Genetic diversity and geographic differentiation in the alternative legume *Tripodion tetraphyllum* (L.) Fourr. in North African populations

E. Bellucci, L. Nanni, E. Bitocchi, M. Rossi & R. Papa

Dipartimento di Scienze Ambientali e delle Produzioni Vegetali, Università Politecnica delle Marche, Ancona, Italy

**Keywords**

*Anthyllis tetraphylla*; conservation; genetic differentiation; inter-simple sequence repeats; Leguminosae; Mediterranean.

**Correspondence**

R. Papa, Dipartimento di Scienze Ambientali e delle Produzioni Vegetali, Università Politecnica delle Marche, Via Breccie Bianche, 60131 Ancona, Italy.
E-mail: r.papa@univpm.it

**Editor**

O. Loudet

Received: 7 January 2010; Accepted: 3 May 2010

doi:10.1111/j.1438-8677.2010.00370.x

**ABSTRACT**

Wild legumes constitute an important component of widespread pastures in the Mediterranean basin. This region is experiencing remarkable effects from climate change, and continuous monitoring of species and population dynamics is important in order to plan and enact valuable conservation programmes. *Tripodion tetraphyllum* (L.) Fourr. [=*Anthyllis tetraphylla* L.] (2n = 16), belongs to the tribe Loteae (Fabaceae), and could be very important for soil protection and sward improvement in abandoned or degraded Mediterranean areas. This alternative pasture legume is very closely related to *Lotus japonicus* and has some important characteristics for survival of the species in difficult and overgrazed Mediterranean areas. In this study, we have investigated the molecular diversity and population structure of *T. tetraphyllum* from North Africa using ISSR markers and plastidial microsatellites. To date, this is the first study concerning the genetic diversity and geographic differentiation of *T. tetraphyllum*. Ninety genotypes from three North African countries were analysed according to ISSRs, cpSSRs and one phenotypic trait. *T. tetraphyllum* shows a clear geographical structure, with differentiation associated with longitudinal differences; moreover, there is a general reduction in genetic diversity from Morocco to Tunisia. With all the markers used, strong differentiation was seen among collection sites. Our data highlight a genetic diversity gradient and cline of distribution, indicating that *T. tetraphyllum* has extended its area of distribution from Morocco to Tunisia.

**INTRODUCTION**

Arid and semi-arid land makes up approximately one-third of the land surface of the Earth. Such areas are expanding, which probably represents another indication of global weather change (Dregne 1986; Archibald 1995; Stafford Smith 1996; Dore 2005). Consequently, much productive land is being degraded and livestock productivity is increasingly suffering from feed shortage.

Wild legumes constitute a crucial component of widespread pastures in the Mediterranean basin. These plants are an important source of fat, protein and minerals for sheep, horses, goats and cows, and can represent up to 80% of total forage material, especially for sheep. Wild legume species in the Mediterranean area occupy about 40% of the grazed meadows (Viano et al. 1995). Pasture and forage legumes also have important roles in arable land where, especially in dry areas, they have the potential to replace fallow in cereal-producing farming systems through their ability to fix nitrogen (Bennett & Cocks 1999). The alternative legumes deserve more attention, and investigations need to be carried out to determine their potential roles for multiple uses in Mediterranean environments.

Among the wild legumes, Loteae are widely distributed throughout the Mediterranean basin. *Tripodion tetraphyllum* (L.) Fourr. (2n = 16) belongs to the Fabaceae, tribe Loteae; it is a predominantly selfing, annual prostrate legume with yellow flowers (Couderc 1980), that occurs on calcareous soils from sea level to 1000 m a.s.l. (Pignatti 1982), and has good palatability (Salsano 1996). Although little is known about this alternative pasture legume, which is usually neglected and regarded as unimportant, *T. tetraphyllum* may be very important for soil protection and sward improvement on slopes and in abandoned or degraded Mediterranean areas (such as rocky soils, steep slopes with very shallow soils, sandy salt soils). Furthermore, this species has ornamental importance because of its peculiar swollen calyx that persists until seed maturity and its corolla colour (Meloni et al. 2000). *T. tetraphyllum* also has important characteristics for survival of the species in difficult and over-grazed Mediterranean areas: the production of numerous pods and seeds, small seed size, hard seeds, early flowering and prostrate growth.

Using Internal Transcribed Spacer (ITS) and Chloroplast Simple Sequence Repeat (cpSSR) markers, *T. tetraphyllum* [=*Anthyllis tetraphylla* L.] was shown to be closely related to *Lotus japonicus* (Nanni et al. 2004), which, together with *Medicago truncatula*, was selected as a model plant species. *L. japonicus* is the focus of large, multinational genome projects and is the second legume species to be sequenced.
Genetic diversity and geographic differentiation in T. tetraphyllum

Bellucci, Nanni, Bitocchi, Rossi & Papa

(Young et al. 2005). Thus, in the near future, the literature on the ecological genetics and molecular ecology of wild species closely related to L. japonicus is expected to grow rapidly, as was the case for Arabidopsis (Bevan & Walsh 2005).

There are many morphological structures used for taxonomic classification in Papilionaceae (including Loteae) that can distinguish T. tetraphyllum from Anthyllis, so that the accepted name for this species is T. tetraphyllum, with the two synonyms, Anthyllis tetraphyllum L. and Physanthes tetraphyllum (L.) Boiss. (Roskov et al. 2005). T. tetraphyllum has dehiscent (as does Lotus) and non-lomentumaceous fruits, and the fruit structure is leguminous, two-seeded, pubescent and lomentumaceous (Lassen 1986). When studying pollen morphology, Diez & Ferguson (1990) confirmed that T. tetraphyllum differs significantly from all of the other species of Anthyllis in the size of its pollen grains. In addition, in a study of several Loteae species (Young et al. 2005), the importance of the direction for development of conservation guidelines and strategies of more general interest. To the best of our knowledge, the present study is the first concerning the population structure of T. tetraphyllum, with the number of individual plants.

The accessions collected from ICARDA reflect a broad range of geographical distributions of T. tetraphyllum in North Africa, from Morocco to Tunisia. There are 16 Moroccan accessions, from 16 collection sites distributed across four provinces (Beni Mellal, Khenifra, Marrakech and Tetouan), with altitudes ranging from 40 m a.s.l. (Tetouan, accession MTTe) to 1420 m a.s.l. (Beni Mellal, accession MdBMa). Algeria is represented by 10 accessions, from 10 different collection sites in six provinces (Algiers, Constantine, Medea, Miliana, Setif and Tlemcen), with altitudes from 150 m a.s.l. (Constantine, accession ACOa) to 1000 m a.s.l. (Miliana, accession AMId). Tunisia is represented by nine accessions from nine collection sites in six provinces (Beja, Bizerte, Jendouba, Kasserine, Sousse and Zaghouan), with altitudes from 5 m a.s.l. (Bizerte, accession TBBb) to 670 m a.s.l. (Kasserine, accession TKAa). The Portuguese accession (POR) obtained from the USDA was collected from the Elvas Plant Breeding Station (PT), 300 m a.s.l. For each accession, four seeds were sown and the plants were grown and cultivated at the SAPROV Department, University Politecnica delle Marche, Italy. Table 1 also summarises the number of individual plants for each accession used for the DNA isolation, and the subsequent molecular analyses. Total genomic DNA was isolated from young leaves of single individual plants using the DNeasy Plant Mini kit (Qiagen GmbH, Hilden, Germany), following the protocols provided by the manufacturer. The final DNA quality and concentrations were determined by running the samples on 0.8% agarose gels, and DNA concentrations were standardised to 5 ng µl⁻¹.

ISSR analysis

DNA amplifications were achieved using five different ISSR primers (listed in Table 2). These were selected after initial screening of a set of about 20 primers, because they gave unambiguous and reproducible intra-population and inter-population patterns. The ISSR PCRs were performed in 25 µl reaction mixtures containing 25 ng genomic DNA, 1× PCR buffer (Epicentre), 2× Enhancer (Epicentre), 2 mM MgCl₂ (Epicentre), 200 µM dNTPs (Epicentre), 15 pmol primer and 1 U Taq polymerase (Epicentre). Amplification was performed in a Perkin Elmer 9600 thermal cycler (Norwalk, CT, USA). The PCR reaction Touchdown programme consisted of: one cycle at 95 °C for 3 min; one cycle at 94 °C for 1 min; annealing for 1 min; and 72 °C for 2 min. In the followed 10 cycles, the annealing temperature was reduced for each cycle by 0.5 °C to the optimal annealing temperature. Then, 26 cycles at 94 °C for 1 min; annealing for 1 min; and 72 °C for 2 min. In the followed 10 cycles, the annealing temperature was reduced for each cycle by 0.5 °C to the optimal annealing temperature. Then, 26 cycles at 94 °C for 1 min; annealing for 1 min; and 72 °C for 2 min. Details of amplifications for each primer are given in Table 2. The reaction products were separated by electrophoresis on 1.5% agarose gels in a 0.5× Tris–borate–EDTA (TBE) buffer, stained for 15 min in ethidium bromide, destained for 15 min, and visualised on a UV transilluminator.

MATERIAL AND METHODS

Plant material and DNA isolation

We studied 36 accessions of T. tetraphyllum, of which the 35 from North Africa were obtained from the International Centre for Research in Dry Areas (ICARDA, Aleppo, Syria) and one accession from Portugal was obtained from the USDA (USDA,ARS,National Genetic Resources Programme, Germplasm Resources Information Network – GRIN). The accessions from ICARDA consisted of samples obtained by bulking seeds from several plants (on average 11 plants) collected in each site. Overall, the DNA of 94 individual plants of T. tetraphyllum, (on average three per accession) was obtained. Table 1 provides accession names, donors, states, provinces, altitudes and geographical coordinates for each accession analysed, and the number of individual plants.

For each accession, four seeds were sown and the plants were grown and cultivated at the SAPROV Department, University Politecnica delle Marche, Italy. Table 1 also summarises the number of individual plants for each accession used for the DNA isolation, and the subsequent molecular analyses. Total genomic DNA was isolated from young leaves of single individual plants using the DNeasy Plant Mini kit (Qiagen GmbH, Hilden, Germany), following the protocols provided by the manufacturer. The final DNA quality and concentrations were determined by running the samples on 0.8% agarose gels, and DNA concentrations were standardised to 5 ng µl⁻¹.

ISSR analysis

DNA amplifications were achieved using five different ISSR primers (listed in Table 2). These were selected after initial screening of a set of about 20 primers, because they gave unambiguous and reproducible intra-population and inter-population patterns. The ISSR PCRs were performed in 25 µl reaction mixtures containing 25 ng genomic DNA, 1× PCR buffer (Epicentre), 2× Enhancer (Epicentre), 2 mM MgCl₂ (Epicentre), 200 µM dNTPs (Epicentre), 15 pmol primer and 1 U Taq polymerase (Epicentre). Amplification was performed in a Perkin Elmer 9600 thermal cycler (Norwalk, CT, USA). The PCR reaction Touchdown programme consisted of: one cycle at 95 °C for 3 min; one cycle at 94 °C for 1 min; annealing for 1 min; and 72 °C for 2 min. In the followed 10 cycles, the annealing temperature was reduced for each cycle by 0.5 °C to the optimal annealing temperature. Then, 26 cycles at 94 °C for 1 min; annealing for 1 min; and 72 °C for 2 min. Details of amplifications for each primer are given in Table 2. The reaction products were separated by electrophoresis on 1.5% agarose gels in a 0.5× Tris–borate–EDTA (TBE) buffer, stained for 15 min in ethidium bromide, destained for 15 min, and visualised on a UV transilluminator.
Chloroplast microsatellite (cpSSR) analysis

A screening with three pairs of chloroplast SSR universal primers, previously used by Nanni et al. (2004) in Anthyllis spp. and by Angioi et al. (2009a,b) in Phaseolus vulgaris, was conducted for all of the accessions of T. tetraphyllum examined with ISSR markers. The primers used (Table S1) were designed based on chloroplast sequences of Nicotiana tabacum from Weising & Gardner (1999) to be amplified across a broad range of species. The cpSSR PCRs were performed in total volumes of 25 μl containing 25 ng genomic DNA, with the reaction mixture and PCR conditions as described in Nanni et al. (2004). After addition to each sample (2.5 μl) of 2.5 μl loading buffer (98% formamide, 2% dextran blue, 0.25 mM EDTA), the PCR products were resolved on 6% polyacrylamide gels run at 100 W constant power for 1.5 h, and then visualised using GENOMIX (Beckman, CA, USA).

Phenotypic characterisation

Each individual of each accession was analysed for colour of the epicotyls. One month from sowing, the T. tetraphyllum plants showed the absence or presence of anthocyanic colouration at the epicotyl level. On the basis of epicotyl colour, three discrete phenotypic classes were identified: absence of colour (class 0), presence of light-red coloration (class 1) and intense red colouration (class 2).

Table 1. Details of the T. tetraphyllum accessions analysed in the present study.

<table>
<thead>
<tr>
<th>accession name</th>
<th>donor</th>
<th>original accession code</th>
<th>state</th>
<th>province</th>
<th>altitude m a.s.l.</th>
<th>geographical coordinates</th>
<th>no. of genotypes analysed per accession</th>
<th>no. of genotypes per province</th>
<th>no. of genotypes per state</th>
</tr>
</thead>
<tbody>
<tr>
<td>MdBMa</td>
<td>ICARDA</td>
<td>14357</td>
<td>Morocco</td>
<td>Beni Mellal</td>
<td>1420</td>
<td>W06 30 N32 03</td>
<td>2</td>
<td>12</td>
<td>36</td>
</tr>
<tr>
<td>MBMb</td>
<td>ICARDA</td>
<td>14360</td>
<td>Morocco</td>
<td>Beni Mellal</td>
<td>1070</td>
<td>W06 27 N32 10</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBMc</td>
<td>ICARDA</td>
<td>14364</td>
<td>Morocco</td>
<td>Beni Mellal</td>
<td>820</td>
<td>W06 44 N32 02</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBMd</td>
<td>ICARDA</td>
<td>14369</td>
<td>Morocco</td>
<td>Beni Mellal</td>
<td>1280</td>
<td>W06 32 N31 55</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBMe</td>
<td>ICARDA</td>
<td>14370</td>
<td>Morocco</td>
<td>Beni Mellal</td>
<td>530</td>
<td>W06 20 N32 35</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBMf</td>
<td>ICARDA</td>
<td>14371</td>
<td>Morocco</td>
<td>Beni Mellal</td>
<td>770</td>
<td>NA</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MKHa</td>
<td>ICARDA</td>
<td>14372</td>
<td>Morocco</td>
<td>Khenifra</td>
<td>1300</td>
<td>W05 12 N33 31</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>MMAa</td>
<td>ICARDA</td>
<td>14359</td>
<td>Morocco</td>
<td>Marrakech</td>
<td>350</td>
<td>W07 49 N33 00</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>MMAb</td>
<td>ICARDA</td>
<td>14362</td>
<td>Morocco</td>
<td>Marrakech</td>
<td>300</td>
<td>W07 34 N32 58</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMAc</td>
<td>ICARDA</td>
<td>14363</td>
<td>Morocco</td>
<td>Marrakech</td>
<td>390</td>
<td>W07 34 N32 44</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMAd</td>
<td>ICARDA</td>
<td>14368</td>
<td>Morocco</td>
<td>Marrakech</td>
<td>350</td>
<td>W07 36 N33 00</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTEa</td>
<td>ICARDA</td>
<td>14358</td>
<td>Morocco</td>
<td>Tetouan</td>
<td>220</td>
<td>W05 19 N35 10</td>
<td>4</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>MTEb</td>
<td>ICARDA</td>
<td>14361</td>
<td>Morocco</td>
<td>Tetouan</td>
<td>140</td>
<td>W05 24 N35 35</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTEc</td>
<td>ICARDA</td>
<td>14365</td>
<td>Morocco</td>
<td>Tetouan</td>
<td>60</td>
<td>W06 00 N35 34</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTEd</td>
<td>ICARDA</td>
<td>14366</td>
<td>Morocco</td>
<td>Tetouan</td>
<td>100</td>
<td>W05 36 N35 04</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTEe</td>
<td>ICARDA</td>
<td>14367</td>
<td>Morocco</td>
<td>Tetouan</td>
<td>40</td>
<td>W05 27 N35 41</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAla</td>
<td>ICARDA</td>
<td>14348</td>
<td>Algeria</td>
<td>Alger</td>
<td>350</td>
<td>E02 35 N36 27</td>
<td>4</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>ACOa</td>
<td>ICARDA</td>
<td>14356</td>
<td>Algeria</td>
<td>Constantine</td>
<td>150</td>
<td>E06 42 N36 47</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AMEa</td>
<td>ICARDA</td>
<td>14346</td>
<td>Algeria</td>
<td>Medea</td>
<td>720</td>
<td>E03 40 N36 17</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AMEb</td>
<td>ICARDA</td>
<td>14347</td>
<td>Algeria</td>
<td>Medea</td>
<td>825</td>
<td>E02 53 N36 07</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMia</td>
<td>ICARDA</td>
<td>14349</td>
<td>Algeria</td>
<td>Miliana</td>
<td>650</td>
<td>E02 20 N36 19</td>
<td>2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>AMib</td>
<td>ICARDA</td>
<td>14350</td>
<td>Algeria</td>
<td>Miliana</td>
<td>550</td>
<td>E02 15 N36 12</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMic</td>
<td>ICARDA</td>
<td>14351</td>
<td>Algeria</td>
<td>Miliana</td>
<td>700</td>
<td>E02 12 N36 06</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMid</td>
<td>ICARDA</td>
<td>14352</td>
<td>Algeria</td>
<td>Miliana</td>
<td>1000</td>
<td>E02 10 N36 00</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASea</td>
<td>ICARDA</td>
<td>14354</td>
<td>Algeria</td>
<td>Setif</td>
<td>950</td>
<td>E04 23 N36 06</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>ATLa</td>
<td>ICARDA</td>
<td>14353</td>
<td>Algeria</td>
<td>Tlemcen</td>
<td>395</td>
<td>W01 40 N34 46</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>TBea</td>
<td>ICARDA</td>
<td>111290</td>
<td>Tunisia</td>
<td>Beja</td>
<td>300</td>
<td>E09 04 01 N36 45 53</td>
<td>3</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>TBeb</td>
<td>ICARDA</td>
<td>111262</td>
<td>Tunisia</td>
<td>Beja</td>
<td>150</td>
<td>E09 05 43 N36 34 50</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBlb</td>
<td>ICARDA</td>
<td>111634</td>
<td>Tunisia</td>
<td>Bizerte</td>
<td>30</td>
<td>E09 40 03 N37 12 33</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>TBlb</td>
<td>ICARDA</td>
<td>111713</td>
<td>Tunisia</td>
<td>Bizerte</td>
<td>5</td>
<td>E09 41 59 N37 06 35</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBlc</td>
<td>ICARDA</td>
<td>111900</td>
<td>Tunisia</td>
<td>Bizerte</td>
<td>10</td>
<td>E09 56 59 N37 08 48</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TJEa</td>
<td>ICARDA</td>
<td>111351</td>
<td>Tunisia</td>
<td>Jendouba</td>
<td>10</td>
<td>E08 50 45 N36 57 14</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>TKAa</td>
<td>ICARDA</td>
<td>109019</td>
<td>Tunisia</td>
<td>Kasserine</td>
<td>670</td>
<td>E09 02 N35 23</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>TSOb</td>
<td>ICARDA</td>
<td>109033</td>
<td>Tunisia</td>
<td>Sousse</td>
<td>100</td>
<td>E10 26 N35 51</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>TZAa</td>
<td>ICARDA</td>
<td>111065</td>
<td>Tunisia</td>
<td>Zaghouan</td>
<td>150</td>
<td>E10 13 29 N36 21 08</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>POR</td>
<td>USDA</td>
<td>P239858</td>
<td>Portugal</td>
<td>Alentejo</td>
<td>300</td>
<td>NA</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.6</td>
<td>5.5</td>
<td>30a</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>94</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NA = Not Available.

aCalculated excluding the Portuguese genotypes.
Data analysis

From the ISSR amplification profiles, only those bands that showed consistent amplification were scored (smeared and weak bands were excluded). The ISSR bands were scored as present (1) or absent (0) for each sample. A total of 32 bands (26 of which were polymorphic) from 250 to 1700 bp were scored. Electrophoretic ISSR patterns were translated into a binary data matrix of 32 bands × 94 individuals. The matrix was then analysed using different computer packages. As *T. tetraphyllum* is a predominantly selfing species (Couderc 1980), all of the individuals studied were considered as homozygous, and the data were therefore analysed assuming a haploid genome.

Genetic diversity, population analysis and differentiation

Genetic diversity analyses for the ISSR molecular marker data were carried out by estimating allele frequencies using the AFLP-SURV 1.0 computer program (Vekemans 2002); the frequency of the marker allele at each locus due to complete self-fertilisation. Estimates of allele frequencies were used, following the approach of Lynch & Milligan (1994), to calculate the effective number of alleles (*n*<sub>e</sub>), the unbiased expected heterozygosity, *H*<sub>eu</sub> (Nei 1987), genetic distances (Nei 1978) and *F*<sub>ST</sub> values (Wright 1969) between populations. To test the significance of differences in estimates of gene diversity, *n*<sub>e</sub>, and *H*<sub>eu</sub>, we used the non-parametric Wilcoxon signed-ranks test for two groups arranged as paired observations (Wilcoxon 1945). To determine the number of haplotypes and for analysis of molecular variance (*AMOVA*), we applied the population genetics package ARLEQUIN, version 3.1 (Excoffier et al. 2005). Using the TFPGA software (Miller 1997), we calculated the Nei’s unbiased heterozygosity, *H*<sub>eu</sub> (Nei 1987) for each locus and at different hierarchical levels, from state to collection site. TFPGA was also used to compute frequency distributions and to calculate genetic diversity for the phenotypic trait (epicotyl colour) and for the chloroplast microsatellites used. The Mantel test (Mantel 1967) of genetic and geographic distances was carried out to evaluate the correlation between the two data matrices, using the GENALEX, version 6 (Peakall & Smouse 2006).

Population structure

With the aim to further investigate population structure, we used the assignment method implemented in *structure*, version 2.1 (Pritchard et al. 2000). This infers the number of clusters (populations), *K*, present in a sample by comparing the posterior probability for different numbers of putative populations specified by the user, and it assigns individuals to these clusters, giving the percentage of membership (*q*). Twenty independent runs for each *K* (from 1 to 12) were performed using 30,000 MCMC (Markov chain Monte Carlo) repetitions and 30,000 burn-in periods, using no prior information and assuming correlated allele frequencies and admixture. The number of clusters (*K*) was estimated by computing the ad-hoc statistic, *ΔK*, based on the rate of change in the log probability of the data between successive *K* values (Evanno et al. 2005). *K* was set to 2, as this number maximised the *ΔK* parameter (Evanno et al. 2005). In this case, the run length was set to 100,000 burn-ins and 100,000 iterations. The output of the software gives, for each individual, the percentage membership for the *K* clusters. We then computed the average percentage membership (*q*) of the three populations and of each collection site. To display the membership coefficients on the North Africa map and to obtain an interpolated map, we used the r software (version 2.8.1, ‘maps’ package), using the Kriging method (R Development Core Team 2008).

Correlation between the membership coefficient *q* (obtained by *structure* analysis based on ISSR markers) and epicotyl colour classes or cpSSR markers was estimated using a non-parametric test (Spearman’s *ρ*, multivariate method) implemented in JMP 7.0 software (SAS Institute Inc. 2007, Cary, NC, USA).

Geographic structure

Spatial autocorrelation between spatial (geographical and altitude) and genetic distances were computed separately using Spatial Genetic Software (SGS), version 1.0d (Degen et al. 2001). Moran’s *I* (Moran 1950; Sokal & Oden 1978) was used for spatial distance classes (in metres), the dimensions of which were 136 km for geographic distances (10 classes) and 94 m for altitude differences (10 classes). The significances of the observed average Moran’s *I* values were assessed by comparing them with the corresponding values derived by randomly permuting (500 bootstrap) genotypes over the spatial coordinates of samplings. Ninety-nine per cent confidence envelopes were estimated. The mean overall pairs, corresponding to a random distribution, constituted a reference value against which spatial genetic structure was measured. Values higher than the 99% confidence limit indicate positive spatial genetic structure (*i.e.* proximal individuals are genetically more similar than expected if the distribution of genotypes was spatially random). Values below the lower confidence limit indicate negative spatial structure. Values within the 99% confidence limits do not deviate significantly from a random distribution.
Table 3. ISSR marker analysis for each of the North African states.

<table>
<thead>
<tr>
<th>state</th>
<th>N total</th>
<th>N polymorphys</th>
<th>% private</th>
<th>rare*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morocco</td>
<td>30</td>
<td>26</td>
<td>86.7</td>
<td>0</td>
</tr>
<tr>
<td>Algeria</td>
<td>29</td>
<td>25</td>
<td>86.2</td>
<td>0</td>
</tr>
<tr>
<td>Tunisia</td>
<td>27</td>
<td>22</td>
<td>81.5</td>
<td>0</td>
</tr>
<tr>
<td>Overall</td>
<td>32</td>
<td>26</td>
<td>81.3</td>
<td>NA</td>
</tr>
<tr>
<td>Mean</td>
<td>29</td>
<td>24</td>
<td>84.9</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA = not applicable.
*Rare marker, occurring in <5% of all individuals, frequency <0.05.

RESULTS

Genetic diversity, population analysis and differentiation

ISSRs

All of the five ISSR primers used gave reproducible banding profiles; the total number of bands from 250 to 1700 bp was 32, 26 of which (81.3%) were polymorphic, with an average of six total bands and five polymorphic bands per primer (Table 2). The percentages of polymorphic band (PPB) values ranged from 57.1% for ISSR-TE to 100.0% for ISSR-15 and ISSR-8082. For the three states analysed (Table 3), the average PPB was 84.9%, and varied little among the states, with the Moroccan accessions having the highest number of polymorphic bands (30; 86.7%). The Moroccan accessions also had a private marker (with a frequency of 0.40) and for the Tunisian population we identified three rare markers (occurring in no more than 5% of all individuals; frequency <0.05).

The diversity for ISSR markers was analysed in the three North African states for a total of 90 genotypes (Table 4). The total number of haplotypes detected was 73 (81.1%): 31 (86.1%) for the Moroccan population, 21 (75.0%) for Algeria and 21 (80.8%) for Tunisia. The highest genetic diversity (H_e) was found in Morocco (H_e = 0.33, standard deviation, SD = 0.18), and Tunisia (H_e = 0.20, SD = 0.19). The same pattern was observed for the average effective number of alleles (n_e), with the Moroccan population showing the highest value (n_e = 1.56, SD = 0.37), the Tunisian population a lower value (n_e = 1.31, SD = 0.33), and the Algerian population at the halfway point (n_e = 1.44, SD = 0.34). Using the non-parametric Wilcoxon signed-ranks test, the values of H_e and n_e for the Tunisian population were significantly higher than for the Moroccan population (P < 0.001). Indeed, the Moroccan population had values of H_e and n_e 1.65-fold and 1.19-fold higher, respectively, than the Tunisian population. The differences between the Algerian and Tunisian populations for both H_e and n_e were significant (P < 0.05, Wilcoxon signed-ranks test); the difference between the Moroccan and the Algerian populations was significant only for n_e (P < 0.05), but marginally not significant for H_e (P = 0.056).

In Table S2, for the ISSR markers, the number of haplotypes, ISSR marker distribution and genetic diversity were calculated for each North African state at different hierarchical levels (state, province and accession). The province with the highest genetic diversity identified was Tetouan, Morocco (H_e = 0.34), near the Straits of Gibraltar. In this case, 10 haplotypes out of 12 genotypes were found, and 85.7% of the ISSR markers were polymorphic. Miliana province showed the higher genetic diversity in an Algerian state (H_e = 0.32), and Beja province showed the highest genetic diversity within Tunisia (H_e = 0.20). A lower genetic diversity was detected for Constantine province (Algeria) with H_e = 0.01 and two haplotypes out of four genotypes analysed, even though only one collection site was available. Considering the accession level, the collection site showing the highest diversity was AMIb (in Miliana Province, Algeria, H_e = 0.27), followed by AALa (in Alger Province, Algeria, H_e = 0.20) and by MBMb (in Beni Mellal Province, Morocco, H_e = 0.18). Two accessions from Morocco (MBMd and MMAd), two from Tunisia (TBlb and TBBlb) and one from Algeria (AMEb) were completely monomorphic (H_e = 0.00, and single haplotype).

CpSSR and epicotyl colour

Two (ccmp2 and ccmp10) out of three cpSSRs used were monomorphic in all of the T. tetraphyllum genotypes analysed. Among the 94 genotypes analysed, amplification with ccmp7 gave four alleles, with a size ranging from 127 to 130 bp. Even with ccmp7, Morocco showed the highest genetic diversity (H_e = 0.58), with three alleles out of four (chlorotypes A, B and C); Algeria also had three out of four alleles (chlorotypes A, B and D), but with an overall H_e = 0.42 and the frequency of allele cp-129 (chlorotype B) equal to 0.70; Tunisia, instead, showed the presence of only two alleles (chlorotypes A and B) (Tables 4 and S2). Moreover, Morocco and Algeria each had a private chlorotype: C (with a frequency of 0.11) for the Moroccan accessions and D (with a frequency of 0.04) for the Algerian accessions.

With respect to the phenotypic marker evaluated (epicotyl colour), we found that the Moroccan accessions presented the three phenotypic classes with similar frequencies and an overall phenotypic diversity of 0.67, while for the Algerian accessions the class ‘absence of colour’ was the most frequent.
We calculated the proportions of membership \((q_1)\) for collection sites and 94 individual genotypes to inferred clusters 1 and 2. The second step of the analysis was assignment of the 36 collection sites, characterised by two groups, \(K = 2\), named cluster 1 and cluster 2. We calculated the proportions of membership \((q_1, q_2)\) for each collection site, where each vertical line represents an accession (collection site) and the percentage membership is the mean of individual \(q_s\). All of the Moroccan collection sites showed a high percentage membership for cluster 1, with the highest \(q_1\) for the collection sites MTEb and MTEe (0.96), and the lowest for collection sites MKHa (\(q_1 = 0.70\)). The Tunisian collection sites showed medium-high percentages of membership to cluster 2, from \(q_2 = 0.96\) for TKaa and TSOa, to \(q_2 = 0.66\) for TBea and TBB, with the exception of only one collection site, TBI, which showed a high percentage membership to cluster 1 \((q_1 = 0.85)\), and collection site TBe showing admixture \((q_1 = 0.44)\). The Algerian collection sites had very variable percentages of membership: three sites (AMMa, AMMB, and AMMb) were assigned to cluster 1 \((average q_1 = 0.92)\), three \((AOCa, ALa and ASea)\) to cluster 2 \((average q_2 = 0.89)\) and the remaining four collection sites had high levels of admixture with average \(q_1 = 0.58\) and \(q_2 = 0.42\).

At the genotype level, the same trend as for the collection sites was observed, with individuals belonging to the same collection site showing the same percentages of membership to the same cluster, with few exceptions. All of the Portuguese genotypes were assigned to cluster 1, with a very high percentage of membership \((q_1 = 0.95)\). If we set a threshold value for identification of the admixed genotype of 0.70 (the lowest in Moroccan collection sites), we observe that the Moroccan genotypes were 94% assigned to cluster 1 (with 6% admixed genotypes), the Tunisian genotypes were 77% assigned to cluster 2 and 12% to cluster 1 (with 11% admixed genotypes), while the Algerian genotypes showed 36% membership to cluster 1, 39% to cluster 2 and 25% admixed genotypes.

The correlation between epicotyl colour classes and STRUCTURE membership coefficient \((q)\) was significant (Spearman’s correlation \(r = -0.35, P = 0.0004\), pair-wise correlation \(r = 0.33, P = 0.001\)), while for the marker cpSSR7 the correlation was slightly significant (Spearman’s correlation \(r = 0.28, P = 0.01\), pair-wise correlation 0.23, \(P = 0.05\)).

**Mantel test**
The Mantel test was applied to the matrix of Nei’s unbiased genetic distances and the matrix of geographic distances, which showed a weak, but significant, correlation between the two matrices \((R = 0.22\) and \(P < 0.01)\). Moreover, we tested the correlation between molecular and altitudinal distances and observed an \(R = 0.08\) with \(P < 0.01\). These significant relationships between genetic structure and both the geographical and altitudinal variations were further investigated and validated using spatial autocorrelation analysis. While no significant trends were found for altitude levels, the spatial autocorrelation analysis confirmed the significant effect for geographic distances obtained by the Mantel test.

**Geographical structure**
Within the three distance classes (below about 400 km), positive Moran’s I values were detected, indicating that ‘proximal’ individuals (including those of the same collection site) were more genetically similar than expected from a random
distribution (see Fig. 2). The degrees of correlation, however, rapidly decreased, and negative $I$ values were found starting from the fourth class (above 450 km), suggesting a negative spatial structure in this range of distances (from 450 to 800 km). From the sixth class (beyond 800 km), the $I$ values measured tended to not fall out of the 99% confidence envelope, indicating that above that distance $I$ values did not deviate significantly from a random distribution.

**DISCUSSION**

This is the first study that has analysed molecular diversity and population structure within *T. tetraphyllum*. We analysed North African populations obtained from the ICARDA (International Centre for Research in the Dry Areas, Aleppo – Syria) germplasm collection centre. Our study has here provided an optimised method for evaluating genetic diversity of
Genetic diversity and geographic differentiation in *T. tetraphyllum*

*T. tetraphyllum* using ISSR markers, which is an economic and rapid approach used on an alternative and neglected legume species, and can be useful for further investigation. Unlike co-dominant markers, estimation of allele frequencies from dominant markers, such as ISSR, presents some statistical difficulties (Lynch & Milligan 1994). These difficulties should be taken into account and can be resolved using appropriate methods, such as Bayesian methods (Zhivotovsky 1999), and better estimators for analysis of dominant markers (Lynch & Milligan 1994).

Ninety genotypes of *T. tetraphyllum* from three North African states were analysed using five ISSR primers, yielding 32 scorable bands; this number is lower than in similar studies based on ISSRs in other plant species, but given the somewhat preliminary flavour of this study, it proved large enough to provide some useful insights. The three populations show percentages of polymorphism ranging from 81.5% (Tunisia) to 86.7% (Morocco), while the overall polymorphism was 81.3%. Other studies in plant species with a large distribution, based on a greater number of individuals and ISSR bands, yielded similar ISSRs values of polymorphism (in Penstemon, Wolfe et al. 1998 and in *Primula obconica*, Nan et al. 2003). The ISSR polymorphism detected revealed a high level of variability, suggesting that ISSR technology is a powerful and efficient approach to assess population structure in *T. tetraphyllum*.

The populations of *T. tetraphyllum* showed clear geographical structures, with the differentiation associated mainly to longitudinal differences, which extended from Marrakech (collection site MMAa, 07°49′ W) in Morocco to Sousse (collection site TSOa, 10°26′ E) in Tunisia. We were thus able to observe a cline in distribution of this species from west to east in the Mediterranean region considered, and the pattern was highlighted using R software, with the plotted structure membership coefficients (see Fig. 1a and b).

Both the spatial autocorrelation and *amova* results indicated a high population structure for nuclear and chloroplast markers. This is in agreement with a largely autogamous breeding system, where a low frequency of out-crossing reduces migration between populations by reducing the pollen flow (Hamrick & Godt 1990). Moreover, the Mantel test and spatial autocorrelation analysis suggested that population structure was coherent with the Isolation-by-Distance model (IBD) of migration proposed by Wright (1946), with a positive and significant Moran I value between plants at distances <400 km. This indicates again that geographic distance was an important factor for genetic divergence, and genetic drift might be the primary evolutionary force during the process. Any genetic differentiation, however, was not simply a function of spatial distance, as indicated by the moderate correlation between genetic and geographic distance matrices. Other factors known to be important include the life-history characteristics, population arrangements (patchy versus continuous), physical barriers to dispersal and habitat factors affecting geographical distribution, such as soil, temperature and salinity. The high levels of variability observed within countries may be required to maintain plasticity in a highly patchy and diverse environment like the Mediterranean Basin (Ortiz-Dorda et al. 2005).

Based on the markers used, the most important outcome of this study is that the results strongly suggest that *T. tetraphyllum* extended its area of distribution from Morocco to Tunisia. Indeed, for all the markers analysed (ISSR, cpSSR, and colour of the epicotyl; Table 4), a general reduction in genetic diversity from Morocco to Tunisia can be seen. Furthermore, in Morocco we can identify a private ISSR fragment and a private cpSSR allele, both with a frequency higher than 10%. These alleles could be lost during the migration of populations of *T. tetraphyllum* from Morocco to Algeria and Tunisia. Similarly in Tunisia, we can find three rare alleles that, on the contrary, showed a frequency higher than 5% in Morocco and Algeria.

Based on results of the genetic diversity gradient observed and of the cline of distribution highlighted by structure analysis, we can state that *T. tetraphyllum* extended its area of distribution from Morocco to Tunisia, and we can also suggest a possible hypothesis in relation to the origin of *T. tetraphyllum*. This hypothesis is also supported by the distribution of two species still classified in the genus Anthyllis, but very close to *T. tetraphyllum*, and belonging to the ‘tetraphylla clade’ (Nanni et al. 2004): *A. hamosa*, the distribution of which is limited to the Iberian peninsula, and *A. cornicina*, which is distributed both in the Iberian Peninsula and in Morocco (Greuter et al. 1989). The hypothesis is that *T. tetraphyllum* arose in Morocco, or that it originated from the Iberian Peninsula. To resolve the question of the origin of *T. tetraphyllum*, a more detail analysis should be conducted, including populations from other Mediterranean countries and in particular from the Iberian Peninsula. It is also interesting that in the province of Tetouan, the nearest to the Gibraltar Strait, we identified the highest value of He (0.34), and we can find almost all of the ISSR markers (28 out of 32), three out of four cp-SSR alleles, all the classes of epicotyl colour and the Morocco private allele, with a frequency of 0.43. Moreover, the ITS sequences and cpSSR analysis (Nanni et al. 2004) showed that the closest species to *T. tetraphyllum* is *A. hamosa*, as previously mentioned, a species endemic in the Iberian Peninsula.

The Mediterranean region is experiencing a remarkable effect of climate change and human activities, and it is thus important that there is continuous monitoring of species and population dynamics, to plan and enact valuable conservation programmes for species such as *T. tetraphyllum*. This study has implications for the use and conservation of this alternative pasture legume through its potential role for multiple uses in the changing environments of the Mediterranean basin. Analysis of distribution of genetic variability suggests that the higher percentage of total variation is still harboured within states and provinces and among collection sites, indicating that in the North African environment, this species is not yet endangered.

**ACKNOWLEDGEMENT**

The authors would like to thank the Genetic Resources Unit of the International Centre for Research in Dry Areas (ICARDA, Aleppo, Syria) for supplying the seeds.

**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article.
Table S1. Primer cpSSRs used for the analysis.
Table S2. Diversity analysis of *T. tetraphyllum* genotypes.

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

REFERENCES
Genetic diversity and geographic differentiation in *T. tetraphyllum*  

Bellucci, Nanni, Bitocchi, Rossi & Papa

Proceedings of the National Academy of Sciences USA, 92, 7759–7763.
Vekemans X. (2002) AFLP-SURV, version 1.0. Distributed by the author. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.