1999; Mulas 2012). Furtherly, the organization of a conservation program in order to preserve the natural genetic diversity of myrtle and plan an efficient and rational exploitation of wild myrtle biomass, without further reduction of the genetic diversity of the specie (Mulas 2012).

The maintaining of high level of genetic diversity provides the basis for survival of populations and species, especially in environments under deep anthropic pressure and potential climatic or environmental changes (Mohammad-Panah et al., 2017). Loss of genetic diversity induces decrease of probability to survive under biotic and abiotic stress condition in their natural habitat with high risk of genetic erosion and species extinction (Lande 1995). Evaluation and conservation genetic diversity program are essential to preserve myrtle species.

Genetic analysis required specific molecular markers able to give a general picture of the state of the genetic diversity and to discriminate which environmental factors could influence the plant diversity distribution (Streiff 1998).

Previous studies have been conducted using ISSR and AFLP as dominant markers (Melito et al. 2013; 2014; 2017).

In the first case, even with a limited number of accessions, population structure analysis, highlighted two principal genetic clusters corresponding to wild and selected myrtle genotypes (Melito et al 2013); however, AFLP markers were used to characterize both a core collection of selected myrtle genotypes based berry production purposes (Melito et al 2014), and a collection of wild myrtle accessions from Sardinia (Melito et al 2017). In these two studies, altitude and geographical origin significantly affect the population genetic structure. In addition, the relationship between genetic variability, morphological and ecological parameter was explored in aromatic species found in the Mediterranean Basin (Rapposelli et al., 2015, Melito et al., 2013, 2014, 2016, 2017). In Sardinia, where myrtle represent an important component of the flora, natural populations have been only partially explored using dominant markers, such as AFLP (Agrimonti et al., 2007; Melito et al., 2014; 2017) and ISSR (Melito et al., 2013a). Population genetic structure of Sardinian myrtle populations appear to be strongly influenced by several parameters such as altitude, geographic location, landscape and sea proximity (Melito et al., 2013; 2014; 2017), while preferred soils are sub-acids or neutrals with a low clay content (Mulas, 2012). Little information are available on the genetic characterization of myrtle populations based on codominant markers isolate only few years ago (Albaladejo et al., 2010). Preliminary studies conducted by Corona et al. 2015 showed that SSR marker were able to differentiate myrtle in different genetic clusters. Previous research have genetically characterized wild accession or myrtle selections using mainly AFLP markers (Agrimonti et al., 2007; Messaoud et al., 2011; Bruna et al. 2007; Melito et al., 2014, 2016, 2017). Little is known about the genetic diversity and population structure of wild and selections of myrtle genotypes from the same geographic area. Melito et al. (2013b), for the first time, reported the result of ISSR characterization of selection and wild populations that included a small number of *M. communis* accessions from Sardinia and Corsica. In this case, the authors observed a clear differentiation between wild accessions and selections collected based on biomass and berry production. The evaluation of the genetic diversity in wild and selected myrtle plants, is extremely important for few reasons: for the constitution of a repository field, obtained from the selection of genotypes based on biomass yield that should maintain the same quality of biomass collected from wild plants; the genetic diversity stored in candidate cultivar collection need to be asses and eventually integrated with missing useful genotypes; finally, a rational conservation program, with the constitution of repository field, needs to be associated to a genetic diversity evaluation (Pressey et al., 2007).

The aims of the present research are: a detailed genetic characterization of wide number of wild accessions and myrtle selections collected from Sardinia using SSR markers; the exploration of the effect of important environmental factors, such as geographic localization and local meteorological conditions (temperatures and rain), on the genetic diversity and population structure.

2. Material and Methods

2.1. Germoplasm collection

A total of 22 *Myrtus communis* populations including 21 wild and 1 group of myrtle selections, were sampled in Sardinia (Italy). About 20 plants for each natural population were collected from 10 localities (Table 1). Wild samples were randomly collected: five plants were randomly sampled along every 20-meters half-alignment for a total of 20 samples per population in order to exclude basal resprouts of the same genotype in the same population. Differences in ecological parameters among the natural sampling sites were explored. Altitude level, geographic localization and climatic conditions were recorded. Meteorological data for each site were provided by Settore Idrografico della Regione Sardegna. The proximal meteorological station to each sample site was used and the average annual climatic parameter are shown in Table~~s~~ 1. Monthly precipitation and temperature (average, maximum and minimum) of historical series (1991-2011) were considered (Supporting Table 1).

Furthermore, 46 selections held in the same field at the Experimental Station of the University Sassari in Oristano, (39° 53'N, 8°37'E) were also analysed. The myrtle selections (Mulas et al., 2002) were originally

picked from different localities in Sardinia based on the berry, biomass production, and the positive attitude for an intensive growing; the selected plants were harvested for agamic propagation and evaluated in a selection program (Mulas et al 2002). The original sampling sites are reported in Table 1. Based on the locality, an identificative code was assigned to wild and domesticated plant as reported in Tables 1.

2.2. DNA extraction and SSR analysis

Total genomic DNA of 460 myrtle samples was extracted. The young leaf tissue, collected during the summer 2015, was obtained using the DNeasy® Plant kit (QIAGEN, Hilden, Germany), following the manufacturer's protocol. DNA concentration and quality was evaluated by agarose 1.0% (w/v) gel electrophoresis.

Fourteen SSR primers were used in this study, following the Albaladejo et al., (2010) protocol (Supporting Table 2). Multiplexed PCR with fluorescent labelled primers (HEX; NED and FAM) were used for the SSR genotyping. The PCR-SSR products were firstly checked in agarose gel before the sequencing at Macrogen, Inc. (Seoul, Korea). Fragment analysis was carried out with Peak Scanner V1.0 (Applied Biosystems) software (Applied Biosystems). Automatic allele calls were subsequently confirmed manually reviewing all electropherograms.

2.3. Genetic analysis

132 POPGENE 1.31 (Yeh et al. 1997) program was used to assess the genetic diversity, as observed and expected
133 heterozygosis observed (*Obs He*, *Exp He* respectively), the effective number of allele for each locus (N_a, N_e), haplotype
134 number and percentage of polymorphic alleles (P). Polymorphic information index was calculated using Cervus 3.0.7
135 (Kalinowski et al. 2007). Linkage disequilibrium (LD) was estimated using POPGENE (Yeh et al. 1997) by means of
136 pairwise linkage tests with 10000 permutations. Pairwise LD was evaluated for the entire myrtle genotype dataset and
137 all the SSR markers used.
138 Fixation index (Fst) was estimated through the Analysis of Molecular Variance using Arlequin 3.1.5 software (Excoffier
139 and Lischer 2010).
140 Population genetic structure was determined using the model-based program, STRUCTURE 2.3.3 software
141 (Pritchard et al. 2000). The number of the most likelihood populations (*K*) capturing the major structure in the myrtle
142 data was determined using Structure Harvester software (http://taylor0.biology.ucla.edu/structureHarvester/) (Earl et al.
143 2012), based on Evanno et al. (2005) method. The software was run using the admixture model adopted different K
144 values ranging from 1 to 10, and evaluated 10 replicate runs per K value. 200.000 Markov Chain Monte Carlo (MCMC)
145 interactions were evaluated. A Burn-in period of 100.000 interactions followed by an additional 500.000 interactions
146 was set. The attribution of each sample to a specific cluster was based on a coefficient of membership ($Q > 0.7$). Input
147 files for the genetic analysis software were prepared using Convert 1.31 software (Glaubitz 2004).To examine patterns
148 of fine-scale genetic structure, molecular data, were used to estimate the coancestry according to Loiselle et al. 1995.
149 For each identified Sardinia area (North, Center and South) coancestry was calculated between all possible pairs of
150 individuals within the subpopulations (max distance 30 m, as from the sampling scheme) and within the whole areas
151 (max distance different for each area according to pupations position) using SPAGeDi software (Hardy & Vekemans,
152 2002). Tests of significance for estimated coancestry values were performed by using randomization procedures to
153 generate populations under the null hypothesis of no spatial genetic structure. For a given distance class, values from
154 399 simulations were used to construct a 95% interval around the null hypothesis of no genetic structure. Coancestry
155 significant values were considered the result of IBD from non-random gene dispersal at the spatial scale examined that
156 is reasonably due to localized pollen and/or seed movement.
157 *2.4. Statistical analysis*
158 Correlations between site and the genetic coefficient of membership (*Q*) were calculated. A Pearson's χ^2 test for 2×2
159 contingency tables was performed for the categorical variables (geographic localization: North, Centre and South; and
160 assigned genetic group based on STRUCTURE analysis: CA and CB) to test correlations with the clusters identified by
161 STRUCTURE (Pritchard et al. 2000). All meteorological variables were normalized before the statistical analysis.
162 Correlations were carried out by JMP 7 software (SAS Institute, Cary, NC, USA). Climatic, geographic and genetic

(Fst) pairwise distance matrices were calculated for the Mantel test using XLSTAT software 2007. The climatic and geographic distance matrices were calculated based on the Euclidean method, while genetic distance matrix was estimated using Fst values (Excoffier and Lischer 2010).

3. Results and Discussion

3.1. SSR analysis and genetic diversity

A total of 460 individuals of myrtle from Sardinia (Italy) were studied (Tables 1). The accessions include 414 wild and 46 myrtle selections from the localities reported in Fig. 1.

Fourteen primers pairs were used (Supplementary Table 2) and 10 of them (Table 2) produced reproducible fragments and were used for the further investigations. Primer *Myrcom* 2, 7, 8 and 10 did not produced a clear data and were not used in the further investigations. The SSR analysis was done on all plant genotypes revealing 68 alleles with a mean of 6.8 allele for each marker. The marker *Myrcom* 11 was the most polymorphic (13 alleles), while *Myrcom* 13 had only two alleles. The average observed heterozygosity (*Obs He*) among all the SSR marker used was 0.23, ranging from zero (*Myrcom* 13) to 0.52 (*Myrcom* 6) (Table 2).

Within the 22 populations the average number of polymorphic loci was 5.89 ranging from 1 (Parco Sette Fratelli and Paulilatino) to 9 (Santa Sofia and Fenosu). The average expected heterozygosity (*Exp He*) was 0.31 with a minimum value of 0.10 (Paulilatino) and a maximum of 0.42 (Santa Sofia) (Table 3). In addition, the *Obs He* showed high variability: from 0.21 (Porto Ferro) to 0.44 (Olia Speciosa), with an average value of 0.32. Myrtle genotype from Paulilatino showed a very high *Obs He* (1.00). In this case, the presence of only one sample induced a high score of this genetic parameter.

These data showed a general high level of *Obs He*, with a significant difference among the 22 myrtle populations (Pearson, $P = 0.005$). Previous study conducted by Melito et al. (2017) in a wide collection of wild myrtle populations showed a lower level of genetic diversity compared to the data found in this study. Differences in *Obs He* could be related to different reasons: first, the different molecular marker used. Melito et al. (2017), in fact, used AFLP markers while in this study the exploration of the genetic diversity was performed through SSR marker, that is a co-dominant marker. Moreover, differentially to previous studies conducted on Sardinian myrtle accessions (Melito et al., 2014; 2017), in this research, both wild and selections genotypes were analyzed. The average *Obs He* varied considering different parameters: the division in selections and wild origin, and the geographical localization of the sampling sites (North, Central and South areas). Significant difference in genetic diversity was found between selections and wild (Pearson, $P < 0.0001$); based on this harboring, as expected, the selections collection showed a lower *Obs He* (0.40) compared to the wild group (0.45). Similar trend was found for the populations division based on the geographic origin

of the sampling sites (North, Central and South area). Minor differences were found among the *Obs He* of the three areas. The North region grouped myrtle accessions with the highest level of genetic diversity (0.45), followed by the central area (0.43) and the South, that collect the lowest *Obs He* (0.42). The average *Obs He* observed in selections and wild myrtle populations is compatible with the principal self-pollination mating system. The highest values of *Obs He* were found in wild myrtle populations suggesting that, despite the out-crossing is the principal mating system (Mulas and Fadda 2004; González-Varo et al., 2010), a degree of cross-pollination occurs in myrtle (Barrett 2013; Cowling 2013).

High variability in haplotype number was also found considering the myrtle genotypes divided in two groups: wild and selections populations. The average number of haplotypes was 32.95 with a minimum of 2 (Paulilatino) and a maximum of 77 (Fenosu). The high number of haplotype found in the selections myrtle collection is the result of the strategy used for the collection program. In this case, plants were selected from wild populations belonging to different Sardinia locations based on plant size, vigor and berry production (Table 1), as a consequence the domesticated population grouped a wide range of myrtle variability as expressed by the high value of haplotype's number.

The average Polymorphic information content (PIC) was 0.48, ranging from 0.12 (*Myrcom* 4) to 0.88 (*Myrcom* 9), revealing the lowest and the highest polymorphic makers. Fifty percent of used SSR markers showed a PIC value higher than 0.5, indicating that these markers are useful for fingerprinting studies. The average Major Allele Frequency (MAF) was 0.68, ranging from 0.35 (*Myrcom* 6) to 0.93 (*Myrcom* 4). Although *Myrcom* 13 presented a MAF of 0.99, considering that the two alleles present only 1bp of difference, this markers should be further investigated and explored in order to acquire more support. The PIC and MAF value indicate a considerable genetic variation among all myrtle collections used.

LD analysis on all the SSR molecular markers used, showed 45 cases of significant LD (α =0.05). Twenty of them showed to be significant ($0.01 \leq P \leq 0.05$), and 25 resulted highly significant ($P < 0.01$). The SSR primer combinations *Myrcom* 4-12/13; *Myrcom* 5-6/14 and *Myrcom* 12-13 had the highest P value.

3.2. Genetic Structure

Population genetic structure was explored using the model-based Bayesian clustering method implemented in STRUCTURE. Evanno et al. (2005) method was applied to identify the principal genetic cluster division. The most likely structure was identified at $K = 2$ (Supplementary Fig. 1). At $K = 2$ the myrtle accessions were separated in two principal clusters (Cluster A and B), capturing 42% (193) and 53% (244) of the individuals, respectively. Little percentage (5%) of admixed genotypes ($Q \leq 0.7$) was detected only in wild populations. Variation in the percentage of individuals assigned to the two clusters was found between wild and selections populations: 52.4% (205) wild

accessions were assigned to cluster B, while 89% (41) was harbored to the cluster A (Pearson, P<0.001). The presence of admixed individual could be associated to a more frequent gene flow between the two clusters in wild populations compared to the selections. As with the increasing of the *K* values, individuals and populations were separated in smaller other genetic clusters (Supplementary Fig 2). The average coefficient of membership for each individual plant and each population in each Structure group is reported in Figure 1A and 1B, respectively. In Figure 1B each population is plotted in the Sardinia map. The clusters A and B were not equally detected in the island: significantly, different distribution of the two genetic clusters was found in the North and South Sardinia (Pearson, P = 0. 0280). Cluster A is in fact predominant in the North, while cluster B showed has a major diffusion in the South (Fig.1 B,C). This geographical oriented distribution of the clusters is even more significant (Pearson, P = 0.003) when a more detailed geographical division of the population sampling sites were done (Fig. 1C). In this case, three geographical areas were identified: North, Central and South areas (Fig. 1C). Based on this division the cluster A was mainly distributed in the South and the center of the Island, while a predominant diffusion of the Cluster B was found in the North area. Overall the genetic cluster identified at *K* = 2 showed that the geographical localization of the samples site had influenced the development of different genetic groups. The geographic trend observed of the two clusters could in part had been influenced by the shape of Sardinia that is longer than wide. Consequently, a more gene flow occurs between East to West than North to South. This observation may be, at least partially, involved in the differential distribution of the cluster A and B. The long distance (North - South) and the mountain localization (in the center of the island) can both have influenced the distribution of the two clusters. These two factors could have act as natural barrier to the myrtle gene flux. As a consequence Cluster B (North area) is isolated the cluster A distributed in center and South area. Previous studies conducted on myrtle candidate cultivar from Sardinia and from Sicily, have shown that geographical distance could induce a different distribution of genetic cluster (Melito et al., 2014; 2016).

3.3. Population divergence, AMOVA and coancestry analysis

Pairwise distance matrix of the genetic diversity, estimated through the Fst value was performed at different levels. At the population level, the average Fst is 0.15 (Table 4). This value is lower to data previously found by Migliore et al., (2013) in a wild *Myrtus nivellei* collection. Despite, these two species (*M. communis* and *M. nivellei*) resulted closely related, *M. novellei* have diverged from *M. communis* since its diffusion in Africa and its colonization of the of Saharan mountains region (Migliore et al., 2012). These two species developed two specific pattern of evolutions from a common ancestor colonizing and surviving to specific environmental conditions associated to the geographic features of the land of diffusion. In addition, beside the different phylogeographic history of the two species, the apparent discrepancy with *M. nivellei* could be associated several other reasons such as the sampling size of these two studies

and the genetic isolation that characterize the Sharan mountain region that promote a high genetic isolation with an higher Fst value (Migliore et al 2013).

Analysis of molecular variance (AMOVA) was used to explore the sources of genetic variance among and at intra population level (Table 4). Few principal parameters, considered potentially important sources of variance were explored: the division in 22 populations, the partition based on STRUCTURE analysis (K = 2), the germplasm origin (wild or selections populations) and finally the myrtle partition based on the geographic origin of each sampling sites (North, Center and South).

AMOVA based on the 22 myrtle populations showed that most of the variance (85.21%) was found "within population" while a lower contribute was given by the among population variance (14.79%). The associated Fst value is 0.15 that is quite low underlining a high exchange of genetic material among the 22 populations. This value is similar to other species belonging to the myrtaceae family, such as *Eucalyptus globulus* (Costa et al., 2017). As in *E. globulus* the Fst value found in this research could be defined moderate, indicating that are still genetic flux among the genotypes considered.

To evaluate the contribution of the clustering partition identified by STRUCTURE, the AMOVA was the run considering the myrtle accessions divided in the two genetic groups identified at K = 2. As shown for the division in 22 populations, the AMOVA considering the K = 2 partition revealed a similar distribution of the genetic variance. Most of the variance was associated to the "within populations" (86.34%) and a minor contribute was related to the "among populations (13.66%). In this case, the Fst (0.14) is quite comparable to the Fst found for the 22 populations.

To complete the exploration of the genetic effect in the distribution of the molecular variance the partition in natural and selections collections was performed. Considering this partition, the within populations (selections and natural) variance had the highest score (96.02%), with a consequent decreasing of the among population variance contribute (3.98%). The associated Fst was decreased (0.04%) comparing to the previous partition. This low Fst data suggest a limited genetic differentiation between the natural and selections genotypes. The selections accessions were selected based on berry production and not based on the genetic diversity. Therefore, considering the relatively recent selection of these myrtle genotypes, this potential candidate cultivar collection has probably conserved a lot of the genetic material from the natural population of origin.

Finally, the geographical area of sampling sites was considered as an environmental factor that could have influenced the molecular variance. Considering the three geographic area, again the same partition of the genetic variance was observed: highest variance associated to the "within geographic area" (95.83%) and lowest for among geographic areas (4.17%). The Fst was found similar to the partition in selections / natural population (Fst = 0.04).

As expected, for all the partition considered, most of the genetic variation is associated to the "intra population level"

287 rather than among them. This data is a consequence myrtle mating system. *Mirtus communis* is mainly an autocrossing
288 specie with a partial self-pollinating mating system (Mulas and Fadda 2004; Lughadha and Proenca 1996). The pollen is
289 driven principally by small insects and coleopters (Melito et al., 2017), that might move for limited distances. This
290 vector of myrtle reproduction explained the high value of intra-population variance along all the sources of variance
291 considered. The low variance based at the "among population" level indicated a reduced population's differentiation.
292 In Table 5 is reported the Fst distance matrix of the 22 populations tested. Telti (TEL) and Santa Sofia (SSF) showed the
293 highest genetic distance value (0.36). TEL and SSF are located respectively in North and Central areas. The
294 geographical distance as well as the presence of different climatic conditions and physical barriers could have induced
295 isolation between these two populations. The presence of the mountain range only in the central part of the Island
296 represents a natural barrier to the gene flow between these two populations.
297 Based on these findings an Fst distance matrix was run considering the partition of population in the three geographic
298 areas of Sardinia (North, Center and South). North and South regions were more divergent (Fst = 0.05) compared to the
299 Central and South region (Fst = 0.04) (data not show). This data is compatible with the predominant distribution of the
300 two genetic cluster CA and CB respectively in Center/South and in the North of the Island.
301 Analysis of fine-scale intra-population genetic structure indicates significant positive autocorrelation among individuals
302 located up to 30-m apart in the three considered areas; the only exception regarded individuals in the last distance class
303 in the South part of Sardinia island (Fig. 2, left). In the spatial window from 0 to 30 m, coancestry decrease is quite
304 clear even if it is never deep. When all the samples of each area (North, Center and South) were compared together,
305 coancestry values decreased quite fast in the first three-four spatial classes, and were not significantly different from
306 zero beyond few kilometers (Fig. 2, right). In particular, they were not significant after about 9 km in the North, 8 in the
307 Centre and 12 in the South. Some negative correlations were observed at high distances in all the studied areas.
308 Obtained results are compatible with the species pollen and seed dispersals systems. Finally, an unexpected, low
309 increase of coaencestry values was observed at 22 and 37 km in Center and South area.
310
311 *3.4. Correlation between genetic and local climate*
312 Several studies have reported that the environmental conditions influence the pattern of genetic variation across plant
313 populations (Jump et al., 2005). Genetic variation and structure has been shown to correlate with environmental factors
314 (which included micro-climatic conditions) in several tree plant species (Savolainen et al 2007). Principal
315 meteorological data, for each collection site was given in Supplementary Table 1. Multivariate analysis was run
316 between the monthly average meteorological data (average, minimum and maximum temperatures, and rain expressed
317 as mm of water) and the coefficient of membership identified at $K = 2$. Pairwise correlation, showed a significant

positive association between the coefficients of membership of the Cluster A and the millimeters of rain of March. Furthermore, a negative correlation was found between the same coefficient of membership and the minimum temperature of the period from February to May and September, showing a possible less thermophilic character of the group (Table 6).

The same analysis was run considering the *Obs He* of each population. The genetic diversity resulted positively correlated to the mm of rain detected during the two summer months of June and July (0.160 and 0.217 respectively), while is negatively correlated to the autumn rain precipitations during the months of September, October and November (-0.108, -0.347 and -0.140 respectively). These results are in agreement with a wide study conducted by Jingfang et al 2018. In this research, the authors have reported a significant correlation between genetic diversity, expressed as *Obs He*, and climatic factors in several woody and herbaceous species. Few climatic factors, such as mean temperatures, precipitation and minimum temperatures etc., have been identified as the most significant for plant genetic diversity. Climatic parameters can deeply influence the genetic diversity of plant species in three principal ways: first, reducing the gene flow affecting the flowering and the pollination (Knight et al., 2005; Ortego et al., 2012; Sork et al., 2010); second the climatic condition could select only the genotype that are able to survive in that specific climatic condition (Davis and Shaw, 2001; Franks and Hoffmann, 2012; Hewitt, 1996); finally climate can select specific alleles, when a population of a specie colonize a new habitat originated by specific climatic conditions (Excoffier and Ray, 2008). Our results suggest that climatic condition have a significant effect on genetic diversity shape of myrtle populations. The correlations found between myrtle *Obs He* and precipitation could be justify considering the specie phenological development. The late flowering time of the species, which close the fruit set in the first weeks of July and may develop fertile seeds with relative abundance when some rains occurs between July and August. On the contrary, if the summer months are completely arid the young fruits drop and the plant produce news flowering fluxes according to the autumn rains support. The late fruits generated by plant re-flowering showed bad maturation, low seed number and strong predation by frugivorous species (Mulas and Fadda 2004).

All the average temperatures are negatively correlated to the *Obs He*. This data suggest that increasing the temperature there is a decreasing of the genetic diversity of myrtle populations. In Table 7 are also reported the average maximum and minimum temperature that are significantly negatively correlated to the *Obs He*.

To explore the relationship genetic divergence among population and environmental factors (geographic and climatic distances matrices) the Mantel test was run (Table 7). A strong positive correlation between genetic and climatic distances matrices was found ($P = 0.0003$) as well as between geographic and climatic data. However, no correlation was detected between genetic and geographic distance matrix. These results are compatible with the strong orographic lack of homogeneity of Sardinia land. Physical barriers, such as mountains, localized in the middle of the island, might

create physical barriers for the gene flow, modifying the genetic exchanges between different geographic areas. The spatial distances between myrtle population has a minor effect in the population genetic differentiation compared to the sampling sites distances. This data is compatible with the identification of few meteorological parameter significantly correlated with the genetic structure and the genetic diversity expressed as *Obs He*, shown in Table 7.

4. Conclusions

This study reported for the first time the exploration of the genetic diversity and population structure of wild and selections of myrtle populations collected in Sardinia, an important island of the Mediterranean Basin. The myrtle selections showed a higher value of genetic diversity in term of number of haplotype comparing to the wild plants. These results give an evidence that the anthropic selection based on phenotypic, biomass and berry production of the myrtle candidate selections, could have a deep impact in the genetic diversity collection. The genetic partition at $K = 2$ highlighted that most of the wild and the early domesticated populations belonged to the two identified clusters. In addition geographic localization of the sampling sites and the climatic conditions of each area significantly affected genetic cluster distribution and the heterozygosity level of myrtle. These findings will be useful to facilitating future strategies of breeding programs.

Acknowledgement

This study was supported by Regione Autonoma della Sardegna (Special Grant L.R. 7/2007-2012: “The myrtle: genetics and metabolomics, an integrated approach for industry development”). The authors would like to thanks Prof. Alfredo Pauciullo for the assistance in the genetic analysis.

Data archiving statement

All primers used in this research followed the research article of Albaladejo et al. (2010).

References

Agrimonti, C., Bianchi, R., Bianchi, A., Ballero, M., Poli, F., Marmiroli, N., 2007. Understanding biological conservation strategies: a molecular-genetic approach to the case of myrtle (*Myrtus communis* L.) in two Italian regions: Sardinia and Calabria. Conserv. Genet. 8, 385–396.

Albaladejo, R.G., Sebastiani, F., González-Martínez, S.C., González-Varo, J.P., Vendramin, G.G., Aparicio, A., 2010. Isolation of microsatellite markers for the common Mediterranean shrub *Myrtus communis* (Myrtaceae). Am. J. Bot. 97(5), e23-e25.

Barrett, S.C.H., 2013. The evolution of plant reproductive systems: How often are transitions irreversible? Proceedings of the Royal Society B, vol. 280, no. 1765, p. 1-8.

Bruna, S., Portis, E., Cervelli, C., De Benedetti, L., Schiva, T., Mercuri, A., 2007. AFLP-based genetic relationships in the Mediterranean myrtle (*Myrtus communis* L.). Sci. Hortic. Amsterdam. 113(4),370-375.

Corona, L., Mele, C., Chessa, I., Mulas, M., 2017. Analysis of Sardinian myrtle (*Myrtus communis* L.) germplasm selections by SSR markers. Acta Hortic. 1172, 165-170.

Costa, J., Vaillancourt, R.E., Steane, D.A., Jones, R.C., Marques, C. 2017. Microsatellite analysis of population structure in *Eucalyptus globulus*. Genome, 60(9).

Cowlin, R.M., Rundel, P.W., Lamont, B.B., Arroyo, M.K., Arianoutsou, M., 1996. Plant diversity in Mediterranean-climate regions. Trends Ecol. Evol. 11(9), 362-366.

Davis, M.B., Shaw, R.G., 2001. Range shifts and adaptive responses to Quaternary climate change. Science. 292(5517), 673–679.

Dettori, C.A., Sergi, S., Tamburini, E., Bacchetta, G., 2014. The genetic diversity and spatial genetic structure of the Corso-Sardinian endemic *Ferula arrigonii* Bocchieri (Apiaceae). Plant Biology 16, 1005–1013.

Earl Dent, A. and von Holdt Bridgett, M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resourc. 4(2), 359-361 doi: 10.1007/s12686-011-9548-7.

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.

Excoffier, L. and Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol.Res. 10(3), 564-567.

Excoffier, L., Ray, N., 2008. Surfing during population expansions promotes genetic revolutions and structuration. – Trends Ecol Evol. 23(7), 347–351.

Franks, S.J., Hoffmann, A.A., 2012. Genetics of climate change adaptation. Annu Rev Genet., 46(1), 185–208.

Gardeli, C., Vassiliki, P., Athanasios, M., Kibouris, T., Komaitis, M., 2008. Essential oil composition of *Pistacia lentiscus* L.: and *Myrtus communis* L.: evaluation of antioxidant capacity of methanolic extracts. Food Chemestry 107, 1120–1130.

Glaubitz, J.C., 2004. Convert: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Mol. Ecol. Notes 4, 309-10.

Gonzales-Varo, J.P., Albaladejo, R.G., Aparicio, A., Arroyo, J., 2010. Linking genetic diversity, mating patterns and progeny performance in fragmented populations of a Mediterranean shrub. J. App. Eco. 47(6), 1242-1252.

Hardy, O.J. and Vekemans X., 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Notes 2, 618-620.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linnean Soc. 58(3), 247–276.

Kalinowski, S.T., Taper, M.L. & Marshall, T.C., 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol. 16, 1099-1106.

Knight, T.M., Steets, J. A., Vamosi, J.C, Mazel, S.J., Burd, M., Cambell, D,R., Dudash, M.R., Johnston, M.O., Mitchell, R.J., Ashman, T.L., 2005. Pollen limitation of plant reproduction: pattern and process. Annu Rev Ecol Evol Syst. 36, 467–497.

Jingfang, T., Jizhong, W., Fangli, L., Feihai, Y., 2018. Relationships between genetic diversity of vascular plant species and climate factors. JoRE. 9(6), 663–672.

Jump, A.S., Peñuelas, J., 2005. Running to stand still: adaptation and the response of plants to rapid climate change. Ecol Lett. 8, 1010–1020.

Lande, R., 1995. Mutation and Conservation. Conserv. Biol. 9(4), 782–791.

Lange, D., 2006. International trade in medicinal and aromatic plants. In: Bogers, R.J., Craker, L.E., Lange, D. (Eds.), Medicinal and Aromatic Plants: Agricultural, Commercial, Ecological, Legal, Pharmacological and Social Aspects. Springer, Dordrecht, 155–170.

Loiselle, B.A., Sork, V.L., Nason, J., Graham, C., 1995. Spatial genetic structure of a tropical understory shrub, Psychotria officinalis (Rubiaceae). Am. J. Bot. 82, 1420–1425.

Lughadha, E.N., Proenca, C., 1996. A survey of the reproductive biology of the *Myrtoideae* (Myrtaceae). Ann Missouri Bot, 480–503.

Maldini, M., Chessa, M.Petretto, G.L.,Montoro, P.,Rourke, J.P.,Foddai, M.,Nicoletti, M.,Pintore, G. 2016. Profiling and simultaneous quantitative determination of anthocyanins in wild *Myrtus communis* L. berries from different geographical areas in Sardinia and their comparative evaluation. Phytochem Anal. 27(5), 249-256.

435 Melito, S., Chessa, I., Erre, P., Podani, J., Mulas, M., 2013a. The genetic diversity of Sardinian myrtle (*Myrtus*
436 *communis* L.) populations. Electron. J. Biotechn. 16, 1–14.
437 Melito, S., Sias, A., Petretto, G.L., Chessa, M., Pintore, G., Porceddu, A., 2013b. Genetic and metabolite diversity of
438 Sardinian populations of *Helichrysum italicum*. PLoS One 8, E79043.
439 Melito, S., Dessena, L., Sale, L., Mulas, M., 2017. Genetic diversity and population structure of wild Sardinian Myrtle
440 (*Myrtus communis* L.) genotypes from different microclimatic areas. Aust. J. Crop Sci. 11(11), 1488-1496.
441 Melito, S., Fadda, A., Rapposelli, E., Mulas, M., 2014. Genetic diversity and population structure of Sardinian myrtle
442 (*Myrtus communis* L.) selections as obtained by AFLP markers. HortScience 49(5), 531–537.
443 Melito, S., La Bella, S., Martinelli, F., Cammalleri, I., Tuttolomondo, T., Leto, C., Fadda, A., Molinu, M.G., Mulas, M.,
444 2016. Morphological, chemical and genetic diversity of wild myrtle (*Myrtus communis* L.) populations in Sicily.
445 Turk. J. Agri. For. 40, 249-261.
446 Mendes, M.M., Gazarini, L.C., Rodrigues, M.L., 2001. Acclimation of *Myrtus communis* to contrasting Mediterranean
447 light environments effects on structure and chemical composition of foliage and plant water relations. Enviro.
448 Exp. Bot. 45, 165–178.
449 Messaoud, C., Khoudja, M.L., Boussaid, M., 2006. Genetic Diversity and Structure of Wild Tunisian *Myrtus communis*
450 L. (Myrtaceae) Populations. Genet. Resour. Crop. Evol. 53(2), 407-417.
451 Messaoud, C., Béjaoui, A., Boussaid, M., 2011. Fruit colour, chemical and genetic diversity and structure of *Myrtus*
452 *communis* L. var. *italica* Mill. morph populations. Biochem. Syst. Ecol. 39, 570– 580.
453 Migliore, J., Baumel, A., Juin, M., Medail, F., 2012. From Mediterranean shores to central Saharan mountains: key
454 phylogeographical insights from the genus Myrtus. J Biogeogr. 39, 942–956.
455 Migliore, J., Baumel, A., Juin, M., Fady, B., Roig, A., Duong, N., Medail, F., 2013. Surviving in mountain climate
456 refugia: new insights from the genetic diversity and structure of the relict shrub *Myrtus nivellei* (Myrtaceae) in
457 the Sahara desert. PLoS ONE 8(9): e73795.
458 Mohammad-Panah, N., Shabanian, N., Khadivi, A., Rahmani, M.S., Emami, A., 2017. Genetic structure of gall oak
459 (*Quercus infectoria*) characterized by nuclear and chloroplast SSR markers. Tree Genet. Genomes 13, 70–81.
460 Mulas, M., Cani, M,R., Brigaglia, N., 1998. Characters useful to cultivation in spontaneous populations of *Myrtus*
461 *communis* L. Acta Hortic. 457, 271–278.
462 Mulas, M., and Melis, R.A.M., 2011. Essential oil composition of myrtle (*Myrtus communis*) leaves. J. Herbs Spices
463 Med. Plants 17(1), 21-34.
464 Mulas, M., 2012. The myrtle (*Myrtus communis* L.) case, from a wild shrub to a new fruit crop. Acta Hortic. 948, 235-
465 242.

466 Mulas, M., Fadda, A., 2004. First observation on biology and organ morphology of myrtle (*Myrtus communis* L.)
467 flower. Agricoltura Mediterranea 134, 223–235.
468 Mulas, M., Francesconi, A.H.D., Perinu, B., 2002. Myrtle (*Myrtus communis* L.) as a new aromatic crop: cultivar
469 selection. Sassari, 30 giugno 1999, 113-121.
470 Ortego, J., Riordan, E,C., Gugger, P,F., Sork, V.L., 2012. Influence of environmental heterogeneity on genetic diversity
471 and structure in an endemic southern Californian oak. Mol Ecol. 21, 3210–3223.
472 Petretto,G.L., Maldini, M., Addis, R., Chessa, M., Foddai, M., Rourke, J.P. , Pintore, G., 2016. Variability of chemical
473 composition and antioxidant activity of essential oils between *Myrtus communis* var. *Leucocarpa* DC and var.
474 *Melanocarpa* DC. Food Chem 197, 124-131,
475 Pressey, R.L., Cabeza, M., Watts, M.E., Cowling, R.M., Wilson, K.A., 2007. Conservation planning in a changing
476 world. Trends in Ecology & Evolution 22(11), 583-592.
477 Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data.
478 Genetics 155(2), 945-959.
479 Rapposelli, E., Melito, S., Barmina, G.G., Foddai, M., Azara, E., Scarpa, G.M., 2015. AFLP fingerprinting and essential
480 oil profiling of cultivated and wild populations of Sardinian *Salvia desoleana*. Genet. Resour. Crop Evol. 62,
481 959-970.
482 Savolainen, O., Volainen, O., Pyhajarvi, T., Knurr, T., 2007. Gene flow and local adaptation in 4 forest trees. Annu Rev
483 Ecol Evol Syst. 38, 595–619.
484 Streiff, R., Labbe, T., Baculieri, R., Steinkellner, H., Glössl, J., Kremer, A., 1998. Within-population genetic structure in
485 *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. Mol Ecol. 7,
486 317–328.
487 Tuberoso, C.I.G., Melis, M., Angioni, A., Pala, M., Cabras, P., 2007. Myrtle hydroalcoholic extracts obtained from
488 different selections of *Myrtus communis* L. Food Chem.101, 806–811.
489 Yeh, F.C., Yang, R.C., Boyle, T.B.J., Ye, Z.H., Mao, J.X., 1997. POPGENE, the user-friendly shareware for population
490 genetic analysis. Molec. Biol. and Biotechn. Center, University of Alberta, Edmonton, Alberta, Canada.
491

Figure legends

Figure 1 A) Barplots (K = 2) of the selections (SEL) and wild myrtle populations (see table 1 for the code); **B)** Average coefficient of membership, based on STRUCTURE results at K = 2. For the selections (SEL) and wild populations (see table 1 for the code associated to each sampling sites) is indicated the average coefficient of membership of Cluster A and B. **C)** Partition of the two genetic clusters (Cluster A and B), identified by STRUCTURE, based on the geographic division of Sardinia in: North and South and North, Center and South.

Figure 2. Correlogram of estimated coancestry for wild *M. communis* plants sampled in North **(A),** Center **(B)** and South **(C)** Sardinia. Left: coancestry values calculated between all possible pairs of individuals in the subpopulations, of the three areas, within 3 m distance classes. Right: coancestry values calculated between all possible pairs of individuals in the three areas within ten distance classes defined in such a way that the number of pairwise comparisons within each distance class was approximately constant. Dashed lines represent upper and lower 95% confidence limits around zero relationship.

Supplementary File

Supplementary Figure 1. Identification of the most probable genetic structure, through the Delta *K* highest value (Evanno et al., 2005).

Supplementary Figure 2. Barplots of the other minor picks identified by Evanno et al. 2005 (*K* = 4; *K* = 6; *K* = 9) of the selections (SEL) and wild myrtle populations (see table 1 for the code).

Supplementary Table 1. Metereological data registrered for each sampling sites. In table are recorded monthly the Average Temperature (Tm), Average maximum temperature (Tmax), Average minimum temperature (Tmin).

Supplementary Table 2. Primers used for the amplification and designed from Albaladejo et al. (2010).

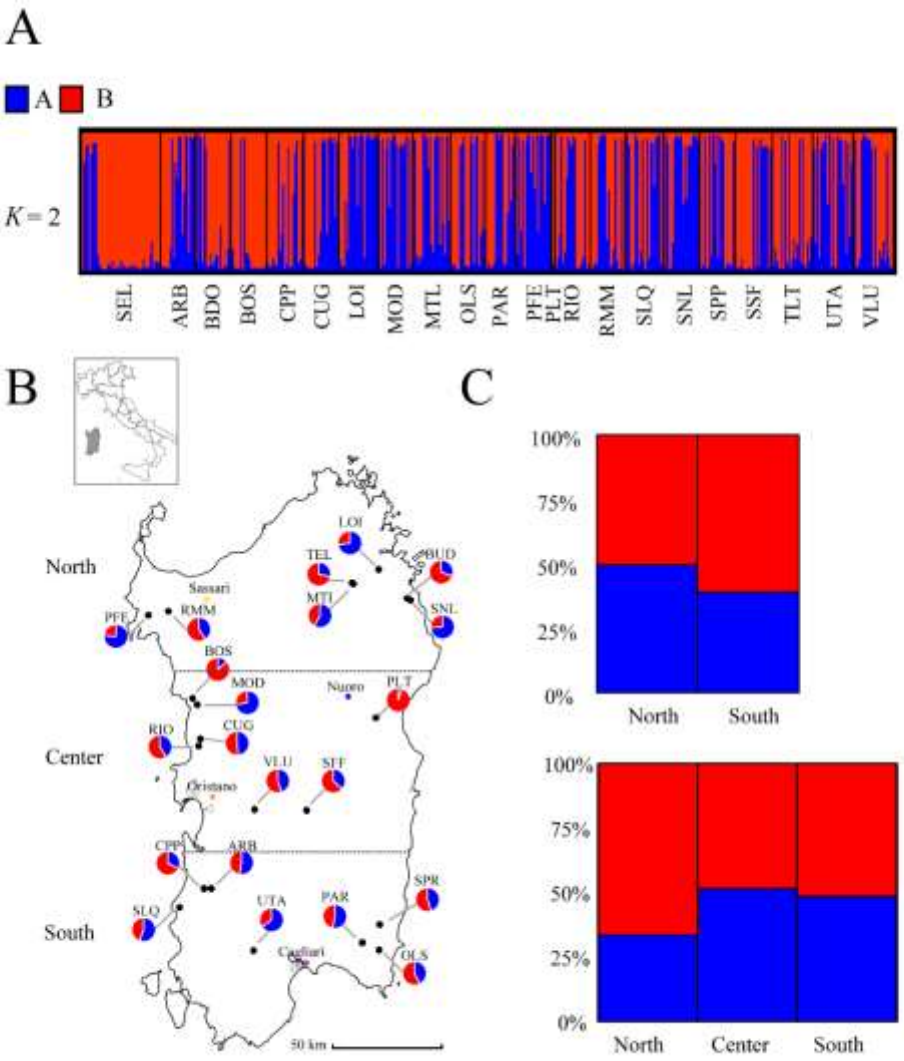

A
A
B
K = 2
SEL
ARB
BDO
BOS
CPP
CUG
LOI
MOD
MTL
OLS
PAR
PFE
PLT
RIO
RMM
SLQ
SNL
SPP
SSF
TLT
UTA
VLU
B
North
Center
South
Sassari
Nuoro
Oristano
Cagliari
50 km
C
100%
75%
50%
25%
0%
North
South
100%
75%
50%
25%
0%
North
Center
South

Figure 1.

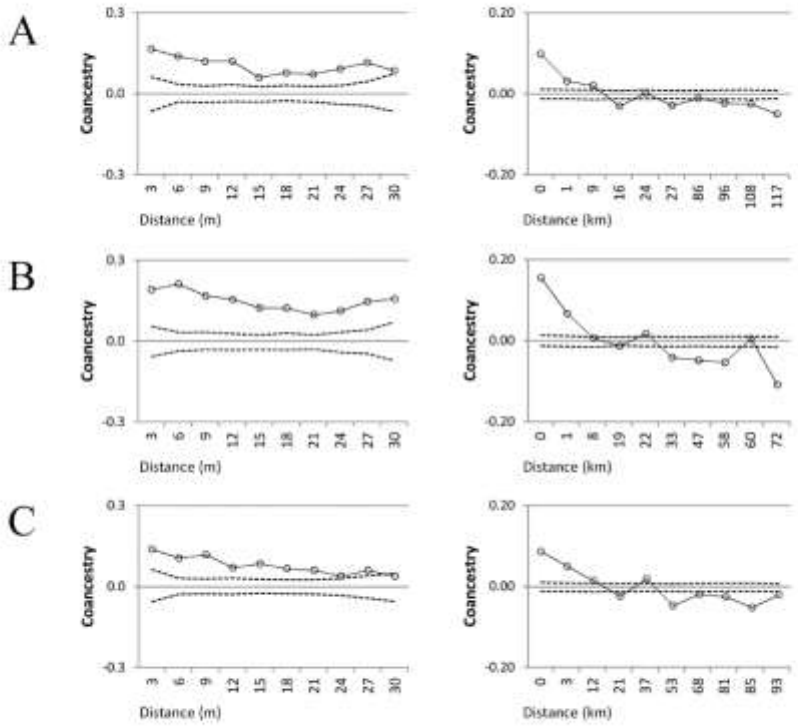

A
B
C
Coancestry
0.3
0.0
-0.3
0.20
0.00
-0.20
3 6 9 12 15 18 21 24 27 30
0 1 9 16 24 27 86 96 108 117
0 1 8 19 22 33 47 58 60 72
0 3 12 21 37 53 68 81 85 93
Distance (m)
Distance (km)

Figure 2.

Tables

Table 1

Myrtle sampling sites characteristics of wild and selected populations. In table are reported: type of plant collection (wild, W and selection, S), locality, the altitude, geographic coordinate of each sampling site (WGS84), the number of plant collected and the associated meteorological conditions for each site. Annual temperatures (°C)(average, T; average maximum, Tmax; average minimum, Tmin) and average annual rain (mm) were reported. For the selection collection the meteorological condition were registered in the experimental station of the University Sassari in Oristano, (39°53'N, 8°37'E), were the myrtles are planted.

Locality	Genotype	Code	Altitude (m.a.s.l.)	Longitude	Latitude	N of plant	T	Tmax	Tmin	Rain
Arbus	W	ARB	271	464500	4375180	20	18.12	22.53	12.61	42.75
Budoni	W	BUD	70	554899	4508931	20	17.28	22.39	12.80	56.50
Bosa	W	BOS	83	456218	4462844	20	16.99	21.33	12.67	50.42
Modolo	W	MOD	103	456218	4462848	20	18.02	22.55	11.04	57.75
Capo Pecora	W	CPP	50	461082	4375180	21	18.02	22.55	11.04	57.75
Cuglieri	W	CUG	235	459720	4444035	20	17.28	23.27	12.19	61.50
Loiri San Paolo	W	LOI	80	541729	4522211	22	15.49	20.53	12.72	65.33
Telti	W	TEL	103	541729	4522213	23	15.78	19.75	12.34	49.25
Monti	W	MTI	506	541729	4522216	21	15.78	19.75	12.34	49.25
Olea Speciosa	W	OLS	82	542915	4346888	20	14.80	20.61	8.93	58.92
Parco Sette Fratelli	W	PAR	422	542915	4346890	17	17.08	22.58	11.56	56.08
Porto Ferro	W	PFE	35	436231	4501232	20	15.19	19.67	10.94	57.92
Rumanedda	W	RMM	45	436231	4501233	20	16.98	22.30	11.67	47.58
Riola Sardo	W	RIO	68	436231	4501235	20	16.98	22.30	11.67	47.58
Siliqua	W	SLQ	68	449519	4366476	21	16.90	22.73	11.23	53.08
Siniscola	W	SNL	40	449519	4366478	21	17.46	24.80	10.13	47.75
San Priamo	W	SPR	59	449519	4366484	20	17.22	23.08	11.33	57.92
Santa Sofia	W	SSF	707	508840	4411011	21	17.91	23.48	12.26	53.92
Uta	W	UTA	151	508840	4411013	23	13.63	19.36	7.87	71.67
Villaurbana	W	VLU	155	508840	4411016	23	17.09	23.11	11.03	42.83
Paulilatino	W	PLT	265	508840	4411019	1	16.09	23.28	11.70	60.67
Asinara	S	SEL1	25	1439500	4493900	1				
Budoni	S	SEL2	70	1559000	4506850	3				
Bosa	S	SEL3	83	1459590	4456380	2				
Cuglieri	S	SEL4	235	1463250	4448970	2				
Isili	S	SEL5	523	1511345	4413070	3				
Laconi	S	SEL6	707	1511350	4413050	6				
Monti	S	SEL7	506	1527500	4517200	2				
Olea Speciosa	S	SEL8	82	1542920	4343130	2	18.12	22.53	12.61	42.75
Orosei	S	SEL9	19	1527500	4517201	4				
Parco Sette Fratelli	S	SEL10	422	1534500	4351175	2				
Rumanedda	S	SEL11	45	1439500	4493900	11				
Sassari	S	SEL12	225	1446430	4478067	1				
Siniscola	S	SEL13	40	1484550	4385190	1				
Telti	S	SEL14	103	1542780	4530430	2				
Uta	S	SEL15	151	1497500	4349950	4				

Table 2

Statistics diversity per locus: n_a = observed number of alleles; n_e = effective number of alleles; *Obs He* = Observed heterozygosity; *Exp He* = Expected heterozygosity, PIC (Polymorphic Information Content), MAF (Major Allele Frequency.

Locus name	Sample size	Alleles			*Obs He*	*Exp He*	PIC	MAF
		Size range (bp)	n_a	n_e				
Myrcom 1	912	210 – 213	4	3.44	0.45	0.71	0.66	0.42
~~*Myrcom 2*~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~
Myrcom 3	908	150 – 160	6	1.80	0.36	0.44	0.40	0.71
Myrcom 4	902	151 – 185	7	1.15	0.12	0.13	0.12	0.93
Myrcom 5	894	252 – 268	8	1.80	0.36	0.44	0.39	0.72
Myrcom 6	912	154 – 167	11	4.25	0.52	0.77	0.73	0.35
~~*Myrcom 7*~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~
~~*Myrcom 8*~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~
Myrcom 9	900	164 – 168	6	2.50	0.04	0.60	0.88	0.59
~~*Myrcom 10*~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~	~~nd~~
Myrcom 11	824	209 – 224	13	3.50	0.20	0.71	0.57	0.46
Myrcom 12	906	146 – 148	4	1.32	0.18	0.24	0.14	0.86
Myrcom 13	908	188 - 189	2	1.01	0	0.01	0.22	0.99
Myrcom 14	888	220 – 277	7	1.70	0.07	0.40	0.70	0.75
Mean	895		6.80	2.24	0.23	0.45	0.48	0.68
St. Dev			3.29	1.13	0.18	0.26	0.27	0.22
Total			68	22.40				

Table 3

Genetic diversity parameters for the selections (S) and wild (W) myrtle population studied. In table are reported the

Population	Origin	NL	NPL	*Obs Het*	*Exp He*	NA
Selection	S	10	9	0.32	0.39	77
Arbus	W	10	8	0.31	0.34	32
Budoni	W	10	4	0.31	0.33	32
Bosa	W	10	7	0.23	0.24	23
Capo Pecora	W	10	4	0.3	0.25	30
Cuglieri	W	10	5	0.3	0.35	34
Loiri Porto San Paolo	W	10	6	0.36	0.39	37
Modolo	W	10	3	0.25	0.29	31
Monti	W	10	8	0.23	0.33	33
Olea Speciosa	W	10	6	0.44	0.29	29
Parco Sette Fratelli	W	10	1	0.23	0.37	32
Porto Ferro	W	10	8	0.19	0.27	26
Paulilatino	W	10	1	1	0.1	2
Riola Sardo	W	10	5	0.27	0.29	29
Rumandedda	W	10	7	0.38	0.38	37
Siliqua	W	10	7	0.27	0.34	36
Sanluri	W	10	5	0.26	0.36	39
San Priamo	W	10	6	0.22	0.31	31
Santa Sofia	W	10	9	0.39	0.42	37
Telti	W	10	7	0.21	0.22	26
Uta	W	10	6	0.26	0.36	42
Villaurbana	W	10	7	0.28	0.32	30
Mean		10	5.86	0.32	0.32	32.95
s.d.		0	2.23	0.16	0.07	12.59

number of loci (NL). number of polymorphic loci (NPL) the observed heterozygosity (*Obs Het*). the expected heteozygosity (*Exp Het*) and the number of haplotypes (NA).

Table 4

Partition of the genetic variance. based on AMOVA analysis. Population genetic structure at $K = 2$. the nature of the genetic material (wild or selections population's). and the geographic origin of the sampling sites were considered as different sources of variance.

Partition	Sources of variation	Df	Sum of square	% of variation	Fst
Overall (22 populations)	Among population	21	224.60	14.79	0.15
	Within population	898	1168.40	85.21	
	Total	919	1392.0		
$K = 2$	Among clusters	1	112.54	13.66	0.14
	Within clusters	918	1478.19	86.34	
	Total	919	1590.73		
Germoplasm origin (wild/selections)	Among wild and selections	1	14.09	3.98	0.04
	Within wild and selections	918	1645.71	96.02	
	Total	919	1659.80		
Geographic origin (North. Center. South)	Among geographic areas	2	50.24	4.17	0.04
	Within geographic areas	917	1609.60	95.83	
	Total	919	1659.84		

Table 5

Pairwise Fst value calculated among the 22 myrtle populations collected in Sardinia All pairwise values are significant different. Selections (SEL) and wild populations (see table 1) were used to build the identity matrix below.

SEL	0.00																					
ARB	0.03	0.00																				
BUD	0.10	0.11	0.00																			
BOS	0.06	0.08	0.15	0																		
CPP	0.06	0.02	0.20	0.10	0																	
CUG	0.03	0.07	0.16	0.16	0.13	0																
LOI	0.08	0.10	0.11	0.16	0.13	0.08	0															
MOD	0.12	0.19	0.08	0.15	0.27	0.16	0.14	0														
MTI	0.06	0.11	0.07	0.10	0.13	0.08	0.06	0.03	0													
OLS	0.18	0.20	0.17	0.29	0.33	0.26	0.26	0.26	0.27	0												
PAR	0.06	0.10	0.13	0.20	0.20	0.10	0.11	0.15	0.12	0.10	0											
PFE	0.06	0.11	0.14	0.09	0.14	0.09	0.11	0.11	0.04	0.31	0.14	0										
PLT	-0.10	-0.03	0.03	0.03	0.02	0.00	0.06	0.04	-0.12	0.26	0.04	-0.1	0									
RIO	0.08	0.15	0.10	0.11	0.21	0.13	0.09	0.07	0.04	0.34	0.19	0.04	0.02	0								
RMM	0.03	0.09	0.11	0.12	0.15	0.06	0.12	0.08	0.07	0.17	0.06	0.11	-0.07	0.12	0							
SLQ	0.06	0.08	0.20	0.10	0.15	0.11	0.11	0.21	0.17	0.22	0.11	0.15	0.16	0.17	0.12	0						
SNL	0.14	0.19	0.12	0.22	0.28	0.21	0.10	0.14	0.14	0.17	0.09	0.22	0.17	0.16	0.15	0.16	0					
SPR	0.05	0.06	0.13	0.14	0.13	0.09	0.10	0.18	0.14	0.09	0.01	0.14	0.07	0.20	0.08	0.05	0.12	0				
SSF	0.14	0.18	0.24	0.30	0.26	0.17	0.18	0.29	0.25	0.24	0.14	0.29	0.16	0.29	0.18	0.20	0.21	0.15	0			
TLT	0.13	0.16	0.13	0.14	0.18	0.16	0.18	0.13	0.05	0.36	0.25	0.06	-0.05	0.11	0.17	0.26	0.30	0.24	**0.36**	0		
UTA	0.08	0.06	0.20	0.08	0.13	0.14	0.15	0.21	0.19	0.24	0.16	0.19	0.13	0.19	0.12	0.02	0.18	0.10	0.21	0.26	0	
VLU	0.10	0.10	0.25	0.14	0.18	0.14	0.18	0.25	0.22	0.28	0.17	0.24	0.16	0.24	0.14	0.07	0.21	0.13	0.22	0.33	0.02	0
	SEL	ARB	BUD	BOS	CPP	CUG	LOI	MOD	MTI	OLS	PAR	PFE	PLT	RIO	RMM	SLQ	SNL	SPR	SSF	TLT	UTA	VLU

Table 6

Pairwise correlation between genetic parameters such as the coefficient of membership (Q) based on STRUCTURE analysis and the observed heterozygosity (*Obs He*) and metereological data. Rain (mm). average temperature tm. average minimum and maximum temperatures (Tmin and Tmax) monthly were used.

Variable	Vs Variable	Correlation	Prob. Sign
Rain March	Average Q (Cluster A)	**0.1109**	0.0204
Tmin February		**-0.0978**	0.041
Tmin March		**-0.1957**	<0.0001
Tmin April		**-0.108**	0.024
Tmin May		**-0.1195**	0.0124
Tmin September		**-0.1041**	0.0296
Rain June	*Obs He*	0.1601	0.0008
Rain July		0.217	<.0001
Rain September		-0.1081	0.0238
Rain October		-0.3465	<.0001
Rain November		-0.1404	0.0033
Tm January		-0.2074	<.0001
Tm February		-0.2697	<.0001
Tm March		-0.21	<.0001
Tm April		-0.2782	<.0001
Tm May		-0.1826	0.0001
Tm June		-0.4484	<.0001
Tm July		-0.095	0.0471
Tm May		-0.1375	0.004
Tm September		-0.2291	<.0001
Tm October		-0.2724	<.0001
Tm November		-0.2285	<.0001
Tm December		-0.2503	<.0001
Tmax January		-0.1989	<.0001
Tmax February		-0.4287	<.0001
Tmax March		-0.176	0.0002
Tmax April		-0.2507	<.0001
Tmax November		-0.195	<.0001
Tmax September		-0.2028	<.0001
Tmax December		-0.1631	0.0006
Tmin January		-0.1678	0.0004
Tmin February		-0.3031	<.0001
Tmin April		-0.2313	<.0001
Tmin August		-0.1509	0.0016
Tmin May		-0.3147	<.0001
Tmin June		-0.2432	<.0001
Tmin July		-0.1251	0.0089

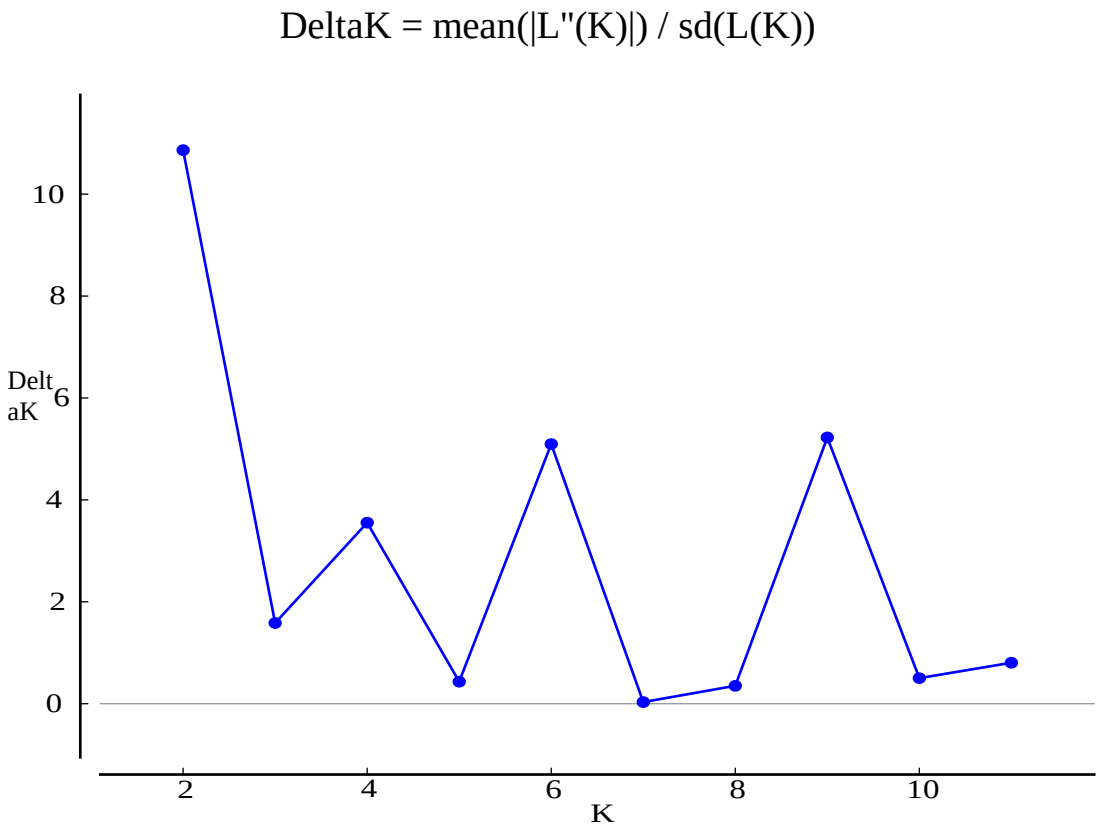
Tmin September	-0.3317	<.0001
Tmin December	-0.1282	0.0073

Table 7

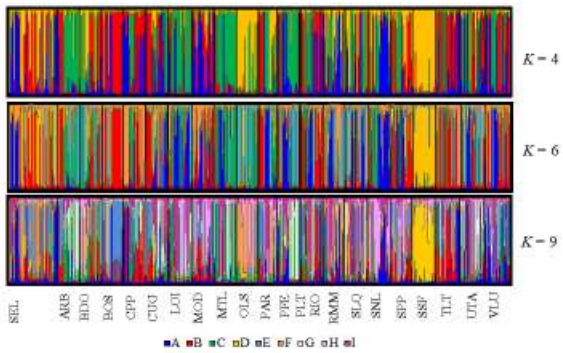
Mantel test to evaluate the correlation between climatic. geographic and genetic distance matrices.

Matrices	Pearson	P value
Climatic vs Geographic	-0.149	0.021
Genetic vs Geographic	n.s.*	n.s.*
Genetic vs Climatic	0.223	0.0003

*n.s. not significant



Supplementary Figure 1 Identification of the most probable genetic structure, through the Delta *K* highest value (Evanno et al., 2005)



Supplementary Figure 2. 1. Barplots of the other minor picks identified by Evanno et al. 2005 ($K = 4$; $K = 6$; $K = 9$) of the selections (SEL) and wild myrtle populations (see table 1 for the code).