Genetic diversity and differentiation processes in the ploidy series of *Olea europaea* L.: a multiscale approach from subspecies to insular populations

C. GARCÍA-VERDUGO,* † M. F. FAY,‡ C. GRANADO-YELA,⁎ † R. RUBIO DE CASAS,⁎ † L. BALAGUER,⁎ G. BESNARD§ ¶ and P. VARGAS†

*Departamento de Biología Vegetal I, Universidad Complutense de Madrid, José Antonio Novais 2, 28040 Madrid, Spain, †Real Jardín Botánico de Madrid, CSIC, Plaza Murillo 2, 28014 Madrid, Spain, ‡Genetics Section, Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK, §Department of Ecology and Evolution, Biophore, University of Lausanne, 1015 Lausanne, Switzerland, ¶Imperial College, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire SL5 7PY, UK

Abstract

Geographical isolation and polyploidization are central concepts in plant evolution. The hierarchical organization of archipelagos in this study provides a framework for testing the evolutionary consequences for polyploid taxa and populations occurring in isolation. Using amplified fragment length polymorphism and simple sequence repeat markers, we determined the genetic diversity and differentiation patterns at three levels of geographical isolation in *Olea europaea*: mainland-archipelagos, islands within an archipelago, and populations within an island. At the subspecies scale, the hexaploid ssp. *maroccana* (southwest Morocco) exhibited higher genetic diversity than the insular counterparts. In contrast, the tetraploid ssp. *cerasiformis* (Madeira) displayed values similar to those obtained for the diploid ssp. *guanchica* (Canary Islands). Geographical isolation was associated with a high genetic differentiation at this scale. In the Canarian archipelago, the stepping-stone model of differentiation suggested in a previous study was partially supported. Within the western lineage, an east-to-west differentiation pattern was confirmed. Conversely, the easternmost populations were more related to the mainland ssp. *europaea* than to the western *guanchica* lineage. Genetic diversity across the Canarian archipelago was significantly correlated with the date of the last volcanic activity in the area/island where each population occurs. At the island scale, this pattern was not confirmed in older islands (Tenerife and Madeira), where populations were genetically homogeneous. In contrast, founder effects resulted in low genetic diversity and marked genetic differentiation among populations of the youngest island, La Palma.

Keywords: archipelago, founder effect, Macaronesia, *Olea europaea*, ploidy level, stepping-stone

Received 7 May 2008; revision received 27 October 2008; accepted 3 November 2008

Introduction

Isolation is a crucial process in the evolution of wild organisms. As the degree of isolation between two given populations increases, gene flow is constrained, allowing other evolutionary forces to operate (Wright 1943; Kimura & Weiss 1964). Two genetic parameters reflect the nature and intensity of the evolutionary change: the genetic diversity, expected to be low in recently established or isolated populations (Hedrick 1999; Maki 2001) and the genetic differentiation, which presumably increases as gene flow diminishes (Mayr 1954; Le Corre & Kremer 1998). The degree of isolation, however, depends on the scale of analysis, and therefore these parameters will vary according to the scale under consideration (Hardy & Vekemans 2001; Grivet & Petit 2002). In this sense, archipelagos offer an ideal geographical framework to study the patterns of genetic variation in wild organisms at different levels of isolation: mainland-archipelago (Inoue & Kawahara 1990; Chiang *et al*. 2006), among islands within an archipelago.
Polyploidy has been shown to be another highly influential component in the evolution of plants, providing remarkable genetic advantages (Solé & Soltis 2000; Ramsey & Schemske 2002). For instance, it is commonly assumed that populations with a higher ploidy have higher levels of genetic diversity (Mallet 2007), especially if they are allopolyploids and their origin is not recent (Ramsey & Schemske 2002; Luttikhuizen et al. 2007). However, this assumption has only been tested in a few cases that analyse two ploidy levels (tetraploids and diploids) (Hardy & Vekemans 2001; Luttikhuizen et al. 2007; but see also Abbott et al. 2007). If genetic diversity was promoted by an increase in ploidy, plants of higher ploidy would theoretically prevent the genetic erosion experienced by small populations in isolation more efficiently than those of lower ploidy (Bever & Felber 1992). A lack of empirical studies analyzing the genetic consequences of isolation in polyploids hinders further discussion of this topic. In addition, the effects of ploidy on the differentiation of closely related taxa have also been recently addressed, with the development of appropriate genetic markers and statistical approaches (Chung et al. 2004; Guo et al. 2005).

The Olea europaea L. complex is an ideal plant group to investigate the effects of isolation in the genetics of populations and taxa. The complex comprises six subspecies of long-lived, outcrossing, woody plants displaying wind-pollination and zoochory (Green 2002), which are suitable features for successful long-distance dispersal (Hess et al. 2000). Three of the subspecies are endemic to the Macaronesian region sensu Suring (1979; but see also Médail & Quézel 1999, for another delimitation of the area). The biogeographical identity of these three taxa is supported by a shared plastid sub-lineage exclusive to this area (Besnard et al. 2007a). On the basis of morphology, no clear differences are found among them, and range distribution is cited as the main criterion for taxon identification (Green 2002). Nonetheless, these subspecies differ in ploidy, from diploid to hexaploid, representing a ploidy series (Besnard et al. 2008; Brito et al. 2008): Olea europaea ssp. maroccana (Greuter & Burdet) P. Vargas et al. (hereafter maroccana) is a hexaploid taxon occupying a total area of a few hundred hectares in the Western High Atlas (Médail et al. 2001); O. europaea ssp. cerasiformis (Webb & Berth.) Kunk. & Sund. (cerasiformis) is a tetraploid endemic to the two Madeiran islands; O. europaea ssp. guanchica P. Vargas et al. (guanchica) is a diploid mainly occurring on six of the seven Canary Islands (Vargas et al. 2001; Green 2002).

Hess et al. (2000) provided a biogeographical framework for inferring the patterns of genetic diversity and differentiation in the Olea lineages endemic to Macaronesia. They presented different conclusions at different spatial scales based on internal transcribed spacer (ITS)-1 and fingerprinting [random amplified polymorphic DNA (RAPD) and intersimple sequence repeat (ISSR)] data: (i) at the subspecies scale, the three Macaronesian subspecies (formerly cerasiformis and laperirina) resulted in three differentiated lineages, (ii) at the Canarian archipelago scale, a stepping-stone model of differentiation from east to west was supported, (iii) strong isolation between populations of different islands suggested that inter-island dispersal has been rare or may not have occurred at all. The main limitation of the study, however, deals with its methodology: an individual-based approach with dominant markers of an unknown polyploid series. Using this approach, analyses of population structure cannot be properly attempted and inferences from neighbor-joining trees could lead to misleading interpretations (Hollingsworth & Ennos 2004; Kosman & Leonard 2005). Thus, no estimations of genetic diversity and no pattern of distribution of such diversity were detailed at any of the spatial scales. In addition, although differentiation between genetic groups (subspecies, islands, populations) was inferred, the degree of such differentiation could not be measured.

In this study, we used dominant [amplified fragment length polymorphism (AFLP)] and codominant [simple sequence repeat (SSR)] markers to study the patterns of genetic diversity and differentiation at three levels of geographical isolation, testing the following hypotheses:

1. At the subspecies scale, we expect that ploidy determined the levels of genetic variation maintained within subspecies. Thus, the hexaploid maroccana may have higher levels of genetic diversity than the tetraploid cerasiformis, and cerasiformis may have higher diversity than the diploid guanchica. Genetic differentiation between subspecies is expected to be limited, as manifested by the moderate morphological differentiation.

2. At the Canarian archipelago scale, we predict that colonization sequence was highly determinant in the distribution of genetic diversity. Consequently, genetic diversity of guanchica populations may be correlated with the geological age of each colonized area, that is, there would be higher genetic diversity in older islands/areas. At this scale, we also test the stepping-stone pattern of differentiation previously suggested using a population-based approach.
At the within-island scale, we expect that time and distance influenced levels of genetic diversity among populations. The complex geological history of each island (temporal scale) and the geographical separation (spatial scale) may have resulted in strong genetic differentiation among populations within each island.

Materials and methods

Study sites and sampling

To represent the natural distribution of each subspecies (maroccana, cerasiformis and guanchica), at least three separate populations were sampled depending on the subspecies distribution (Fig. 1a). Five to 12 individuals were sampled in each population, except for the island of El Hierro, where only three individuals were found (Table S1, Supporting information). Populations of maroccana occur in a restricted area of the southern slopes of the Western High Atlas (Médail et al. 2001), of which three were sampled. Populations of cerasiformis only occur on the southern slopes of Madeira, of which three were sampled (Fig. 1b). A single individual found on the island of Porto Santo was included in the easternmost population from Madeira (STC), given its geographical and genetic proximity (Rubio de Casas et al. 2006). Populations of guanchica were sampled in six of the seven Canary Islands (no wild populations are known on Lanzarote). Following historical records, population sizes and conservation criteria, the 13 most intact populations were sampled across the distribution of guanchica. Additionally, 12 isolated individuals were used for individual-based analysis. The representation of a few individuals from different populations throughout the geographical range of each subspecies is an effective sampling strategy when genetic diversities of taxa are compared (Pons & Petit 1995; Luttikhuizen et al. 2007). Thus, we randomly chose two-11 individuals per population for the SSR analysis (Table S1).

The cultivated olive tree (O. europaea ssp. europaea var. europaea) has been introduced in the island of Gran Canaria (Hess et al. 2000) and is also cultivated in Morocco (Médail et al. 2001; Ancochea et al. 2004). My, millions of years; Ky, thousands of years.
(et al. 2001). Since gene flow between wild and cultivated trees has been reported within subspecies of *O. europaea* (Besnard & Bervillé 2000), forms suspected to be introgressed by cultivars were avoided by choosing characteristic populations of *guanchica* on the basis of the fruit size and leaf morphology (Green 2002). The individuals in the selected populations produced small fruits (less than 1 cm long) and leaves that clearly resemble those of *guanchica* (see Rubio de Casas et al. 2006). In addition, ecological characteristics (remoteness from crops, optimal climatic conditions, inaccessible locations) were considered.

**DNA extraction and fingerprinting protocols**

Total genomic DNA was extracted from c. 0.03 g silica gel dried leaf material with the Plant DNeasy Minikit (QIAGEN Inc.), following the manufacturer’s protocol.

Standard AFLP protocols following Vos et al. (1995) with modifications by Schönswetter et al. (2003) were used. As 2C values of the *O. europaea* subspecies range from 2.93 (in diploids) to 7.88 pg (in the hexaploid *marocccana*) (Besnard et al. 2008) the standard 3 + 3 selective bases protocol was used as recommended in Fay et al. (2005). A trial comparing the same samples at different concentrations resulted in no difference in either height of the chromatogram peaks or number of fragments (data not shown). Thirty primer pairs were initially screened for their potential to produce scorable fragments. Of these, four primer pairs (EcoRI/MseI plus ACC/CCT, ACT/CAC, AAC/CTT and AGG/CAT) were chosen according to reproducibility, levels of fragment polymorphism and previous analysis (Rubio de Casas et al. 2006). Two to three replicates from the restriction–ligation phase were included in all the subsequent reactions, and DNA samples from different extractions of the same individuals were also included (complete replicates) to test reproducibility (Chung et al. 2004). In order to avoid subjective interpretation in fragment scoring, all AFLP profiles were scored under an extraction code, and their identity was revealed after scoring the whole data set. Thus, scoring tests allowed us to determine the reproducibility of AFLP reactions and accuracy in scoring at the same time. Fragments were scored from 140 to 450 bp to minimize the occurrence of fragment size homoplasy, more likely with shorter fragments (Vekemans et al. 2002). The reactions were separated on a 7% polyacrylamide gel using an ABI 3100 automated sequencer. GeneScan version 2.1 and GenoTyper version 2.0 (Applied Biosystems) were used to score the alleles. The MAC-PR method (Esselink et al. 2004) was employed to obtain the allele-dosage of each polyploid individual in six of the seven SSR loci (we did not manage to succesfully apply the MAC-PR method in DCA13).

**Data analyses**

The first level of analysis was the regional delimitation, focusing on the three Macaronesian subspecies, with the samples covering the distributions of the subspecies. At this level, we first conducted two analyses with no assignment of samples to populations or taxa. These approaches allowed us to explore the pattern of differentiation among subspecies. A principal coordinate analysis (PCoA) based on scored AFLP and SSR genetic phenotypes was performed as implemented in GENALEX 6 (Peakall & Smouse 2006). A neighbor-joining dendogram with SSR individual phenotypes based on the distance measure of Nei & Li (1979) was constructed using PHYLIP version 3.6 package (Felsenstein 2005). In addition, we also conducted analyses with each sample assigned to a given population. To explore the partitioning of the genetic variance into different geographically and genetically distinguishable groups, we analysed the populations using SAMOVA (Dupanloup et al. 2002). Various SAMOVA were run, increasing the number of K groups until the percentage of explained variance among groups reached a limit. After genetically homogeneous groups were established, we determined the genetic diversity for each regional group, by using Shannon’s index (for AFLPs) and the allelic richness (for SSRs). Shannon’s index was calculated as indicated in Abbott et al. (2007), and allelic richness corrected for sample size was calculated using the method described in Petit et al. (1998).

Sharing of private AFLP fragments and SSR alleles among subspecies was calculated, excluding those only scored in less than three individuals per subspecies (rare
fragments or alleles). Such analyses allow the detection of shared ancient polymorphisms and dissection of the genetic contribution of close relatives to polyploid taxa (Guo et al. 2005; Paun et al. 2006). We also followed a population-based approach using $F_{ST}$ coefficients (Gömöry et al. 2007). Samples of subspp. europaea, laperrinei and cuspidata were grouped by taxon, and considered as artificial populations in the analyses to provide external groups. Pairwise $F_{ST}$ values between populations were calculated by using AFLPSurv version 1.0 (Vekemans et al. 2002) and null-allele frequencies were estimated following a Bayesian approach (Zhivotovsky 1999). We used 10 000 bootstrap replicates to determine branch support in the consensus tree. Genetic diversity estimates (expected heterozygosity, $H_j$), standard errors and percentage of polymorphic loci for each population were calculated as implemented in the software.

At the Canarian archipelago scale, the distribution of within-population diversity throughout the archipelago was investigated by correlating within-population diversity for each population ($H_j$) with the age of the last volcanic activity in each area (Fig. 1c). To determine the patterns of genetic differentiation across the archipelago, a Bayesian mixture analysis was performed in BAPS version 5.0 (Corander et al. 2006), including only AFLP profiles of europaea and guanchica. Posterior admixture analysis was conducted for the mixture results as recommended in Corander & Marttinen (2006).

At the within-island scale, a fine-scale analysis for each island was performed to infer the contribution of geographical distances and geological events (volcanic activity) in the distribution of genetic diversity. We used the age of the last volcanic event as an estimate for population age, following Goméz et al. (2003, and references therein). In order to determine the correlation between genetic [expressed as $F_{ST}/(1 - F_{ST})$] and population-age divergences in each island, we performed Mantel tests on the islands that had at least three populations with different geological ages (Tenerife and La Palma, Fig. 1c). Similarly, pairwise genetic and geographical distance (kilometres) matrices were analysed by means of a Mantel test for those islands that had at least three populations (Tenerife, La Palma and Madeira). The level of significance was assigned after 1000 permutation tests, as implemented in GENALEX 6 (Peakall & Smouse 2006), and the slope of each linear function was interpreted as the differentiation pattern for that island, that is, the greater the slope, the greater the pattern of differentiation by distance within that island.

Results

AFLP and SSR polymorphisms

The four AFLP primer combinations yielded 490 scorable fragments, 97.8% being polymorphic. Reactions were highly reproducible (overall mismatch rate = 0.048; Table S2, Supporting information) and AFLP profiles from different extractions showed a low mismatch rate (0.078). Any DNA fragments that were differently scored between two replicated samples (17 positions) were removed from the analysis. The hexaploid maroccana displayed a higher number of fragments per individual for all primer combinations (54 on average; Table S2). The seven SSR primer combinations yielded a total number of 124 alleles. Number of alleles per subspecies, excluding those alleles present in only one individual, were: 94 in maroccana, 50 in cerasiformis, 62 in guanchica and 35 in europaea.

The regional scale analysis

Exploratory PCoA analyses based on AFLPs and SSRs rendered similar results, but AFLPs gave greater resolution than SSRs, probably due to the higher number of both loci and samples included in the former. The subspecies maroccana and cerasiformis formed genetically individualized clusters. Axis 1 accounted for a high percentage of variance (40% in AFLP and 28% in SSR analyses; Fig. 2) and clearly separated maroccana and cerasiformis from the rest of the samples. Individuals of europaea and guanchica from easternmost populations were placed close to each other in the AFLP analysis. Samples from Tenerife, La Gomera, La Palma and El Hierro were positioned at the extreme edge of that group. High diversification within guanchica was revealed since these samples were spread along the two PCo dimensions in both analyses. The neighbor-joining tree based on SSR phenotypes was congruent with the pattern of differentiation provided by the PCoA (Fig. S1, Supporting information): maroccana and cerasiformis were sister groups, and samples from europaea and guanchica (easternmost islands) were included in the same branch and these were separated from a group of Gran Canaria and another formed by western guanchica samples.

The SAMOVA of AFLP phenotypes performed on the populations showed that the partitioning of the variance among groups reached a limit when the number of groups equalled four: populations of maroccana (AL, ZA, IMO; group 1), populations of cerasiformis (STC, CDL, PDS; group 2), and a split within guanchica in such a way that populations from the easternmost islands (TIN, BET, BC; group 3) were again separated from those occurring on the western islands (AN, GU, ER, TE, VH, AR, SJ, FA, CAL, FG; group 4). A high proportion (24%) of the AFLP variation was partitioned among these groups, and 9% of the variation among populations revealed a certain population structure (Table 1).

Table 2 summarizes the estimates of genetic diversity within the sampled populations (2A) and within the genetically homogeneous groups previously revealed by SAMOVA.
At the population level, the highest values of Shannon’s index (0.190–0.203), expected heterozygosity ($H_j = 0.25–0.26$), and percentage of polymorphic loci (64.7–75.7%) were found in $maroccana$ populations. The population TIN ($H_j = 0.27$), which is the closest to the mainland, reached similar values (Table 2). The lowest values of diversity were found at the western limit of the Canarian archipelago (on the island of La Palma), with a maximum value of 0.167 (Shannon’s index) and 0.19 ($H_j$) in population FA and a minimum of 0.083 (Shannon’s index) and 0.11 ($H_j$) found in population CAL. Populations of $cerasiformis$ showed similar values of genetic diversity estimates to those reported for $guanchica$ populations (0.153–0.188, Shannon’s index). Both AFLP and SSR gave similar estimates of genetic diversity in the groups determined by samova (regional genetic diversity, Table 2B). Regardless of the fingerprinting technique, hexaploid $maroccana$ exhibited the highest levels of genetic diversity, whereas tetraploid $cerasiformis$ and the two groups of diploid $guanchica$ showed similar values of diversity. Although this result appears to be incongruent with the representation provided by the PCoA for AFLP data (Fig. 2), it should be noted that the latter is an exploratory technique for visualizing genetic similarity among samples and does not properly reflect the genetic diversity of a given group. Eastern $guanchica$ showed a higher genetic diversity (0.189) than western $guanchica$ (0.178) according to Shannon’s index, probably due to the high variation reported for TIN population.

Despite the genetic distinction among subspecies revealed in previous analyses, sharing of private fragments and alleles established a close relationship between the hexaploid $maroccana$ genome and those of $guanchica$ and $cerasiformis$. Thus, 22% of the AFLP fragments and 29% of the SSR alleles of the hexaploid were exclusively shared with both $guanchica$ and $cerasiformis$ (Fig. 3). Fifteen fragments (9% of the total scored for $maroccana$) and 17 alleles (16% of the total) were exclusively shared by $maroccana$ and $cerasiformis$. Twenty-one fragments (13%) and 14 alleles (13%) were found to be characteristic of $guanchica$ and $maroccana$. Conversely, the highest proportion of exclusively shared fragments and alleles found in the tetraploid $cerasiformis$ and the diploid $guanchica$ was due to $maroccana$ (Fig. 3). Pie charts also revealed the rich genomic diversity of $maroccana$, which displayed private markers in 30% of the total SSR alleles and 12% of the AFLP fragments. Additionally, $guanchica$ shared with $europaea$ the highest proportion of

---

**Table 1** Results of hierarchical analysis of molecular variance (SAMOVA) conducted on AFLP profiles of 19 populations representing the Macaronesian subspecies of Olea. Four groups (‘$maroccana$’, ‘$cerasiformis$’, ‘east $guanchica$’ and ‘west $guanchica$’) were revealed by the analysis. Tests of significance were based on 1023 permutations.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among groups</td>
<td>3</td>
<td>1959.08</td>
<td>15.98</td>
<td>24%</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Among populations</td>
<td>15</td>
<td>1398.84</td>
<td>5.77</td>
<td>9%</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>149</td>
<td>6418.75</td>
<td>43.08</td>
<td>66%</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
exclusive fragments (7%) and alleles (5%). The different percentages between AFLP and SSR analysis could be due to the different nature of the markers (dominant vs. codominant; total DNA vs. single locus screening) and the sample size in each case.

The relationships that can be inferred from the population-based analysis of pairwise \( F_{ST} \) (Fig. 4) are consistent with those revealed by our previous analyses. First, *maroccana* and *cerasiformis* were sister groups (81% bootstrap) and second, their separation from the clade *europaea + guanchica* was well supported (81% bootstrap). Finally, a clear east–west differentiation is revealed in western *guanchica*. Individuals of *europaea* were closely related to those of *guanchica* from Gran Canaria and Fuerteventura, splitting *guanchica* into two different groups with medium to strong support (Fig. 4).

**The Canarian scale analysis**

We found a significant correlation between geological age and within-population diversity of AFLPs across the Canarian archipelago (Spearman correlation \( R = 0.67, P < 0.05 \)). BAPS analysis recognized three different genetically homogeneous groups within the archipelago (Fig. 5). Posterior admixture analyses revealed active gene flow between these three clusters, being more pronounced at their limits. Thus, cluster 1 comprises samples of *europaea* and the easternmost populations (Fuerteventura and Gran Canaria), although some gene flow from cluster 2 was revealed by admixture analysis. Populations of Tenerife were assigned to cluster 2, although AN contained at least two individuals closely related to cluster 1, and TE was remarkably related to clusters 1 and 3. The island of La Gomera is geographically close to Tenerife, and genetically was intermediate between clusters 2 and 3, showing a mixed origin for the two sampled populations (VH and AR).

**The within-island scale analysis**

Individual Mantel tests of AFLPs supported a significant relationship between population age-divergences and pairwise genetic differentiation in the island of La Palma (\( R^2 = 0.99, P < 0.05 \)), whereas no significant relationship was found within the island of Tenerife (\( R^2 = 0.07, P > 0.05 \)). When the Mantel test was performed between geographical distance and pairwise genetic differentiation, a strong linear pattern of differentiation by distance among the populations within the three islands was confirmed (Fig. 6a). Tenerife and Madeira presented similar linear

---

**Table 2** Estimates of genetic diversity based on AFLPs within each population (A) and within the genetically homogeneous groups (B) determined.

<table>
<thead>
<tr>
<th>Subspecies</th>
<th>Population (island)</th>
<th>( N )</th>
<th>( P ) (percentage)</th>
<th>Shannon’s index</th>
<th>( H_j )</th>
<th>B) Regional genetic diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>guanchica</td>
<td>TIN (Fue)</td>
<td>5</td>
<td>63.9</td>
<td>0.212</td>
<td>0.27</td>
<td>east guanchica</td>
</tr>
<tr>
<td>guanchica</td>
<td>BET (Fue)</td>
<td>9</td>
<td>65.7</td>
<td>0.181</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>BC (GC)</td>
<td>12</td>
<td>68.8</td>
<td>0.187</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>AN (Ten)</td>
<td>12</td>
<td>66.3</td>
<td>0.179</td>
<td>0.21</td>
<td>west guanchica</td>
</tr>
<tr>
<td>guanchica</td>
<td>GU (Ten)</td>
<td>6</td>
<td>65.2</td>
<td>0.179</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>ER (Ten)</td>
<td>12</td>
<td>69.2</td>
<td>0.194</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>TE (Ten)</td>
<td>6</td>
<td>62.9</td>
<td>0.194</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>VH (Go)</td>
<td>12</td>
<td>72.0</td>
<td>0.185</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>AR (Go)</td>
<td>12</td>
<td>53.1</td>
<td>0.171</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>SJ (Pal)</td>
<td>12</td>
<td>51.8</td>
<td>0.152</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>FA (Pal)</td>
<td>12</td>
<td>53.5</td>
<td>0.167</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>CAL (Pal)</td>
<td>7</td>
<td>42.0</td>
<td>0.083</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>guanchica</td>
<td>FG (Hie)</td>
<td>3</td>
<td>46.1</td>
<td>0.152</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>cerasiformis</td>
<td>STC (Mad)</td>
<td>10</td>
<td>68.8</td>
<td>0.188</td>
<td>0.23</td>
<td>cerasiformis</td>
</tr>
<tr>
<td>cerasiformis</td>
<td>CDL (Mad)</td>
<td>7</td>
<td>54.1</td>
<td>0.153</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>cerasiformis</td>
<td>PDS (Mad)</td>
<td>8</td>
<td>61.8</td>
<td>0.161</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>maroccana</td>
<td>IMO</td>
<td>12</td>
<td>75.7</td>
<td>0.203</td>
<td>0.25</td>
<td>maroccana</td>
</tr>
<tr>
<td>maroccana</td>
<td>ZA</td>
<td>6</td>
<td>67.6</td>
<td>0.199</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>maroccana</td>
<td>AL</td>
<td>5</td>
<td>64.7</td>
<td>0.190</td>
<td>0.26</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: \( N \), sample size; \( P \), percentage of polymorphic loci; \( H_j \), expected heterozygosity based on AFLP; *allelic richness was corrected by sample size following the method described in Petit et al. (1998) for \( g = 10 \).
functions \((1.1 \times 10^{-3})\) and \((1.6 \times 10^{-3})\), respectively), whereas La Palma clearly showed a steeper slope \((15.0 \times 10^{-3})\). Fig. 6b shows the levels of genetic diversity within populations, similar among the populations of Madeira and Tenerife and higher than those found in La Palma.

**Fig. 4** Neighbour-joining tree based on AFLP pairwise \(F_{ST}\) differences between populations, as implemented in **aflpsur**. Subspecies **cuspida** served as the outgroup taxon. Numbers beside nodes indicate bootstrap support percentage recovered after 10 000 replicates. Populations are coded as in Fig. 1.

**Fig. 5** Results of the admixture analysis performed with **baps** based on the AFLP genetic phenotypes scored for **europaea** and **guanchica** samples. Mixture analysis resulted in three clusters that were used to perform the subsequent admixture analysis as indicated in Corander & Marttinen (2006). Each vertical bar corresponds to an individual, with its population coding below and its subspecies/island affiliation above, and each of the three clusters is represented by a different colour.
Insights into genetic diversity of polyploids and subspecies differentiation

Dominant (AFLPs) and codominant (SSRs) markers have been highly consistent in the estimation of genetic diversity and the inference of subspecies differentiation (PCoA, sharing of private fragments/alleles), suggesting that a strong genetic signal can be revealed regardless of the inheritance of the marker. Both techniques detected higher levels of genetic diversity in hexaploid maroccana than in its tetraploid and diploid relatives, supporting our first hypothesis. Nevertheless, our expectations of genetic diversity were not fully confirmed in the Olea europaea complex, since estimates of genetic diversity in tetraploid cerasiformis were similar to those estimates for diploid guanchica (Table 2). These two subspecies exhibited moderate levels of genetic diversity, falling in the heterozygosity range 0.20–0.25 (for dominant markers) usually described for outcrossing, long-lived endemics (Nybom 2004). Several examples illustrate that founder effects and a restricted distribution area produce a depauperation in the genetic pool of populations (Mayr 1954; Maki 2001; Nielsen 2004), even in polyploid organisms (Maki et al. 1996; Buza et al. 2000). In the case of tetraploid cerasiformis, almost restricted to a single island, we hypothesize that polyploidy has probably buffered the genetic depauperation due to isolation. As a result, estimates of genetic diversity were not remarkably low, but lower than we expected for the tetraploid. In diploid guanchica, gene flow from europaea appears to be a factor promoting genetic enrichment in the easternmost islands, as revealed in several analyses.

All the analyses performed supported a marked genetic differentiation of the three Macaronesian subspecies. Whatever the assumptions (individual samples with no population assignment vs. population partitioning) or the data sets (AFLPs vs. SSRs), the three subspecies were clearly differentiated from each other. Samova revealed strong geographical and genetic discontinuities among the three taxa, with a high proportion of the AFLP variation partitioned among groups/subspecies (Table 1). These results are compatible with limited or even no gene flow among populations of these subspecies for a long period of time. The current pattern of isolation could date from the Pliocene (Besnard et al. 2007a), allowing taxa to diverge allopatrically since then. Although wild olive trees show remarkable mechanisms for long-distance gene flow by pollen (wind pollination) and seed (zoochory) (Rubio de Casas et al. 2006), geographical distances separating these subspecies may have been successful barriers. Recent studies have demonstrated that pollen dissemination in wind-pollinated plants is often limited to short distances (Koening & Ashley 2003; Besnard et al. 2007b, and references therein), and long-distance seed dispersal has probably been limited among these areas (Hess et al. 2000). In addition to geographical isolation, we cannot rule out the likelihood of sexual isolation by ploidy (for reviews see Ramsey & Schemske 2002; Mallet 2007).

Limited morphological differentiation among these subspecies of O. europaea (Médail et al. 2001; Green 2002) is, however, associated with a clear genetic distinction, as has been described within other species displaying different ploidy levels (Albach 2007; Luttikhuizen et al. 2007). In the case of Olea, different ploidy and strong geographical isolation probably resulted in relatively ancient processes of genetic divergence and subtle morphological forms. Thus, analyses of shared private fragments and alleles (Fig. 3) supported the sharing of ancient polymorphisms rather than recent gene flow among the subspecies. The high similarity observed between the guanchica and cerasiformis genomes (Fig. 3) could be explained by the sharing of a common ancestor, from which cerasiformis may have arisen by autopolyploidy. Plastid DNA analyses support such a possibility (Besnard et al. 2007a, but see also the phylogenetic signal of ITS sequences). An autotetraploid
origin may be therefore inferred as an alternative explanation for the moderate levels of genetic diversity found in *cerasiformis*. In addition, the PCoA (Fig. 2), figures reported in the pie charts (Fig. 3), the $F_{st}$ tree (Fig. 4) and relationships revealed by plastid analysis between the three Macaronesian subspecies (Bensard et al. 2007a) suggested a relatively ancient allopolyploid origin for the hexaploid *maroccana*, with *cerasiformis* and *guanchica* (or continental, extinct relatives) lineages probably being involved in this process. According to this result, allopolyploidy is a plausible reason for the high levels of genetic diversity found in *maroccana*. The potential allopolyploid origin of this taxon has implications in the interpretation of such genetic diversity. Since disomic inheritance is expected in allopolyploids, allelic frequencies cannot be truly determined, and measures of genetic diversity (heterozygosity) and some indices for genetic differentiation (e.g. $F_{st}$) are not reliable (De Silva et al. 2005; Obbard et al. 2006). As genetic diversity was compared among subspecies by means of indices with no inheritance assumption (i.e. Shannon’s index and allelic richness), and genetic differentiation was assessed by several congruent analyses, our approach supports these conclusions. Although these methods have been used to study genetic diversity and differentiation among ploidy levels (Abbott et al. 2007; Kloda et al. 2008), further discussion on the causal factors affecting diversity cannot be properly addressed (Obbard et al. 2006). In any case, markers more suitable for studying hybridization events (e.g. low-copy genes) should be explored further in order to clarify the processes involved in the origin of these polyploids.

**Differentiation pattern and distribution of diversity across the Canarian archipelago**

Unlike the regional scale, gene flow has promoted genetic admixture over the oceanic barriers between the six Canarian islands, as shown by the Bayesian analyses (Fig. 5). This result contradicts to some extent the strong inter-island isolation proposed by Hess et al. (2000). In addition, the individual-based analyses and the $F_{st}$ tree (Fig. 4) supported a clear stepping-stone model of differentiation in the western part of the archipelago but not for the easternmost islands. According to these analyses, populations from the easternmost islands (Fuerteventura and Gran Canaria) were more closely related to *europaea* than to *guanchica* from the westernmost islands (Figs 4 and 5). This pattern explains the genetic proximity found in a previous study between *guanchica* samples from eastern populations and *europaea* (Lumaret et al. 2004). Two possible explanations could be responsible for such a pattern. First, although we were extremely cautious in population sampling, gene flow from cultivated trees (*europaea* var. *europaea*) into wild *guanchica* populations may have taken place. Introggression from cultivars could explain the high levels of genetic variation found within population TIN. However, Hess et al. (2000) did not find genetic similarities between *guanchica* and olives cultivated on the eastern islands (no olive orchards are known on the western islands). On the other hand, a natural introduction from mainland *europaea* populations (var. *sylvestris*) is also plausible. Gran Canaria suffered a recent volcanic eruption (c. 300 Ky) that mainly affected the eastern part of the island (Ancochea et al. 2004), where most of the wild olives currently occur. Lack of the private fragments in population BC that were scored on both neighbouring islands (data not shown), and the recent geological origin of this area, suggest a recent event of colonization. Additional sampling of mainland populations and cultivars from the Canary Islands could clarify the split within *guanchica*.

Our analyses showed that genetic diversity in populations of *guanchica* was significantly correlated with the last volcanic activity in the area in which each population occurs. Wild olive populations displayed diversity values across the archipelago from 0.27 (TIN, 18.7 million years since last volcanic event) to 0.11 (CAL, 3000–4000 years). Despite being significant, the coefficient of correlation ($R = 0.67$, see results) showed that within-population diversity was not totally explained by the factor ‘date since last volcanic event’. Thus, other factors not considered in the correlation will also contribute to determine such genetic diversity, mainly the genetic exchange among populations. For instance, the FG population (380 000 years), on the island of El Hierro, exhibited heterozygosity values (Table 2) similar to those reported for AN (8–6 million years) and AR (4.5 million years). Dispersal events from different source populations to El Hierro could enhance within-population diversity regardless of the young age of this area. Such an explanation is congruent with the high plastid haplotype diversity of *Pinus canariensis* (Gómez et al. 2003) and *Cistus monspeliensis* (M. Fernández-Mazuecos & P. Vargas, unpublished) on this island.

**Genetic differentiation within an island**

Two different situations arise from the analyses performed on three oceanic islands (Fig. 6). Madeira and Tenerife, in which wild olive populations were sampled over similar distances (c. 40 and 60 km, respectively) displayed similar rates of genetic differentiation by distance ($1.1 \times 10^{-3}$ and $1.6 \times 10^{-3}$, respectively). Particularly, populations in Tenerife did not show a pattern of differentiation determined by geological divergence among populations, in contrast to that of another endemic wind-pollinated Canarian endemic, *P. canariensis* (Gómez et al. 2003). In the absence of conspicuous barriers, populations can be genetically connected, even over long distances if sufficient time is involved (Gömöry et al. 2007). Populations on Tenerife had areas available for colonization over the last 8 million years...
The expectation is frequently assumed in theoretical studies. However, it does not necessarily imply higher genetic diversity, although such a pattern has been observed in studies of the formation of Madeira (5 million years; Rubio de Casas et al. 2006). Despite different timing of colonization, genetic homogenization has been achieved within Tenerife and Madeira, suggesting that internal geographical barriers and geological events have not been effective causes of population isolation in either subspecies. It has been suggested that human impact in some Canarian forest species since c. 2000 years BP could have resulted in remarkable vegetation changes on Tenerife (de Nascimento et al. 2007). While the effect of human activities is less likely associated with the natural pattern of distribution described for quenchica (i.e., isolation by distance), a tangible implication of human impact appears to be related to the highly fragmented distribution and scarcity of present-day populations. Similar genetic diversity and low genetic differentiation among these populations lead us to reject our initial hypothesis (time and distance effect on genetic diversity), and instead to envisage a scenario in which wild olive populations were formerly more conspicuous and connected within each island (Rodríguez & Marrero 1990).

A different pattern is observed on the young island of La Palma. Regardless of their origin (either from other islands or from populations on La Palma itself), our populations could not have colonized the studied area before 20,000 years ago due to its recent volcanic origin (Carracedo et al. 2001; Fig. 1c). This may account for the markedly low genetic diversity within populations on La Palma (cf. Nybom 2004), and the differentiation pattern strongly explained by geological divergence among populations ($R^2 = 0.99$). The most striking example of La Palma was that of population CAL. A small population of 20 individuals (C. García-Verdugo, personal observation) occurs in an area colonized less than 3,000–4,000 years ago (Carracedo et al. 2001), and our analyses revealed the lowest heterozygosity values (0.11) in the study. In addition, founder effects on La Palma appear to be responsible for a great genetic differentiation by distance (Hedrick 1999), up to 10 times higher than those found in the other islands. In another wind-pollinated species, Marquardt & Epperson (2004) reported $F_{ST}$ values of 0.005 between Pinus strobus populations separated by 2 km, whereas our populations, separated by 16 km, displayed $F_{ST}$ values of 0.18. Contrary to olders islands (Tenerife and Madeira), volcanic events in La Palma may have led to recent colonization, and patterns of within-population diversity and among-population differentiation determined by geological age.

Conclusions

The comparison between three ploidy levels of closely related taxa revealed that higher ploidy levels per se do not necessarily imply higher genetic diversity, although such expectation is frequently assumed in theoretical studies. Neutral processes, such as genetic drift or founder effects, and the pathway of formation of polyploid taxa can determine the acquisition and maintenance of that assumed high diversity, as our results suggested. In this study, we found that an alloployploid origin of maroccana is likely to be related to the high genetic diversity found in this continental hexaploid, in contrast to similar diversity levels in the tetraploid (cerasiformis) and diploid (quenchica) subspecies.

In the Canary archipelago, gene flow from the mainland and the volcanic processes in the region have been shown to be factors involved in the distribution of genetic diversity within quenchica. As evidenced by a stepping-stone model of differentiation across the western islands and the gene flow reported between eastern and western islands, the ability of long-distance dispersal has been a successful feature to overcome inter-island isolation. Consequently, partial connectivity by long-distance dispersal has prevented inference of dramatic evolutionary changes of lineages within this taxon, contrary to most of the taxa previously studied in the archipelago.

Within-island analysis have revealed a similar pattern between two insular subspecies occurring in different archipelagos: an isolation-by-distance pattern was described among populations on older islands (Madeira and Tenerife), and we therefore inferred that gene flow is also a predominant mechanism in the pattern of differentiation at this scale. Populations occurring on a young island (La Palma), however, are largely affected by recent geological events, and show higher among-population differentiation and lower genetic diversity. In this latter case, genetic differentiation within the western quenchica lineage might be enhanced in a few generations if gene flow remains restricted in this island.

Within-island analyses should be further conducted at different spatial scales in other endemics also displaying wide geographical distributions. Fine-scale studies of species in which populations have not diverged to the point of speciation provide valuable information in interpreting factors promoting early evolutionary changes in oceanic archipelagos.

Acknowledgements

The authors are indebted to all those people who made possible the sampling of populations, especially M. I. Blanch Matute, A. Justo, P. de la Nuez and R. Domínguez. C. Rodríguez, E. Cano (RJBM), and R. Smith (Kew Gardens) provided valuable assistance in the lab. C. García-Verdugo sincerely thanks A.B. Kelly for his hospitality during his stay at Kew Gardens. J. C. Carracedo provided useful geological data, and Borja (Biosfera), A. Fernández (Garajonay National Park), A. Santos Cabildo and J. A. Bermejo (Excmo. Cadbildo de La Palma) provided information regarding the location of populations. We also would like to thank M.E. Pérez-Corona (UCM), E. Manrique (CCMA), J.A. Delgado and
References


C. García-Verdugo, R. Rubio, C. Granado, L. Balaguer and P. Vargas are particularly interested in the evolution of Mediterranean plants. G. Besnard is focused on olive phylogeography and the molecular bases of C4 photosynthesis evolution. M. Fay’s research interests are conservation genetics, phylogenetics and genome evolution.

Supporting Information

Additional supporting information may be found in the online version of this article:

**Fig. S1** Neighbour-joining dendogram based on Nei-Li distances among individual SSR phenotypes.

**Table S1** List of location of populations and number of sampled individuals (N) for every conducted analysis (AFLPs and SSRS), providing subspecies, ploidy level, island, coordinates, collectors/DNA source and voucher code. Total sampling size and total sampling size for every endemic subspecies is detailed. **“**denotes isolated individuals which were included in the individual-based analyses.

**Table S2** AFLP primer combinations selected for the analyses, results from tests of reproducibility and number of fragments per individual displayed in each subspecies. Name of each subspecies is abbreviated as follows: maroccana (mar), cerasiformis (cer), guanchica (gua), europaea (eur), laperrinei (lap) and cuspidata (cus).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.