Pak. J. Bot., 41(4): 1577-1586, 2009.

A MOLECULAR PHYLOGENETIC STUDY OF RED BUDS (*CERCIS* L., FABACEAE) BASED ON ITS nrDNA SEQUENCES

FATİH COŞKUN* AND CLIFFORD R. PARKS

Department of Biology, Balıkesir University, 10145, Balıkesir, Turkey
University of North Carolina at Chapel Hill, Chapel Hill, N.C., 27599-3280, U.S.A..

Abstract

As a member of the large family Fabaceae, red buds are widely cultivated in the world due to their ornamental value. This study included 13 *Cercis* taxa from different parts of the world. Our analysis of the ITS nrDNA sequences proved useful in understanding the phylogenetic relationships of *Cercis* taxa and resolved most of the branches in the phylogenetic tree. The lowest sequence divergence within ingroup taxa was between *C. canadensis* ssp. *canadensis* and *C. californica* ssp. *californica*, 0.0014%. This was assuring that these taxa belong to the same species. The highest sequence divergence within ingroup taxa was 0.028% between *C. canadensis* ssp. *mexicana* and *C. chingii*. ITS data indicated that *C. chuniana* and *C. occidentalis* are interestingly close relatives, on one hand and *C. siliquastrum* and other North American *Cercis* taxa along with *C.griffithii* are close relatives on the other.

Introduction

Red buds, *Cercis* L., are ornamental plants widely cultivated in the world, specifically in the northern gardens e.g., in North America, varieties of *Cercis canadensis* and in Europe, varieties of *C. siliquastrum*. The bright white to reddish pink color of different plant cultivars is especially attractive to gardeners and garden lovers in early spring. This genus is distinctive in showing the cauliflory, the production of flowers directly on the stem or trunk before the growth of leaves. *Cercis* belongs to the subfamily Caesalpinioideae and is a genus of about 10 species of shrubs or small trees widely scattered across the North Temperate Zones of Eurasia and North America (Fig. 1). Table 1 shows the detailed distribution of this genus in the world based on current literature and floras (Li, 1944; Ball, 1968, Chamberlain & Yaltirik, 1970; Isely, 1975; 1988; 2002).

Li (1944) worked on the taxonomy and distribution of the genus *Cercis* in China and constructed a key to the species of the genus based on morphology and geographic distribution. He emphasized that eastern Asia's likelihood to be the center of development for the genus was not less than North America and/or Eurasia. By his time, about 25 species were described. He recognized only 5 species for China: *Cercis racemosa* Oliver, *C. chuniana* Metcalf., *C. chinensis* Bunge, *C. chingii* Chun and *C. pauciflora* Li (now recognized as a synonym of *C. chinensis* Bunge). Li also recognized two North American viz., *C. canadensis* L. and *C. occidentalis* Torrey ex Gray red bud species based on the literature available in his time. He also did not make an effort to subgroup the species of the genus *Cercis*.

Isely (1958, 1973 and 1975) has also studied the Leguminosae of the United States and he constructed keys to the genera of the subfamilies of Leguminosae and keys to the species of those genera. In his work, Isely (1975) constructed a key to the species of *Cercis* distributed in the United States and gave special references to the other species of

*E-mail: fcoskun@balikesir.edu.tr Fax: +90 (266) 612-1265.

this genus distributed in Eurasia known at that time. Although Isely (1975) constructed keys to the species of *Cercis*, he also did not attempt to group them into subgenera and/or sections based on their relationships. He based his key purely from morphological and geographical data. A recent study of red buds was also carried out simultaneously with our work. Davis *et al*., (2002) worked on phylogeny of *Cercis* using one chloroplast marker (3' end of *ndh*F gene) and one nuclear marker (ITS). A current review of the literature on this genus is given by Coskun (2003).

The present report will discuss the phylogenetic relationships in *Cercis* by presenting evidence from ITS nrDNA sequences. Our questions are as follows: What kind of phylogenetic relationships available among ingroup and outgroup taxa (e.g., presence of monophyly, paraphyly, or polyphyly?).

Materials and Methods

Thirteen *Cercis* taxa were called from all the geographic regions of the world in which this genus is present (Fig. 1 and Table 1). It included eight *Cercis* species with three varieties accepted widely in the literature viz., *C. canadensis* var. *canadensis*, *C. canadensis* var. *texensis*, *C. canadensis* var. *mexicana*, *C. occidentalis*, *C. chinensis*, *C. chingii*, *C. chuniana*, *C. glabra*, *C. racemosa* and *C. siliquastrum*. Our analysis also included two additional taxa: one taxon from western United States, *C. californica*, subspecies *californica* (not listed in the literature) and one from eastern Asia, *C. yunnanensis*, now recognized as the synonym of *C. glabra* (for this synonymy, see Flora of China, 1988, Vol. 39, page 142).

Cercis plant materials collected and used in this study were vouchered as herbarium specimens and were deposited in the Herbarium of the University of North Carolina (NCU).

Outgroup selection: Outgroup selection for the study group of plants included in this work was based on the previous workers analyses. Based on morphology, Bentham (1840), Polhill *et al*., (1981) Wunderlin & Larsen (1981) suggested that the closest relatives of red buds are orchid trees, *Bauhinia* L. This arrangement was supported by Doyle (1995) based on rbcL DNA sequence data (phylogenetic data). The current model using morphological and molecular data would suggest that both of these genera are basal groups of plants in the Caesalpinioids (Polhill *et al.,* 1981; Wunderlin & Larsen, 1981; Doyle, 1995). Thus one *Bauhinia* species, *B. faberi*, was sampled as an out-group that is closely related to *Cercis* in this study.

Genomic DNA isolation and amplification: Total genomic DNA was initially extracted with a modified version of the 'hot' CTAB method outlined in Doyle & Doyle (1987) for all plants included in this work. Either 2 g fresh or 0.5 g silica gel-dried leaf tissue was ground in liquid nitrogen, and added to 20 mL hot (65ºC) 2x CTAB buffer as described in Doyle & Doyle (1987). Then the mixture was incubated in 65ºC for 10 minutes and extracted with 24/1 ratio of chloroform/isoamyl alcohol, respectively. The DNA was then precipitated with 2/3 volume isopropyl alcohol at -20ºC overnight. DNA extracts were suspended in 500 to 1000 µL of sterile distilled, deionized water (ddH_2O) and stored at -20ºC. Later, Qiagen company's DNeasy Plant Mini Kit was used to extract plant genomic DNAs following the manufacturer's protocol.

Primer name	5' to 3' Primer sequence	Primer designed by	Based on (the source publication)
Forward			
ITS5A (Angiosperm)---	CCTTATCATTTAGAGGAAGGAG	Kenneth J. Wurdack	White *et al*., 1990
Reverse			
ITS4----------	TCCTCCGCTTATTGATATGC	Bruce G. Baldwin, 1992	White *et al*., 1990

Fig. 2. ITS primers used in this study with their designers.

Molecular marker analyzed in this study is ITS nuclear ribosomal DNA (nrDNA) for *Cercis* species. Polymerase Chain Reaction (PCR) amplifications of ITS region of nuclear ribosomal DNA (nrDNA) were performed using a new primer, ITS5angiosperm (ITS5a, designed by Kenneth Wurdack) and a well-known primer, ITS4 (White *et al*., 1990), for all taxa included in this work (Fig. 2).

Double stranded DNA amplifications were performed in 35 μL volume containing 28 μL sterile deionized, distilled water, 3.5 μL 10x Taq DNA polymerase PCR buffer (GibcoBRL, Life Technologies or Qiagen companies), 1.05 μL MgCl2 GibcoBRL (Life Technologies or sometimes used 'Q solution' which includes $MgCl_2$, by Qiagen), 0.7 μL 200 μM dNTPs in equimolar ratio (either by Qiagen or GibcoBRL), 2 μL of each 10 μM primer, 0.175 μL Taq DNA polymerase enzyme (either Qiagen or GibcoBRL). For some amplifications of the GC-rich DNA templates, 0.5 to 3 μL 10% Bovine Serum Albumine (BSA) and/or DiMethylSulfOxide (DMSO) were added to the total reaction volume depending on the experience of initial trials of the PCR amplifications. During amplification of ITS nrDNA region, the following PCR amplification protocols were performed in the thermal cycler machine (Perkin-Elmer Applied Biosystems, Inc. model 377): the first cycle was at 95ºC for 1 minute and 15 seconds for denaturation of double stranded DNA. The following 30 more cycles were performed using 1 minute at 94ºC for more denaturation time, 1 minute at 55ºC for annealing, and 2 minutes and 30 seconds for primer extension; an additional 8 minutes of extension time followed the final cycle. In order to check whether PCR Master Mix was contaminated with any DNA or not, negative controls were used in all PCR amplifications. In order to judge the fact that optimum PCR amplification conditions are provided, positive controls were also included in most sets of amplifications.

PCR products were purified using 'Qiaquick PCR purification Kit' (Qiagen) following the instructions directed by the company. Both strands of DNAs were sequenced for all taxa and the sequences were generated from two or three different individuals for each taxon.

Initially, cycle sequencing reactions were performed at Clifford R. Parks Lab., Coker Hall, Department of Biology at UNC-Chapel Hill using Perkin-Elmer Applied Biosystems, Inc. according to manufacturer's protocols (i.e., Cycle sequencing 1: at 96ºC for 4 min.; Cycle sequencing 2: at 96ºC for 30 sec., at 50ºC for 15 sec., and at 60ºC for 4 min. in total of 30 cycles). First, cycle-sequenced products were cleaned by using Sephadex columns, vacuum dried and mailed to Iowa State University's DNA Sequencing Facility for final automated sequencer-generated data collection. Later, purified PCR products were sent to UNC-Chapel Hill DNA Sequencing Facility for cycle sequencing reactions and automated sequencer-generated data collection. Sequence data generated through automated methods were manually edited for each taxon using the

commercial software Sequencher version 3.1 for Macintosh computers, 1998 (Gene Codes Corporation) and assembled into consensus sequences (contigs). Generated DNA sequences were submitted to the Genbank and accession numbers obtained from Genbank were given in Table 1.

Data analysis: The ITS region of nrDNA consensus sequences were first aligned using the software "MultAlin" by Corpet (1988), available free on Internet at http://prodes.toulouse.inra.fr/multalin/multalin.html. Then they were visually checked and manually edited, if necessary.

The data analysis followed using PAUP* Version 4.0b8 for Macintosh (PPC) (Phylogenetic Analysis Using Parsimony and Other Methods, Swofford, 2001). Pairwise distances using Jukes-Cantor model as estimator were generated using PAUP* software.

Branch-and-Bound searches were executed to find the most parsimonious ITS trees. Branch-and-Bound search computed *via* "stepwise addition sequence" using "furthest" option, keeping minimal trees only, and saving all trees. Heuristic search used stepwise addition with "simple" addition sequence, 'swapping on best trees only' option and employing the 'Tree Bisection-Reconnection (TBR)' algorithm for branch swapping. Parsimony analyses included following search options: General search options with collapsing branches if maximum length is zero. Character state optimization followed Accelerated transformation (ACCTRAN). Stepmatrix options utilized allowing assignment of states not observed in terminal taxa to internal nodes using all states in stepmatrix. All informative base-pair differences were used in the analysis, multistate taxa were interpreted as "uncertainty", and gaps were treated as "missing data". During the analyses, several statistical measures were utilized including: bootstrap (Felsenstein, 1985) with 1000 replicates; consistency indices (Kluge & Farris, 1969); decay indices (Bremer, 1994) calculated by using the software Autodecay version 3.0 by Torsten & Wikstrom (1995) and PAUP* 4.0b10 (2002); retention indices (Farris, 1989), homoplasy indices (Kluge & Farris, 1969) and g1 statistic (Hillis & Huelsenbeck, 1992), obtained by generating 1,000,000 random trees.

Results and Discussion

Analyses of the ITS region of nrDNA showed pairwise differences ranging from 0.57% between *C. californica* subsp. *californica* and *C. canadensis* var. *texensis* to 2.59% between *C. canadensis* var. *mexicana* and *C. glabra* (Table 2). The lowest sequence divergence within ingroup taxa was between *C. canadensis* ssp. *canadensis* and *C. californica* ssp. *californica*, 0.0014%. This was assuring that these taxa are members of the same species. The highest sequence divergence within ingroup taxa was between *C. canadensis* ssp. *mexicana* and *C. chingii*, 0.028%. A complete, aligned data matrix of ITS nrDNA region of *Cercis* taxa can be seen in Appendix 1. Branch-and-Bound search of the ITS data generated two most parsimonious phylogenetic trees with a 0.94 consistency index value (Fig. 3). The analysis supported the genus *Cercis* as a monophyletic sister group to *Bauhinia* and related *C. siliquastrum* and *C. griffithii*, with all North American taxa except western red bud, *C. occidentalis*. Western red bud also showed a close affinity with *C. chuniana* and a close relationship with *C. yunnanensis*, *C. glabra*, *C. racemosa* and *C. chingii*. Bootstrap analysis resulted in a high support for a clade including *C. canadensis*, *C. canadensis* var. *texensis*, *C. canadensis* var. *mexicana*, *C. siliquastrum* and *C. griffithii*. Monophyly of *C. racemosa* and *C. glabra* was clear based on the ITS data and received 90% bootstrap support. The monophyletic group *C. occidentalis* and *C. chuniana* received more than a moderate bootstrap support (66%) (Fig. 3).

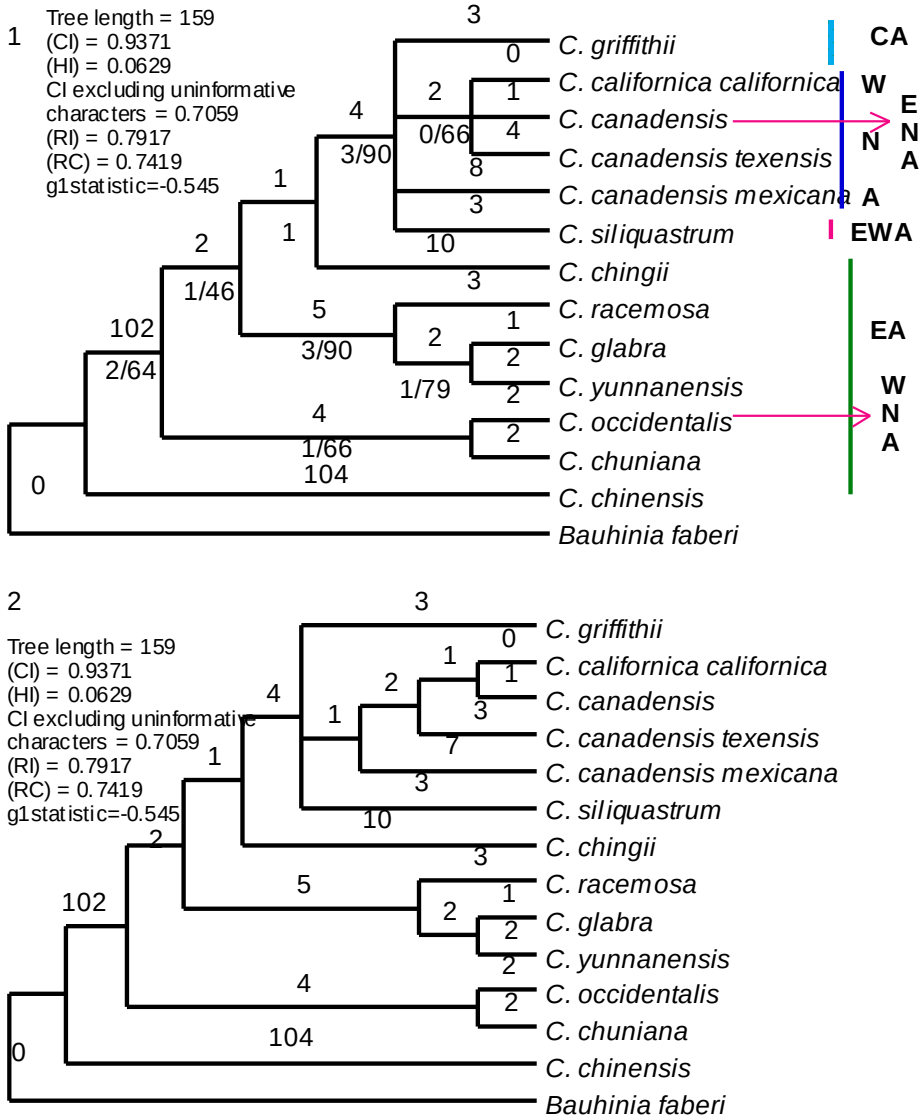

Fig. 3. Two equally most parsimonious ITS trees of *Cercis* taxa after a Branch-and-Bound Search. Branch lengths above branches, decay index values below the branches to the left or alone, and Bootstrap values to the right. CA: Central Asia, EA: Eastern Asia, ENA: Eastern North America, EWA: Europe and Western Asia, WNA: Western North America.

Analyses of data set indicated that ITS region of nrDNA is a useful marker to estimate the *Cercis* phylogeny. It resolved well for most of the relationships among the *Cercis* taxa. Interestingly, *C. occidentalis* showed a close affinity with eastern Asian red buds than rest of the North American red buds (Fig. 3). A well supported clade consisting of North American *Cercis* taxa (excluding *C. occidentalis*), and Eurasian taxa (*C. siliquastrum* and *C. griffithii* with one eastern Asian species, *C. chingii*) exhibited close

relationship with each other. *C. yunnanensis* formed a monophyletic group (79% bootstrap support) with *C. glabra*. This result supports the synonymy of *C. yunnanensis* with respect to *C. glabra* (see also Chinese version of the Flora of China for this synonymy, Vol. 39, p. 142). ITS data also supports the recognition of the North American *Cercis* taxa (excluding *C. occidentalis*) as varieties of one species. Based on the phylogenetic data analysis, it appears that *C. occidentalis* and *C. chuniana* shared the same common ancestor (Fig. 3).

Davis *et al*., (2002) worked on the phylogeny and biogeography of the genus *Cercis*, simultaneous with our study using different number of taxa and using another molecular marker. They found similar but not the same results obtained by our study. Different results were possibly due to the following conditions: Although they used the same ITS marker, their sequences were shorter than the sequences generated in this study since they used different forward primer for sequencing reactions. They also employed less number of taxa than this work such as the use of 11 number of taxa in their analysis whereas this study used 14 number of taxa. They did not include the central Asian red bud (*C. griffithii*), *C. chuniana* and *C. californica* ssp. *californica* whereas our study did not use *C. gigantea* as they did. Both works supports the derivation of North American and Eurasian taxa from eastern Asian taxa (Fig. 3). Although they attempted to employ a molecular clock and assign certain time intervals on phylogenetic tree branches, they did not try to classify this genus into subgenera or sections so didn't we in this work. In order to classify this genus into subgroups, a more elaborate analysis including morphological, molecular, biochemical and biogeographical data would prove helpful in understanding the systematics of this genus.

Acknowledgement

This study was supported by National Science Foundation (NSF Grant Int. 9710505 given to Cliff R. Parks) and by Turkish Higher Educational Council via Balıkesir University, Turkey (A Scholarship for Ph.D. studies in the U.S. granted to Fatih Coskun). Authors thank to J.C. Raulston Arboretum, U.S. National Arboretum, North Carolina Botanical Gardens, Forest Farm Nursery of Oregon, and Prof. Dr. Munir Ozturk for assistance in obtaining germplasm for this research.

References

Ball, P.W. 1968. *Cercis* L. In: *Flora Europea* (Ed.): T.G. Tutin, V.H. Heywood, N.A. Burges, D.M. Moore, D.H. Valentine, S.M. Walters and D.A. Webb. Vol. 2: 83.

Bentham, G. 1840. Enumeration of plants collected by Mr. Schomburgk in British Guiana. *Hook. Journ. Boot*., Vol. 2: 38-99.

Chamberlain, D.F. and F. Yaltirik. 1970. *Cercis* L. In: *Flora of Turkey and the East Aegean Islands*, (Ed.): P.H. Davis. 3: 8-9, 13. Edinburgh at the University Press.

Chung-kuo Chih Wu Chih, L. 1988. Flora of China (Chinese version). Flora Reipublicae Popularis Sinicae. Tomus 39. Science Press, 1988 (Dicotyledoneae, Leguminosae (1) redacted by Chen Te-chao).

Corpet, Florence. 1988. Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res*., 16 (22): 10881-10890.

Coskun, F. 2003. The phylogeny and biogeographyof *Osmanthus*, *Cercis* and *Tilia* based on ITS nuclear ribosomal, *ndh*F and *trn*L-F chloroplast DNA Sequence Data. Unpublished Doctoral Dissertation, University of North Carolina, Chapel Hill, N.C., U.S.A.

W. Fritsch, J. Li and M.J. Donoghue. 2002. Phylogeny and biogeography of *Cercis*): Evidence from nuclear ribosomal ITS and chloroplast *ndh*F sequences. *Systematic* 7(2): 289-302.

Doyle, J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19: 11-15.

Doyle, J.J. 1995. DNA data and Legume Phylogeny: a progress report. In: *Advances in Legume Systematics*. (Eds.): M. Crisp and J.J. Doyle. Vol. 7: Phylogeny, 11-30. Royal Botanic Gardens, Kew.

Eriksson, T. and N. Wikström. 1995. AutoDecay, version 3.0.3. Stockholm University, Stockholm.

Farris, J.S. 1989. The retention index and the rescaled consistency index. *Cladistics*, 5: 417-419.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using bootstrap. *Evolution*, 39: 783-791.

Flora of China. 2002. The Flora of China Project. Available online: http://flora.huh.harvard.edu/china/mss/mssindex.htm .

Hillis, D.M. and J.P. Huelsenbeck. 1992. Signal, noise and reliability in molecular phylegenetic analysis. *J. Heredity*, 83: 189-195.

Isely, D. 1958. Leguminosae of the North-Central United States. III. Mimosoidea and Caesalpinioidea. *Iowa State Coll. Jour. Sci*., 32: 355-393.

Isely, D. 1973. Leguminosae of the United States: I. Subfamily Mimosoidea. *Memoirs of the New York Botanical Garden*, 25(1): 1-152.

Isely, D. 1975. Leguminosae of the United States: I. Subfamily Caesalpinioidea. *Mem. N. Y. Bot. Gard*., 25(2): 1-228.

Kluge, A.G. and J.S. Farris. 1969. Quantitative phyletics and the evolution of Anurans. *Systematic Zoology* (currently named as *Systematic Biology*, 18: 1-32.

Li, Hui-Lin. 1944. Taxonomy and distribution of the genus *Cercis* in China. *Bulletin of the Torrey Botanical Club*, 71(4): 419-425.

Polhill, R.M. and P.H. Raven and C.H. Stirton. 1981. Evolution and systematics of the Leguminosae. In: *Advances in Legume Systematics*. Part 2. Royal Botanic Gardens, Kew, England.

Swofford, D.L. 2001. PAUP*. *Phylogenetic Analysis Using Parsimony* (*and Other Methods). Version 4.0b8. Sinauer Associates, Sunderland, Massachusetts.

White, T.J., T. Bruns, S. Lee and J. Taylor. 1990. Amplifications and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR protocols: a guide to methods and applications*, (Eds.): M. Innis, D. Gelfand, J. Sninsky and T. White. 315-322. Academic Press, San Diego, California, USA.

Wunderlin, R.P., K.Larsen and S.S. Larsen. 1981. Tribe 3. CERCIDEAE Bronn (1822). In: *Advances in Legume Systematics*. Part 1. Royal Botanic Gardens, Kew, England.

(Received for publication 9 September 2007)