A phylogenetic analysis of genus *Onobrychis* and its relationships within the tribe *Hedysareae* (Fabaceae)

Nadeesha LEWKE BANDARA¹*, Alessio PAPINI¹***, Stefano MOSTI², Terence BROWN³, Lydia Mary Josephine SMITH¹
¹The John Bingham Laboratory, National Institute of Agricultural Botany, Cambridge, UK
²Department of Evolutionary Biology, University of Florence, Florence, Italy
³Faculty of Life Sciences, Manchester Interdisciplinary Biocentre, University of Manchester, Manchester, UK

Abstract: Results about a phylogenetic analysis of the genus *Onobrychis* Mill., tribe *Hedysareae* DC. are presented. The systematic knowledge of tribe *Hedysareae* is still incomplete, with difficult circumscription of genera and species. Analyses were undertaken using both nuclear (ITS) and chloroplast (matK) markers for a set of 78 accessions covering 41 *Onobrychis* species, besides previously sequenced *Hedysareae* accessions. The phylogenetic methods used were maximum parsimony, maximum likelihood, and Bayesian analyses to produce phylogenetic trees and robustness indices. The genus *Onobrychis* was resolved as paraphyletic, with species of the genera *Eversmannia* Bunge and *Hedysarum* L. nested within it. The position of the section *Membranacea* of genus *Hedysarum* was as a sister group to *Onobrychis* and *Eversmannia*, separated from other accessions of *Hedysarum*. Variation in the 2 markers was sufficient to resolve infrageneric groups in *Onobrychis* and *Hedysarum*, but we were unable to completely resolve certain species in *Onobrychis*, particularly those within the sect. *Onobrychis*. The cause of this difficult species delimitation may be related to recent speciation, hybridization, and introgression events, particularly between cultivated species and their wild relatives, and the presence of cryptospecies as suggested by intraspecific polyploid series.

Key words: ITS sequences, Leguminosae, matK sequences, molecular phylogenetic
Hedysarum and Onobrychis were separated taxonomically on the basis of fruit morphology, in addition to pollen structure, chromosome number, and biochemical features (Polhill, 1981; Yildiz et al., 1999).

Different approaches that have been used to define the taxonomy of Onobrychis in terms of species and infrageneric taxa circumscription have led to contradictions and uncertainty. This is probably due to the fact that only a limited number of characters have been considered in any one of the available taxonomic descriptions (Boissier, 1872; Ball, 1968; Hedge, 1970; Rechinger, 1984). We used the sectional treatment of Schischkin and Bobrov (1972), with updates by Yildiz et al. (1999) and Ahangarian et al. (2007).

The most frequently used characters are: annual or perennial habit, number of ovules, adnate or free stipules, size, the proportion or character of the indumentum, fruit morphology, and seed number. Yildiz et al. (1999), for example, outlined a classification based mainly on fruit morphology using a sample of 40 species for 5 sections of a total of 162 species classified into 2 subgenera and 8 sections. In addition to the other morphological data, Dolya and Vasilissa (2000) and Avci et al. (2013) used pollen morphology, while Irfan et al. (2007) used electrophoretic analysis of total seed proteins to study the systematics of Onobrychis. Unfortunately, the number of species included in these last 2 studies was too low to draw clear general conclusions on the genus.

A detailed taxonomic investigation of the genus Onobrychis based on molecular markers is still lacking. More recently, a molecular investigation using rDNA internal transcribed spacers (ITS) molecular data (Ahangarian et al., 2007) considered the tribe Hedysareae. The sample set included 11 species of Onobrychis.

The ITS sequences have been shown to elucidate phylogenetic relationships, especially at the species and genus levels (Baldwin et al., 1995; Gültepe et al., 2010; İkinci et al., 2011). Important results in Leguminosae have led to contradictions and uncertainty. This is probably due to the fact that only a limited number of characters have been considered in any one of the available taxonomic descriptions (Boissier, 1872; Ball, 1968; Hedge, 1970; Rechinger, 1984). We used the sectional treatment of Schischkin and Bobrov (1972), with updates by Yildiz et al. (1999) and Ahangarian et al. (2007).

2. Materials and methods

2.1. Sampling material and total DNA extraction

Seed samples were collected from different locations, including the Mediterranean area, North America, Iran, and other areas of Asia. The seeds were stored at the National Institute of Agricultural Botany (NIAB) Gene Bank (Cambridge, UK). Additional samples were obtained from leaves of dried specimens of the Bu-Ali Sina University Herbarium, Iran (for all specimens used in the analysis, see Table 1S in the supplementary material at http://www.unifi.it/caryologia/tjb).

Genomic DNA was isolated either from approximately 40 mg of fresh leaves or from herbarium sample leaves, using the modified Tanksley method (Fulton et al., 1995). Plant tissues were stored at –80 °C until DNA extraction. The microprep buffer was prepared by mixing DNA extraction buffer (0.35 M sorbitol, 0.1 M Tris, 5 mM EDTA), nuclei lysis buffer (0.2 M Tris, 0.5 M EDTA, 2 M NaCl, 2% CTAB), 5% sarcosyl, sodium bisulfite, and RNase. This microprep buffer was incubated at 65 °C. Frozen dried leaf samples were milled using the QIAGEN Geno/Grinder with 500 µL of microprep buffer. Milled samples were incubated at 65 °C for 30 min and then DNA purification continued using chloroform:isoamylalcohol, isopropanol, and 70% ethanol steps. DNA concentrations were estimated by gel electrophoresis on 1% agarose. We used 1 DNA sample of more than 10 ng/µL for each accession.

2.2. Amplification of ITS and matK region

DNA fragments were amplified as follows: the nuclear ribosomal RNA internal transcribed spacer regions, which includes ITS1 spacer – 5.8S rRNA gene – ITS2 spacer, were amplified and later sequenced using 4 primers according to White et al. (1990). The primers trnK685F GTATCCGACTATGTCATTGTA and trnK2R* CCCCGAACTAGTCGGATGG were used for the amplification of the matK sequence as forward and reverse primers, respectively, as suggested by Wojciechowski et al. (2004) for Fabaceae. For sequencing, we used only trnK685F for about 700 bp for the run, corresponding to about half of the matK DNA fragment. The set of matK sequences was much smaller than the ITS set.

The ITS amplification was performed as follows: 180 s at 95 °C; followed by 28 cycles of 30 s at 95 °C, 60 s at 42
2.3. Sequence alignment and phylogenetic analysis

The boundaries of the fragments (about 700 bp for matK and 560 bp for ITS sequences) were determined by comparison with previously published sequences. All new accessions with a corresponding GenBank accession number are reported online in Table 1S (supplementary material: http://www.unifi.it/caryologia/tjb/).

Optimal multiple alignment was obtained with CLUSTALW 1.81 (Thompson et al., 1994) and checked by eye. The matrices were combined with the Python (Python version 2.6.4; Biopython 1.57) program combinex1_0.py, written by one of the authors, A Papini, which was released under GPL license and is available at www.unifi.it/caryologia/PapiniPrograms.html. The matrices are available by the authors as Table 2S for the combined matrix with matk+ITS+indels-derived characters and Table 3S with only ITS sequences+indels-derived characters (supplied as supplementary material at http://www.unifi.it/caryologia/tjb/Tab2Hedysaroid_comb.nex).

Three representatives of genus Caragana were used as outgroups for the phylogenetic analysis: Caragana korshinskii, Caragana microphylla, and Caragana arborescens. These outgroups were chosen according to the relationships of Onobrychis and allied genera outlined in recent molecular studies by Wojciechowski (2003), Wojciechowski et al. (2004), Lavin et al. (2005), and Ahangarian et al. (2007). Sequences described in these studies (75 ITS and 7 matK sequences) were also used in the analysis and not directly produced by us (GenBank accession numbers are supplied in Table 1S).

Parsimony analysis was performed with PAUP* version 4 (Swofford, 2002). A preliminary heuristic search was performed with multtrees off and 100 replicates with random addition. The obtained trees were used as a start for a successive analysis with multtrees on and 10 replicates (default settings in PAUP for hs command).

All characters were weighted equally, and character state transitions were treated as unordered. Gaps were treated as “simple indel coding” after Simmons and Ochoterena (2000), coding them with the software Gapcoder (Young and Healy, 2003). This process codes indels as separate characters at the end of the same DNA sequences data matrix (see Table 2S, supplementary material).

A maximum likelihood (Felsenstein, 1981) search was conducted as follows: MrModeltest 2.0 (Nylander, 2004) was used to test the best model of sequence evolution (based on the Akaike information criterion, Akaike, 1974). The model with the best score was used for settings in a maximum likelihood (ML) phylogenetic analysis in PAUP. The model obtained was used to calculate the likelihood value of the maximum parsimony trees.

The analysis was executed with the GARLI package, which is based on a stochastic genetic algorithm–like approach to simultaneously find the topology, branch lengths, and substitution model parameters that maximize the log-likelihood (lnL). The package was used on a server provided by the Cipres portal (Miller et al., 2009 for the site address). For maximum likelihood analysis, indel-derived characters were excluded.

Bootstrap (Felsenstein, 1985) resampling was performed setting search = faststep (with no TBR branch-swapping because of computational time limits) with 10 random taxon entries per replicate and the multtrees option in effect (with 10,000 replicates) under parsimony criterion.

A decay analysis was performed for Bremer support (Bremer, 1988) with AutoDecay version 5.0 (Eriksson, 2001) to assess the internal support for relationships obtained in the maximum parsimony heuristic analyses. MrModeltest 2.0 results were also used as an evolutionary model for the Bayesian analysis with MrBayes (Huelsenbeck and Ronquist, 2001). We used the same model for the indel-coded characters of the matrix as we did for restriction sites (coded as binary character states), as implemented in MrBayes. Bayesian analysis is particularly useful to treat mixed character sets (Nylander et al., 2004).

The Bayesian phylogenetic analysis was used to assess the robustness of tree topology and the support for clades. The posterior probability of the phylogenetic model was estimated using Markov chain Monte Carlo sampling with the Metropolis–Hastings–Green algorithm. Four chains were run, 3 heated and 1 cold, for 10⁶ generations and were sampled every 100 generations. Following the analysis, the posterior probabilities were checked in the output of MrBayes (in the file .p produced by the software) to estimate the number of trees that should be discarded as “burn-in” when the values reached stationarity (that is, it did not vary anymore out of a range). When stationarity was reached (quite stable values of the log likelihood scores), it was possible to evaluate how many of the beginning trees to discard as “burn-in.” After the “burn-in” trees were discarded, the remaining trees were used to estimate posterior probabilities.

The model obtained was used to test the best model of sequence evolution (based on the Akaike information criterion, Akaike, 1974). The model with the best score was used for settings in a maximum likelihood (ML) phylogenetic analysis in PAUP. The model obtained was used to calculate the likelihood value of the maximum parsimony trees.

The analysis was executed with the GARLI package, which is based on a stochastic genetic algorithm–like approach to simultaneously find the topology, branch lengths, and substitution model parameters that maximize the log-likelihood (lnL). The package was used on a server provided by the Cipres portal (Miller et al., 2009 for the site address). For maximum likelihood analysis, indel-derived characters were excluded.

Bootstrap (Felsenstein, 1985) resampling was performed setting search = faststep (with no TBR branch-swapping because of computational time limits) with 10 random taxon entries per replicate and the multtrees option in effect (with 10,000 replicates) under parsimony criterion.

A decay analysis was performed for Bremer support (Bremer, 1988) with AutoDecay version 5.0 (Eriksson, 2001) to assess the internal support for relationships obtained in the maximum parsimony heuristic analyses. MrModeltest 2.0 results were also used as an evolutionary model for the Bayesian analysis with MrBayes (Huelsenbeck and Ronquist, 2001). We used the same model for the indel-coded characters of the matrix as we did for restriction sites (coded as binary character states), as implemented in MrBayes. Bayesian analysis is particularly useful to treat mixed character sets (Nylander et al., 2004).

The Bayesian phylogenetic analysis was used to assess the robustness of tree topology and the support for clades. The posterior probability of the phylogenetic model was estimated using Markov chain Monte Carlo sampling with the Metropolis–Hastings–Green algorithm. Four chains were run, 3 heated and 1 cold, for 10⁶ generations and were sampled every 100 generations. Following the analysis, the posterior probabilities were checked in the output of MrBayes (in the file .p produced by the software) to estimate the number of trees that should be discarded as “burn-in” when the values reached stationarity (that is, it did not vary anymore out of a range). When stationarity was reached (quite stable values of the log likelihood scores), it was possible to evaluate how many of the beginning trees to discard as “burn-in.” After the “burn-in” trees were discarded, the remaining trees were used to estimate posterior probabilities.
removed from the data set, the remaining trees were used to produce a 50% majority-rule consensus tree with PAUP, in which the percentage support indicated a measure of the Bayesian posterior probabilities. The stationarity was reached at approximately generation 30,000, and so the first 300 trees (or the “burn-in” period of the chain) were discarded. Phylogenetic inferences are therefore based on those trees sampled after generation 30,000 for both the combined data set and the data set for only ITS.

The Templeton (Wilcoxon signed-ranks) test (Templeton, 1983), implemented in PAUP, was used to test the alternative less parsimonious topologies with respect to the most parsimonious tree. This test was used to evaluate the significance of an alternative position of taxa of *Onobrychis* s.l.

A partition homogeneity test was performed to check compatibility between the plastid sequence matK and the ITS sequences with PAUP version 4 (Swofford, 2002), with heuristic search, 100 replicates, and swap=none to reduce the computational effort.

The trees were edited for better readability with the program FigTree v1.3.1 by Andrew Rambaut, Institute of Evolutionary Biology, University of Edinburgh: http://tree.bio.ed.ac.uk/software/figtree/.

Supplementary materials (Figures S1–S4 and Tables S1–S3, with their legends, are in the file SupplOnobrychis.html) are available at www.unifi.it/caryologia/tjb/.

### 3. Results

#### 3.1. Sequence analysis

The total alignment with both markers consisted of 67 taxa and 1501 characters, of which 717 resulted from nucleotide sequence alignment of matK, 643 from the ITS sequences (ITS1+5.8SrDNA+ITS2), and another 140 characters as a result of indel coding (36 for the matK and 103 for the ITS). The partition homogeneity test in PAUP (Swofford, 2002) showed that the matK (plastid genome-encoded) and the ITS gene set were congruent at P = 0.01 (just P-value = 1 – (99/100) = 0.010).

![Figure 1](attach.png)

**Figure 1.** Majority rule consensus tree obtained from the Bayesian trees (excluding the “burn-in” trees) from the total evidence matrix formed by matK+ITS1+5.8SrDNA+ITS2 and indels coded as simple gaps. Robustness is indicated above branches: the first number corresponds to the Bayesian support, the second to the bootstrap (maximum parsimony) support, and the third to the decay values. The value is empty for values lower than 50% for Bayesian and bootstrap support and lower than 1 for the decay values. If only one number is present, it corresponds to the Bayesian support. In green, *Onobrychis* subgenus *Onobrychis* section *Onobrychis*; in yellow, *O*. subgenus *O*. section *Lophobrychis*; in pink, *O*. subgenus *Sisyrosema* section *Hymenobrychis*; in blue, *O*. subgenus *Sisyrosema* section *Heliobrychis*. 

984
3.2. MatK/ITS phylogenetic tree
The phylogenetic analysis, on the basis of the total evidence (matK+ITS) with the heuristic search, produced 109 trees 1111 steps long. Three of these trees were those with the best maximum likelihood value (calculated without considering indels) on the basis of the evolutionary models found with MrModeltest. One of these 3 trees is supplied as supplementary material (Figure S1).

The tree obtained as majority rule consensus trees of the Bayesian analysis trees (obtained with MrBayes) is shown in Figure 1.

Genus *Onobrychis* plus *Hedysarum boreale* (apparently inserted in *Onobrychis* subgenus *Onobrychis*) was supported with 100% Bayesian and bootstrap support, value of decay = 17. In fact, the analysis with matK alone (Figure S4, supplementary material) resulted in *H. boreale* clustering together with the other 2 accessions of *Hedysarum* considered in the analysis and not within *Onobrychis*.

Genus *Onobrychis* subgenus *Onobrychis* section *Onobrychis* plus *Hedysarum boreale* (in green color in Figure 1) had 100% Bayesian and bootstrap support and decay value = 11. O. subgenus *Onobrychis* section *Lophobrychis* Hand.-Mazz. was not monophyletic, since *O. pulchella*, *O. alba* subsp. *laconica*, and *O. crista-galli* formed a clade with *O. petraea*, while *O. aequidentata* and *O. caput-galli* were sister groups to section *Onobrychis*.

O. subgenus *Sysirosema* Bunge was supported as monophyletic, with 100% Bayesian and bootstrap support and autodecay value = 18. O. subgenus *Sysirosema* section *Hymenobrychis* DC. (in pink in Figure 1) was supported as monophyletic with 100% Bayesian support, 98% bootstrap support, and autodecay index = 11. O. subgenus *Sysirosema* section *Heliobrychis* Bunge (in blue in Figure 1) was a sister group to section *Hymenobrychis* and monophyletic with 100% Bayesian and bootstrap support and autodecay index = 6, even though only 2 accessions were sampled. The interspecific relationships in *Onobrychis*, especially within the subgenus *Onobrychis*, were not resolved. In fact, in some cases different accessions of the same species, such as *O. vicifolia*, clustered in a different point of the tree without forming monophyletic groups.

The maximum likelihood tree obtained with GARLI was very similar to that shown in Figure 1 (data not shown). A strict consensus tree of maximum parsimony for 1,014,420 trees (search stopped after 90 min) obtained only with matK sequences (including indel-derived characters) is supplied in the supplementary material as Figure S4 (www.unifi.it/caryologia/tjb/FigS4.pdf). In this tree, the accession of *Hedysarum boreale* clustered together with the other 2 accession of *Hedysarum* used in the analysis and not together with genus *Onobrychis*.

3.3. ITS phylogenetic tree
The analysis of the ITS data set showed that the genus *Onobrychis* was not monophyletic because of the presence, within *Onobrychis*, of 1 accession of *Eversmannia subspinosa* and 2 accessions of *Hedysarum*, *H. boreale* and *H. candidissimum* (Figure 2). The so-formed clade had 93% Bayesian support. *H. membranaceum* was a sister group to *Onobrychis* + *Eversmannia* with 93% Bayesian and 62% bootstrap support and autodecay index = 2 (Figure 2). A Templeton test was then performed with PAUP to test an alternative position of *H. membranaceum*, inserting this last species within *Hedysarum* s. s. The alternative tree was significantly different and 10 steps longer with respect to the maximum parsimony tree.

Within the genus *Onobrychis*, the subgenus *Onobrychis* was also monophyletic (98% Bayesian support, 86% bootstrap support, and decay index = 7) (Figure 2). Subgenus *Sysirosema* was monophyletic with 100% Bayesian support and decay index = 13. *Eversmannia subspinosa* was supported as a sister group to subgenus *Onobrychis* (88% Bayesian support and decay index = 2). An alternative hypothesis with the *Eversmannia* sister group to the whole genus *Onobrychis* produced a 2-step-longer tree. The difference was not statistically significant after the Templeton test.

3.4. Relationships within *Onobrychis*
Section *Onobrychis* plus 1 accession of *O. cyri* (shown in light green in Figure 2) formed a monophyletic group (89% Bayesian support). Section *Lophobrychis* (shown in yellow in Figure 2) was not monophyletic, since *O. caput-galli* and *O. aequidentata* were not included in it, but were sisters to section *Onobrychis*. Section *Dendrobychis* DC. (in dark green) was divided into 2, with 3 accessions of *O. cornuta* clustered within the main part of section *Lophobrychis* and *O. arnacantha* (considered as belonging to subgenus *Sysirosema*) in an unresolved position with respect to the recognized sections of this subgenus. *O. petraea* clustered together with *Lophobrychis* + part of *Dendrobychis*.

O. subgenus *Sysirosema* (Figure 2: fuchsia, blue, gray, red, and a basal dark green branch) was supported as monophyletic with 100% Bayesian and bootstrap support and autodecay value = 13, with the exception of *O. arnacantha* (section *Dendrobychis*, in dark green), taxonomically assigned to subgenus *Onobrychis*. Subgenus *Sysirosema* was formed by sections *Hymenobrychis* (in fuchsia) + *Heliobrychis* (in blue) + *Laxiflorae* (Sirj.) Rech.f. (in red) + *Afghanicae* Sirj. (in gray). O. subgenus *Sysirosema* section *Hymenobrychis* (in pink in Figure 2) was supported as monophyletic with 100% Bayesian support and decay index = 4, provided that we consider *O. acaulis* (taxonomically, this is considered to belong to section *Anthyllium* Nábělek) inserted in *Hymenobrychis*. O. subgenus *Sysirosema* section *Heliobrychis* (in blue in
Figure 2. Majority rule consensus tree obtained from the Bayesian trees (excluding the "burn-in" trees) from the ITS matrix formed by ITS1+5.8SrDNA+ITS2 and indels coded as simple gaps. Robustness is indicated above branches: the first number corresponds to the Bayesian support, the second to the bootstrap (maximum parsimony) support, and the third to the decay values. The value is empty for values lower than 50% for Bayesian and bootstrap support and lower than 1 for the decay values. If only one number is present, it corresponds to the Bayesian support. N= corresponds to the available data about the chromosome number. In green, Onobrychis subgenus Onobrychis section Onobrychis; in yellow, O. subgenus O. section Lophobrychis; in light blue, O. subgenus O. section Dendrobrychis; in fuchsia, O. subgenus Sisyrosema section Hymenobrychis; in blue, O. subgenus Sisyrosema section Heliobrychis; in brown, O. subgenus Sisyrosema section Laxiflorae; in gray, O. subgenus Sisyrosema section Afghanicae; in red, Hedysarum membranaceum. For karyological data references see the text (Section 4.2).
Figure 2) was supported as monophyletic (100% Bayesian support, 87% Bootstrap support, and decay index = 4) and sister group to section Hymenobrychis. Section Laxiflorae was in the basal position of the subgenus in an unresolved position with respect to O. arnacantha and the clade formed by sections Hymenobrychis + Heliochyris + Afghanicae. Section Afghanicae (represented here by a single accession, O. numularia) was an outgroup to Hymenobrychis + Heliochyris.

Even in the ITS data set (with more taxa than the combined set), the interspecific relationships in Onobrychis and particularly within subgenus Onobrychis were not easily resolved. The ITS marker does not produce trees that keep all the accessions of the same species into monophyletic groups.

The maximum likelihood tree obtained with GARLI is supplied as supplementary material (Figure S2) together with one of the maximum parsimony trees with the best maximum likelihood score obtained with PAUP on the basis of the MrModeltest settings (Figure S3). These trees supported H. membranaceum as a sister group to genus Onobrychis+Eversmannia and the position of Eversmannia as a sister group to Onobrychis subgenus Onobrychis and of O. petraea within section Lophobrychis. Some of the maximum parsimony trees (as in Figure S3) positioned genus Ebenus as a sister group to O. subgenus Sysirosema. This alternative topology had Bayesian and bootstrap support lower than 50% and decay index of <1, but the Templeton test showed that, at least with the parsimony criterion, the difference was not statistically significant (data not shown).

3.5. Indels in the combined matK and ITS matrix
The combined matrix is provided as Table S2 (supplementary material: http://www.unifi.it/caryologia/tjb/): the indicated indel positions are referred to in Table S2. Three indels were shared by the outgroups and Alhagi (1153, 1157, and 1297 of the combined alignment). Three indels were shared by the outgroups, Alhagi, and 2 accessions of Hedysarum (919, 947–949, and 976 of the combined alignment). The indel in position 1146–1153 was shared by Onobrychis sect. Lophobrychis plus O. petraea. An indel in 1197 was shared only by the 2 accessions of Hedysarum vicifolia.

The ITS matrix was composed of 153 accessions for 897 positions, of which 1–285 belong to the ITS1, 286–463 to the 5.8S rDNA, 464–711 to the ITS2, and 712–897 to the indels, coded as simple gaps.

3.6. Indels found only in the ITS matrix
Relative only to the ITS alignment, an indel in 35–36 was shared by Alhagi plus Eversmannia. Five indels in 71–73, 108–112, 208, 299–300, and 473 characterized the whole genus Ebenus. An indel in 87–88 characterized Hedysarum membranaceum+Onobrychis subgenus Sysirosema. An insertion in position 80 was shared by Eversmannia, Onobrychis subgenus Sysirosema, and Hedysarum membranaceum. An insertion in 471–477 was shared by O. petraea + section Lophobrychis.

4. Discussion
4.1. Phylogenetic relationships
The results of the partition homogeneity test showed that the 2 data sets, the plastid partial matK sequence and the nucleus encoded sequence ITS, were congruent only at P = 0.01 (P-value = 1 – (99/100) = 0.010). This P-value is just the threshold at which combining 2 data sets would improve phylogenetic confidence after Cunningham (1997). After Rokas et al. (2003), concatenating more sequences in a single matrix would reduce the total number of maximum parsimony trees. For these reasons an initial analysis was performed with the combined data set, followed by a focus on the larger ITS data matrix. We also supplied the only matK strict consensus tree of 1,014,420 trees (PAUP maximum parsimony heuristic search with outgroup Caragana, search stopped after 90 min), as in Figure S4 in the supplementary material.

The combined (matK+ITS) phylogenetic analysis showed that an accession of Hedysarum boreale clustered together with genus Onobrychis, as already observed by Ahangarian et al. (2007). However, the analysis using matK only (Figure S4, supplementary material) positioned H. boreale together with the other 2 accessions of Hedysarum. This matK sequence (AY386892 by Wojciechowski, Lavin, and Sanderson from a sample from Arizona: Wojciechowski 259) was obtained from a different sample with respect to the ITS sequence. This last was the same used by Ahangarian et al. (2007) in his study, that is U50482 for the ITS1 and U50483 for the ITS2, both by Sanderson and Wojciechowski from the same sample “Wojciechowski and Sanderson 131”. The sample of H. boreale may be of hybrid (intergeneric!) origin, since the maternally inherited plastid sequence resulted in a different phylogenetic position with respect to the nuclear ITS. Alternatively, the original samples (or at least one of them) of H. boreale may have been wrongly identified.

The accession of Hedysarum candidissimum (within Onobrychis in our results) was published in GenBank by Ahlquist and Wojciechowski, voucher M.Nydedegger 42636 (MSB). Since this sequence was not yet employed in other phylogenetic analysis, it was not discussed further in this study. As H. boreale, H. candidissimum may also be
really more related to genus Onobrychis than to the rest of Hedysarum and further DNA markers may be necessary to clarify their positions (Martin F Wojciechowski, personal communication).

After Martin F. Wojciechowski (personal communication), Hedysarum and Onobrychis, both containing well over 100 species, are probably paraphyletic and relationships are fluid. In view of this uncertainty, these species were omitted from the following discussion.

In the total evidence matrix, O. subgenus Onobrychis section Lophobrychis was not monophyletic since O. pulchella, O. alba subsp. laconica, and O. crista-galli formed a clade with O. petraea, while O. aequidentata and O. caput-galli, belonging to this section after most treatments, were sister groups to section Onobrychis. However, very variable chromosome numbers have been documented for O. aequidentata: 2n = 14, 16, and 28 (De Montmollin, 1984; Romano et al., 1987; Baltisberger, 1991; Abou-El Enain, 2002). The same values were documented for O. caput-galli (Heyn, 1962; Slavivk et al., 1993; Abou-El Enain, 2002). Such variation (even apparently with different base chromosome numbers: 7 and 8) may suggest the presence of a different species poorly characterized from a morphological point of view, or even the presence of hybrids or species of hybrid origin, under the names O. aequidentata and O. caput-galli. The not “orthodox” position within the section Lophobrychis may be justified by such variability. The difficult phylogenetic reconstruction due to species of hybrid origin is well known in “difficult” genera such as Quercus (Fagaceae) (Simeone et al., 2013) or Rebutia (Cactaceae) (Mosti et al., 2011).

On the basis of the molecular data here presented, genus Onobrychis cannot be considered monophyletic without inserting it at least the genus Eversmannia (which was nested within Onobrychis). Conversely, the Templeton test showed that an alternative tree topology with Eversmannia as sister group to Onobrychis was 2 steps longer (with parsimony), but not significantly different. Eversmannia may be seen as a specialized version of Onobrychis with many autapomorphies, even at the molecular level. This uncertainty suggests that a larger sampling of Eversmannia in particular and Hedysareae in general is necessary to clarify the position of this genus before suggesting its eventual transfer into Onobrychis.

Ahangarian et al. (2007) separated Hedysarum membranaceum as sister group of the genus Sulla (88% bootstrap support), while Eversmannia was put into a sister group of the genus Onobrychis (see maximum parsimony trees). However, these analyses were undertaken on a much smaller sample with respect to the data presented here and only under the maximum parsimony criterion. The position of H. membranaceum as a sister group of Onobrychis was confirmed by the Templeton test (alternative positions resulted in trees with a statistically significant difference).

H. membranaceum is the only species within the monotypic section Membranacea B.Feldsch. of the genus Hedysarum. This species is restricted to north Africa and is morphologically distinct from the other Hedysarum species and the other genera of Hedysareae in that it has pods with wide wings up to 3 mm wide, short inflorescences, ovate standard, wings with a short auricle, a keel with a short claw, and a protruded hilum in the seed (Choi and Ohashi, 2003). Choi and Ohashi (1996) noted that H. membranaceum Cass. et Bal. is intermediate between sect. Fruticosa and other species of Hedysarum in terms of pollen morphology, petal shape, and a suffrutescent habit. H. membranaceum turned out to be the most divergent species among the Mediterranean representatives of the genus Hedysarum on the basis of ISSR data (Chennaoui-Kourda et al., 2007) and in a previous analysis with ITS data restricted to 8 species of Hedysarum s. l. (Chennaoui et al., 2007). The molecular data, together with the results of the Templeton test and the morphological features (Choi and Ohashi, 2003), indicate that this taxon may be recognized as a separate genus with respect to Hedysarum if further morphological and/or molecular characters confirm the here-observed phylogenetic position.

The position of O. petraea, previously positioned in O. section Onobrychis subsection Macropterae Hand.-Mazz., was nested within Lophobrychis+Dendrobrychis. On the basis of these data, we suppose that the sectional treatment may deserve some changes. On the basis of the molecular data alone, in order to form a monophyletic clade corresponding to a section, part of Lophobrychis, Dendrobrychis, and subsection Macropterae of section Onobrychis (O. petraea) should be kept together. Further molecular data may be useful to further clarify the relationships of these species.

Onobrychis subgenus Sysirosema was resolved as monophyletic with high support and should therefore be maintained. At the sectional level, O. subgenus Sysirosema section Helioibrychis was monophyletic and a sister group of section Hymenobrychis. This result confirmed that of Arslan and Ertuğrul (2010), who used seed storage proteins as molecular markers. The monophyly of sections Laxiflorae and Afghanicae is not excluded by the here-presented results, even if the relationships of section Laxiflorae were not completely clear. A gametophytic chromosome count of n = 7 for O. laxiflora (Kathoon and Ali, 1991) may suggest this number as the basal number for the subgenus Sysirosema. Section Afghanicae is positioned as a sister group to Helioibrychis + Hymenobrychis.

The presence of O. arnacantha (section Dendrobrychis, in dark green in Figure 2), taxonomically assigned to subgenus Onobrychis, in a basal position close to subgenus
Sysirosema was quite unexpected. The same result was obtained by Ahangarian et al. (2007), who suggested that some morphological features, similar to those of other species of section Dendrobrychis (as O. cornuta, inserted in the analysis), were due to parallelism.

O. acaulis was nested within section Hymenobrychis despite it currently being classified as belonging to section Anthyllium. Such incongruence may be due to erroneous assignment of O. acaulis to section Anthyllium, or a wrong assignment of the status of the section to this last group of species. Section Anthyllium did not appear to be sufficiently separated from section Hymenobrychis, but better sampling within this section is necessary before eventual definitive taxonomic rearrangements.

An alternative tree topology with genus Ebenus as a sister group to O. subgenus Sysirosema appeared in some maximum parsimony trees. This position had Bayesian and bootstrap support lower than 50% and decay index of <1, but the Templeton test showed that, at least with the parsimony criterion, the difference was not statistically significant. For this reason it was also decided that the current status of Eversmannia should not be changed (together with a current insufficient sampling of species within this genus and the lack of data for other small genera within Hedysareae).

Results from this study showed that the genus Hedysarum was polyphyletic, and to a greater extent than was already proposed by Choi and Ohashi (2003). Phylogenetic analysis showed that it was resolved into 4 different clades. One clade separated from Hedysarum, corresponding to genus Sulla (already separated from the rest of Hedysarum by Choi and Ohashi, 2003), Hedysarum s. s. with the type species H. alpinum, and a last group resulted in a sister group to Taverniera and H. kumaononense. We did not make further taxonomic decisions in relation to Hedysarum (apart from H. membranaceum) since sampling in the here-presented data favored Onobrychis. Moreover, some important genera related to Hedysarum, such as Sartoria and Corethodendron, are not represented here.

The key to the genera of Hedysarum and allied genera as proposed by Choi and Ohashi (2003) still remains valid even though taxonomic rearrangements have been proposed as a result of this study.

4.2. Karyotype evolution in Hedysareae

The known chromosome numbers are indicated in Figure 2. On the basis of the data available from the IPCN chromosome numbers database, genera Allhagi, Sulla, and Taverniera have a basic number of n = 8. These 3 genera are those resulting sister groups to the rest of Hedysareae after our phylogenetic analysis (Figure 2). Astragalus and Chesneya of the related Astragallean clade (Sepet et al., 2011) also have a basic number of n = 8. A count of 2n = 18 in Sulla coronaria (as Hedysarum coronarium) is known (Issolah et al., 2006), such that Arslan et al. (2012) also consider n = 9 to be a possible basis chromosome number in Hedysarum s. l. The count 2n = 16 was also found in the genus Sartoria (not sampled here; Arslan et al., 2012). Hedysarum s. s. (the clade containing the genus type H. alpinum) has n = 7, apart from 1 count of n = 8 for an accession of H. alpinum. The same number (7) was also documented for the genus Ebenus. This last genus was the sister group of the clade containing Hedysarum membranaceum+Onobrychis+Eversmannia. A count of n = 8 is available for Eversmannia. The situation is more complex in Onobrychis. Within the subgenus Sysirosema, the chromosome number is always n = 7 in section Hymenobrychis, apart from a count of 2n = 16 for Onobrychis galegifolia, not sampled here and which would deserve further testing, and some counts of 2n = 16 for O. subnitens (Ranjbar et al. 2012), while n = 8 would be the base chromosome number in section Heliochrys.

This last section appears to be homogeneous regarding chromosome number and DNA sequence evidence, while it appeared quite variable in morphological characters (Karamian et al. 2012). Within the subgenus Onobrychis, n = 8 is known for sections Dendrobrychis and Lophobrychis. These 2 sections clustered together in the phylogenetic analysis and were the sister groups to section Onobrychis, which only had known chromosome counts of n = 7. Comparing the phylogenetic analysis with the karyotype data, we can assume a base number n = 8 for the tribe (present in the more basal genera), changing to 7 in Hedysarum s. s. and Ebenus. This number would be maintained in Onobrychis subgenus Sysirosema, while n = 8 in sect. Heliochrys would be a derived condition. The change to n = 8 would occur also in Eversmannia, a possible sister group to subgenus Onobrychis in some of the presented trees based on the ITS (Figure 2). The same number is maintained in sect. Dendrobrychis and sect. Lophobrychis, while n = 7 in section Onobrychis would be again a derived condition.

A variation in chromosome number and ploidy level is known for some species, particularly for Onobrychis subgenus Onobrychis section Onobrychis. Accessions of O. altissima, for instance, have 2n = 14 (Arslan et al., 2012) and others have 2n = 28 (Hejazi et al., 2010), while a variation of 2n = 22, 27, 28, 29 was found in O. vicifolia. After Ranjbar et al. (2010), O. altissima is considered to be closely related to the cultivated sainfoin (O. vicifolia) and may be a progenitor of it, both, based on morphological similarity, a close relationship between the 2 species was postulated by Hedge (1970). Gömürgen (1996) found also cases of meiotic chromosome instability in O. armena. This chromosome number variation even within species may suggest the presence of cryptospecies with similar morphology but a different chromosome number, at
least in some of the species of this section. The presence of cryptospecies may partly explain the failure of the ITS sequence to distinguish between species within section Onobrychis. Moreover, polyplody within some taxa may be due to ITS polymorphism, even with the possible formation of paralogues. This phenomenon may cause difficult phylogenetic reconstruction and possible conflicts between nuclear and chloroplast sequences, as in the genus Quercus (Bellarosa et al., 2005; Simeone et al., 2009; Papini et al., 2011).

5. Conclusion

The main results of this analysis are: genus Onobrychis may be not considered monophyletic, since Eversmannia was nested within it, while the position of the monotypic section Membranacea B. Feldtsch. of genus Hedysarum (containing only one species, H. membranaceum Coss. & Bal.) was supported as a sister group of Onobrychis+Eversmannia. The 2 markers that were employed (ITS and matK) were able to distinguish between genera within the tribe Hedysareae and intrageneric relationships in Onobrychis and Hedysarum. Nevertheless, these markers were unable to fully distinguish between species of Onobrychis, particularly in sect. Onobrychis. Published cytogenetic data may suggest that this result is linked to difficult species circumscription (various chromosome numbers in polyploid series are present within the same species). The phylogenetic analysis showed that the most basal clades of the tribe have n = 8 as a basic chromosome number, with n = 7 appearing in a clade corresponding to Hedysarum s. s. and Ebenus. These last groups were clustered as a sister group to Onobrychis+Eversmannia+Hedysarum membranaceum. The chromosome numbers followed a variable pattern in Onobrychis. Subgenus Sysirosema section Hymenobrychis has a chromosome number of n = 7 while section Heliobrychis has n = 8. This chromosomal switch may be the basis of the separation of these sections. In subgenus Onobrychis, the sections Dendrobrychis and Lophobrychis have n = 8, while the more derived section Onobrychis has again n = 7. A count of n = 8 was reported for Eversmannia, in agreement with the molecular phylogenetic analysis (indicating a sister group position with respect to subgenus Onobrychis).

We did not take taxonomic decisions in relation to Hedysarum, since sampling in the here-presented data favored Onobrychis. Moreover, some important genera related to Hedysarum, such as Sartoria and Corethodendron, are not represented here.

Acknowledgments

We acknowledge the support from the EU (MRTN-CT-2006-035805; ‘Healthy Hay’ project) and from Dr Massoud Ranjbar, renowned expert and taxonomist defining new species of Onobrychis, Department of Systematic Botany, Bu-Ali Sina University, Iran, for support and supply of herbarium samples for analysis. We also thank the anonymous reviewers who helped us to improve the quality of the manuscript.

References


