Genome size and phylogenetic relationships between the Tunisian species of the genus *Calligonum* (Polygonaceae)

Hassen GOUJA1,2,3, Alfredo GARCÍA-FERNÁNDEZ3*, Teresa GARNATJE3, Aly RAIES2, Mohamed NEFFATI1

1Pastoral Ecology Laboratory, Arid Regions Institute, Medenine, Tunisia
2Microbial and Biomolecular Laboratory, Faculty of Sciences of Tunis, University of Tunis El-Manar, El-Manar, Tunisia
3Botanical Institute of Barcelona (IBB-CSIC-ICUB), Barcelona, Spain

Abstract: The species of the genus *Calligonum* L. (Polygonaceae) that appear in Tunisian deserts play an important role in maintaining local ecosystems and supply important natural resources in these regions. In this study, 31 individuals belonging to the 3 hypothetical species of the genus *Calligonum* (*Calligonum arich* Le Houér., *Calligonum azel* Maire, and *Calligonum comosum* L’Hér.) were collected from 12 populations in 6 localities of the Tunisian desert and examined to assess the relationship between species. Phylogenetic analyses using 1 nuclear (ITS) and 2 plastid regions (*trnL-trnF* and *rbcL*) and genome size assessments are used in this study to evaluate the relationships between these hypothetical species. C-value results suggest the existence of 3 different species, which is also supported by phylogenetic tree topology. The paraphyletic pattern of *C. comosum* and the plausible origin of *C. arich* from *C. azel* suggest the existence of several mechanisms of isolation and speciation in the Sahara Desert for this genus. Additional studies are necessary to evaluate the population size, demographic tendency, and conservation status of these desert species and their genetic relationships with other congener.

Key words: *Calligonum*, desert plants, genome size, ITS, *rbcL*, *trnL-trnF*, Polygonaceae

1. Introduction

*Calligonum* L. is widely distributed in North Africa, South Europe, and West and Central Asia with approximately 60–80 species (Brandbyge, 1993), with rapid evolution and diversification spots in Asia’s deserts (Mabberly, 1990). This accelerated differentiation process has important physiological and morphological consequences, which include the C4 photosynthesis mechanism, high number of stems, and hybridisation and introgression processes (Pyankov et al., 2000). In the Saharan areas, *Calligonum* species are dominant perennial shrubs and trees in active sand dunes and in stabilised sand fields in North Africa (Le Houérou, 1959), and they can tolerate extreme drought conditions by the loss of leaves and branches during the driest months (Le Houérou, 1959; Dhief et al., 2009). They appear to be suitable for rehabilitation of degraded lands as well as use in dry land and agroforestry systems, due to their high tolerance to xerophytic conditions (Dhief et al., 2009) and resistance to summer drought (Le Houérou, 1959; Dhief et al., 2009). Furthermore, Liu et al. (2001) and Badria et al. (2007) highlighted the importance of the genus as a source of natural resources, drugs, and antidotes.

In Tunisia, 3 species of the genus have been described: *Calligonum arich* Le Houér., *C. azel* Maire, and *C. comosum* L’Hér. (Le Houérou, 1959; Pottier-Alapetite, 1979). These species occur in different dune slope positions in the eastern Grand Erg of the Tunisian desert (Le Houérou, 1959). *Calligonum comosum* and *C. azel* are leafless shrubs or small trees with white flowers and young succulent branches serving as assimilation organs reaching a height of >1.5 to <2.2 m (interdune slopes) and >3.5 to <5.5 m (dune slopes), respectively (Dhief et al., 2012). The young branches of *C. comosum* stems are fragile, articulated, fasciculate, and green, whereas the older ones are white, although certain plasticity has been described for some morphological traits in *C. comosum*, including stem colour (Taia and Moussa, 2011). Contrarily, the stems of *C. azel* are mostly buried in the sand and old branches carry a whitish bark peeling off in pieces. Moreover, *C. arich* is a tree, up to 8 m tall with red flowers that grow at dune crests, and the older branches carry a brownish-yellow bark peeling off in pieces (Le Houérou, 1959, Dhief et al., 2012). Recent decrease in the population size and changes in the spatial distribution of the 3 studied species can be
explained by the intensive use of these plants for fuel and charcoal production by local nomadic populations (Le Houérou, 1959; Pottier-Alapetite, 1979).

While 2 of these species are widely accepted, Calligonum azel and C. comosum, the existence of C. arich is doubtful. On the one hand, Le Houérou (1959) and Pottier-Alapetite (1979) considered this species as a Tunisian endemism of the Grand Erg Oriental, together with the databases (International Plant Names Index, 2012; Tropicos, 2012). On the other hand, Greuter et al. (1989) proposed its elimination and, in the current version of the catalogue of flora from Tunisia (Le Floc’h et al., 2010), its existence is doubtful due to the insufficient information available about this species. Previous studies based on morphological and anatomical characters (presence or absence of an ochrea, the tepal arrangement, and anatomy of stomata) have failed to clarify the relationships among these 3 species (Dhief et al., 2009; Le Floc’h et al., 2010) even though cluster analyses carried out with RAPD markers found some genetic differences among the 3 species (Dhief et al., 2011a). These unresolved results lead into the main objective of this study, to evaluate the existence of C. arich among the species of the genus Calligonum in Tunisia.

For this goal, we have considered 2 different approaches: describing the phylogenetic relationships among the 3 species of the genus Calligonum; and estimating the differences in the genome size, if any, among the studied species.

2. Materials and methods
2.1. Plant materials
In this study, we revealed the phylogeny of 31 individuals of 3 Calligonum species (C. arich, C. azel, and C. comosum). Young green branches for each species were collected from 6 sites of the southern desert of Tunisia (Table 1; Figure 1). The samples were collected from adult individuals, selecting green healthy branches (without parasitism or signs of drought stress), and were cleaned in the laboratory and kept in a freezer at –80 °C. In addition, 5 individuals from each species were collected separately from the Kamour site and preserved at 4 °C in order to measure the genome size. Specimen vouchers of the studied material were deposited in the herbarium of the University of Barcelona (BCN). The classification system used was the catalogue of the Tunisian flora (Le Floc’h et al., 2010).

<table>
<thead>
<tr>
<th>Population</th>
<th>Species</th>
<th>Number of individuals analysed</th>
<th>Coordinates</th>
<th>GenBank accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>ITS</td>
<td></td>
</tr>
<tr>
<td>Kamour</td>
<td>C. comosum</td>
<td>2</td>
<td>32°33′60″N 09°27′57″E</td>
<td>KC585426-27</td>
</tr>
<tr>
<td></td>
<td>C. azel</td>
<td>3</td>
<td></td>
<td>KC585435-37</td>
</tr>
<tr>
<td></td>
<td>C. arich</td>
<td>5</td>
<td></td>
<td>KC585444-48</td>
</tr>
<tr>
<td>Douz</td>
<td>C. comosum</td>
<td>1</td>
<td>33°14′20″N 09°20′13″E</td>
<td>KC585419</td>
</tr>
<tr>
<td>El Ouera</td>
<td>C. comosum</td>
<td>3</td>
<td>32°40′18″N 10°35′41″E</td>
<td>KC585420-22</td>
</tr>
<tr>
<td>J’bil</td>
<td>C. comosum</td>
<td>2</td>
<td>32°59′20″N 09°00′29″E</td>
<td>KC585417-18</td>
</tr>
<tr>
<td></td>
<td>C. arich</td>
<td>2</td>
<td></td>
<td>KC585441-43</td>
</tr>
<tr>
<td>El Borma</td>
<td>C. comosum</td>
<td>3</td>
<td>31°38′40″N 09°27′44″E</td>
<td>KC585428-30</td>
</tr>
<tr>
<td></td>
<td>C. azel</td>
<td>2</td>
<td></td>
<td>KC585433-34</td>
</tr>
<tr>
<td></td>
<td>C. arich</td>
<td>3</td>
<td></td>
<td>KC585438-40</td>
</tr>
<tr>
<td>Tiert</td>
<td>C. comosum</td>
<td>3</td>
<td>30°47′06″N 10°16′54″E</td>
<td>KC585423-25</td>
</tr>
<tr>
<td></td>
<td>C. azel</td>
<td>2</td>
<td></td>
<td>KC585431-32</td>
</tr>
</tbody>
</table>

Table 1. Populations of Calligonum species considered for this study, number of individuals sampled for each population, spatial coordinates, and GenBank accession numbers for the sequences employed.
2.2. DNA extraction procedure, PCR protocol, amplification, and sequencing

Genomic DNA was extracted following the CTAB method described by Doyle and Doyle (1990) with some modifications. DNA extractions were checked by electrophoresis using 0.8% agarose gel and DNA concentration was estimated at an optical density of 260 nm using a NanoDrop 1000 spectrophotometer (Thermo Scientific Inc., Wilmington, DE, USA).

A preliminary screening was carried out to find optimal and variable nuclear (ITS) and chloroplast (trnF, rbcL, ndhF, psbA-trnH, matK) regions, based on previous studies in the genus *Calligonum* and family Polygonaceae (e.g., Sanchez et al., 2009; Tavakkoli et al., 2010; Tian et al., 2011; Sun and Zhang, 2012). DNA amplification reactions of *nr* DNA ITS (ITS1, 5.8S, and ITS2) and 2 plastid fragments (trnL-F and rbcL regions) were performed in a 25-µL reaction mixture, according to Tavakkoli et al. (2010), containing 2.5 µL of a 50-ng genomic DNA template, 12.8 µL of sterile water, 2.5 µL of 2 mM MgCl₂, 2.5 µL of 10X Gene Taq universal buffer, 2.5 µL of 2.5 mM dNTP mixture, 1 µL of each primer (5 pmol/µL), and 0.2 µL of AmpliTaq DNA polymerase at 5 U/µL (Applied Biosystems, Carlsbad, CA, USA). The *nr* DNA (ITS1, 5.8S, and ITS2) was amplified using the universal ITS1-ITS4 pair of primers (White et al., 1990). The polymerase chain reaction (PCR) conditions, as performed in a thermocycler (S 1000 Thermal Cycler, Bio-Rad Laboratories, Inc., Hercules, CA, USA), were 2 min at 94 °C; 40 cycles consisting of denaturation for 1 min at 94 °C, annealing for 40 s at 53 °C, and elongation for 1 min at 72 °C; and a final extension for 7 min at 72 °C. The plastid trnL-trnF region was amplified using the universal “c” and “d” primers of Taberlet et al. (1991) under the following conditions: 2.5 min at 95 °C for initial denaturation, followed by 38 cycles of 1 min at 95 °C, 45 s at 53 °C for annealing, 2 min at 72 °C for extension, and then a final 7-min incubation at 72 °C. The rbcL gene was amplified by PCR using primers rbcL-1F and rbcL-724r (3 min at 94 °C, follow by 38 cycles of 1 min of denaturation at 94 °C, 50 s of annealing at 50 °C, 1 min of extension at 72 °C, and then 8 min for final extension at 72 °C, as per Chase et al. (1993)). All amplified products (plastid and nuclear) were checked by electrophoresis on a 1.2% agarose gel (w/v). Each region was sequenced with the BigDye Terminator Cycle Reaction Kit v. 3.1 (PE Biosystems, Foster City, CA, USA) using the same primers as described for PCR. Nucleotide sequencing was carried out at the Centres Científics i Tecnològics of the University of Barcelona on an ABI PRISM 3730XL DNA analyser (PE Biosystems).

2.3. Sequence alignment and phylogenetic analyses

The trnL-trnF, rbcL, and ITS sequences were edited and aligned using BioEdit v. 7.1.3.0 (Hall, 1999) followed by manual adjustment. Alignment of each dataset required the introduction of numerous single and multiple-base indels (insertions/deletions). Positions of indels were treated as missing data for all datasets (Tables 1 and 2). For both nuclear and chloroplastic matrices, *Calligonum junceum* and *Calligonum arborescens* were considered

Figure 1. Photographs of the study area depicting: a- *Calligonum comosum* in interdune, b- *C. azel* in dune slope, c, d- *C. arich* in dune crest.
as outgroups (GenBank accession numbers: C. junceum trnL-trnF JQ009289, rbcL GQ206217, ITS AB542774; C. arborescens trnL-trnF JQ009294, rbcL EU840536, ITS JN 187105; Sanchez et al., 2009; Tavakkoli et al., 2010; Tian et al., 2011; Sun and Zhang, 2012).

Combination of ITS and plastid datasets was assessed using the incongruence length difference (ILD) test (Farris et al., 1994) implemented in PAUP (Swofford, 2003). The test was performed using the heuristic search option involving simple addition sequence and tree bisection-reconnection (TBR) branch swapping with 1000 replicates and excluding invariant characters.

Maximum parsimony analyses were carried out using PAUP. The heuristic search option was employed, using TBR branch swapping with 1000 replications of random addition sequence and excluding uninformative characters. Full heuristic search was used for branch support with 1000 bootstrap replicates each with a simple addition sequence.

Complementary, independent Bayesian phylogenetic analysis was performed in MrBayes 3.1.2 (Huelsenbeck et al., 2001) using a GTR+G model determined from jModeltest v. 0.1.1 (Posada, 2008) under the Akaike information criterion (AIC; Akaike, 1979), to ascertain phylogenetic relationships. The Markov chain Monte Carlo (MCMC) sampling approach was used to calculate posterior probabilities (PPs). Four consecutive MCMC computations were run for 1,000,000 generations, with tree sampling every 100 generations. Data from the first 1000 generations were discarded as the burn-in period. PPs were estimated through the construction of a 50% majority rule consensus tree. The output trees were edited with FigTree v. 1.3.1 (Rambaut, 2006).

Finally, a neighbour-net analysis was carried out considering the uncorrected p-distance between individuals and the same outgroup species from the Bayesian analyses, using SplitsTree 4.10 (Huson and Bryant, 2006). Branches’ support was tested using bootstrapping with 1000 replicates.

2.4. Flow cytometry measurements and analysis

Five individuals were analysed by flow cytometry of each species of Calligonum from the locality of Kamour. Fresh young leaves were used in this analysis and each sample was measured 2 times. The standard method used for DNA content measurement by flow cytometry was as described by Garnatje et al. (2004). Petunia hybrida Vilm. ‘PxPc6’ (2C = 2.85 pg; Marie and Brown, 1993) and Lycopersicon esculentum Mill. ‘Montfavet 63-5’ (2C = 1.99 pg; Catrice et al., 2006) were used as the internal standards and were firstly analysed separately to locate their peak positions. Leaf tissue of each Calligonum sample (about 50 mm²) and the internal standard (approximately half the quantity) were chopped with a razor blade in a plastic petri dish in 600 µL of LB01 buffer (8% Triton X-100) (Dolezel et al., 1989), supplemented with 100 µg/mL ribonuclease A (RNase A, Boehringer, Meylan, France). To eliminate cell debris, nuclei were filtered 2 times through a 30-µm nylon filter; 60 µg/mL propidium iodide was subsequently added (Invitrogen Corporation, Carlsbad, CA, USA). Samples were kept on ice for 10 min before measurement. Fluorescence analysis was carried out using an Epics XL flow cytometer (Coulter Corporation, Hialeah, FL, USA) at the Centres Científics i Tecnològics of the University of Barcelona.

One-way ANOVA analysis was carried out to test the differences in the genome size values, if any, among the 3 studied species. This analysis was performed with Stata 10.0 (Stata Corporation, College Station, TX, USA).

3. Results

3.1. Phylogenetic analyses

The ILD test did not show significant incongruence between plastid and ITS regions (P = 0.98), which allowed the combination of both datasets. The aligned matrix with 31 accessions containing the 3 considered regions is 1749 bp long (620 bp of ITS, 522 bp of trnL-trnF, and 607 bp of rbcL), of which 122 are potentially informative, being the ITS regions most useful to the phylogenetic reconstruction. Table 2 details the number of indels and polymorphic and informative sites for each region in the 3 species.

The tree obtained with Bayesian inference is shown in Figure 2. Calligonum arich shows statistical support for a monophyletic independent branch (PP = 100%); meanwhile, C. azel shares a common origin with C.

<table>
<thead>
<tr>
<th>Table 2. Summary of the number of polymorphic sites, potential informative sites, and indels for each region (ITS, trnL, and rbcL).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regions</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>ITS</td>
</tr>
<tr>
<td>trnF</td>
</tr>
<tr>
<td>rbcL</td>
</tr>
</tbody>
</table>
**arich** (but lacks statistical support). *Calligonum comosum* shows a paraphyletic pattern with no significant value in the branches (Figure 1). Similar results to those shown by Bayesian analysis in the strict consensus tree were obtained with the maximum parsimony approach and bootstrapping coefficients (data not shown). Figure 3 shows the obtained neighbour-net. All populations from the same species are clustered in the same group with significant statistical support.

### 3.2. Genome size

The 2C values obtained for the tree of *Calligonum* species ranged from 2.30 ± 0.10 pg for *C. azel* to 2.45 ± 0.07 pg for *C. arich* and 4.68 ± 0.17 pg in *C. comosum*. Half-peak coefficient of variation (HPCV) values were lower than 5% in all cases, which indicates the high quality of the analyses. Slight significant differences (F = 4.03, P = 0.04) were obtained comparing the genome size values (1Cx), assuming *C. arich* to be diploid.

### 4. Discussion

#### 4.1. Phylogenetic analyses of *Calligonum* from Tunisia

Previous studies have suggested the existence of the 3 species of the genus *Calligonum* in Tunisia, even though some differences have also been found between the different analyses carried out. On the one hand, the cluster analysis developed by Dhief et al. (2012) based on morphological traits showed *C. comosum* and *C. azel* grouped together (similarity of 29%), sharing some phenotypic traits (e.g., the stem, flower and fruit colour), whereas *C. arich* remained isolated. On the other hand, using some chemical descriptors (essential oils), *C. arich* and *C. comosum* were grouped with a similarity of 38% without *C. azel* (Dhief et al., 2011b). Both analyses of morphological diversity and of chemical polymorphism showed different grouping patterns. However, results from phylogenetic analyses support an alternative scenario. Nonmonophyletic clades were obtained for *C. comosum* in the phylogenetic reconstructions (Figure 2). Our work, together with previous studies carried out with polymorphic markers (RAPD; Dhief et al., 2011a), has shown higher values of genetic diversity in this species than in the other 2, which might support partially the tree topology and the network pattern obtained (Figures 2 and 3) in this study. The wider distribution of *C. comosum* along the Tunisian desert and other regions (Koller, 1956), its interdune position, the phenotypic plasticity described (Taia and Moussa, 2001), and its high production of flowers and fruits (Dhief et al., 2009, 2012) may have facilitated sexual reproduction, the gene flow between populations, or potential genomic processes such as population isolation, introgression,

---

**Figure 2.** Majority rule consensus tree from Bayesian inference based on sequences from the merged dataset of the 3 regions (ITS, *trnL*-F, and *rbcL*). The numbers above the branches are the PPs.
or hybridisation at the contact zones with other taxa. Hybridisation processes between current species (or common ancestors) might also explain the polyphyletic pattern found in *C. comosum*, but there is no evidence in DNA sequences. *Calligonum comosum* exhibits a nuclear DNA amount higher than those of the other species and the genome size of this taxon is nearly double the 2C value of *C. azel* and *C. arich*. Considering that the nuclear DNA amount is not purely duplicated during chromosome doubling (Levy and Feldman, 2002), we might hypothesise that *C. azel* or *C. arich* could be involved in the origin of the tetraploid *C. comosum*. Despite the lack of support from the phylogenetic tree (Figure 2), this species is well defined by the previous karyological (Ferchichi, 1997) and cytogenetic characteristics or the neighbour-net approach (Figure 3). Furthermore, its tetraploid karyotype would be prone to accumulation of mutations (Otto, 2007), which facilitate the divergence and the genetic differences between populations (Figures 2 and 3), which may occur more frequent in nuclear than in chloroplast regions (Table 2).

All populations belonging to *C. arich* merged into a well-supported clade (PP = 100%, BS = 99%; see Figures 2 and 3, respectively) but the relationships with *C. azel* are not clear. In this context, a pair of hypotheses are possible: the first is that *C. azel* could have originated from *C. arich* through a polyploidisation process. The genome size values (doubled in *C. azel* compared to *C. arich*) could also support this evidence. The second hypothesis bears the inclusion of *C. arich* into *C. azel*, making it monophyletic, but neither the previous morphological traits and chemical studies (Dhief et al., 2009, 2012) nor the genome sizes agree with this possibility. Since the taxonomical status of *C. arich* could have important implications in the policy of conservation, further karyological, cytogenetic, and phylogenetic analyses with additional genetic markers would be necessary to establish the chromosome number or cytogenetically characterise and chronologically date *C. arich*.

4.2. *Calligonum* genome size

The present study showed the first measures of C values for the genus *Calligonum* (see RGB Kew Plant DNA C-values, Bennett and Leitch, 2010). The small coefficients for HPCV obtained support the reliability of the measurements done, although intrapopulation variability in the genome size and its correlation to certain factors have already been described (e.g., García-Fernández et al., 2012; Tabur et al., 2012) and could not be discarded in certain *Calligonum* populations. Measures with flow cytometry might have some technical and logistical problems for these
species (maybe associated with the presence of alkaloids, polysaccharides, and secondary metabolites (Liu et al., 2001)), due to the difficult of obtaining fresh material from desert populations and their preservation for measurement (Doležel et al., 2007).

The significant statistically differences obtained suggest the existence of 3 independent species of the genus *Calligonum* growing in Tunisia. Mao et al. (1983) proposed \( x = 9 \) as the basic chromosome number for the genus. Furthermore, Ferchichi (1997) carried out chromosomal counting in *C. comosum* (2n = 36, tetraploid) and *C. azel* (2n = 18, diploid), which revealed the different ploidy levels. Other cytogenetic processes that modify the chromosome number, such as mixoploidy or aneuploidy, have been described in some *Calligonum* species and are frequent in other genera of the tribe Atraphaxideae (Xinmin et al., 2009). All these processes might have played relevant roles in the evolution of the genome size, highlighting their key role during genome duplication and polyploidization throughout plant evolution (Soltis and Soltis, 1999). During these processes, genome size must be doubled together with the chromosome number, although a certain quantity of DNA could be eliminated (Leitch and Bennett, 2004) to avoid potential mutations that are frequent in polyploids (Sharma and Sen, 2002).

Polyploids are very common in angiosperm species (Soltis and Soltis, 1999) and might favour the colonisation of new and harsh environments (St Clair and Howe, 2011; te Beest et al., 2012). In spite of the relevant physiological cost of the increase in genome size (i.e. Gregory, 2001; Otto, 2007), the rise in the number of chromosomes supposes potential new alleles and an increase of the genetic variation may play its role in selection under new conditions and stress factors (Otto, 2007). The tetraploid chromosome number of *C. comosum* might have been involved in its ability to colonise different areas of the Tunisian desert. Similar responses to the high levels of stress and chromosome evolution have been found in other genera of the tribe Polygonaceae, such as *Fagopyrum* Mill. (Yamane and Ohnoshi, 2003) and *Rumex* L. (Stehlik and Blattner, 2004)

### 4.3. Concluding remarks

Some cytogenetic and phylogenetic evidence for the independence of the 3 *Calligonum* species in the Tunisian desert are reported in this study. These results are very helpful to preserve, promote, and find the appropriate strategy for the use of these species in future programs of rehabilitation of degraded Tunisian areas. Further studies are needed to establish the relationships among these 3 species and the related congeneric ones.

### Acknowledgements

This work was subsidised by MICINN (Spanish Government; Project CGL2010-22234-C02-01 and 02). We thank the technical staff of the Range Ecology Laboratory, Arid Lands Institute of Medenine, who helped us to develop the fieldwork, and the staff of the Laboratory of Biosystematics and Molecular Systematics at the Botanical Institute of Barcelona for their support and help. We also want to express our gratitude to Jaume Comas and Ricard Álvarez for technical support in flow cytometry. We also thank Alastair Plant for language assistance. AG-F benefited from a postdoctoral contract (Project CGL2010-22234-C02-01). We also acknowledge the commentaries of the 3 anonymous reviewers on previous versions of the manuscript.

### References


