Phylogenetic position of *Dipentodon sinicus*: evidence from DNA sequences of chloroplast *rbc*L, nuclear ribosomal 18S, and mitochondria *mat*R genes

Yalin Peng¹, Zhiduan Chen², Xun Gong³, Yang Zhong⁴, and Suhua Shi^{1,*}

- ¹Key Laboratory of Gene Engineering of The Ministry of Education, Zhongshan University, Guangzhou 510275, The People's Republic of China
- ²Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100873, The People's Republic of China

³ Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, The People's Republic of China

⁴ Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, School of Life Sciences, Fudan University, Shanghai 200433, The People's Republic of China

(Received June 25, 2002; Accepted February 24, 2003)

Abstract. Phylogenetic position of the monotypic genus *Dipentodon* has long been controversial. We investigated its position with 125 accessions representing 50 genera and 40 families of eudicots in the APG system. Four data sets—including the chloroplast gene *rbc*L, the nuclear 18S ribosomal DNA, and the mitochondrial gene *mat*R, as well as the combined matrix—were used in the study with the maximum parsimony (MP) and Bayesian inference (BI) analyses. The phylogenetic trees based on individual genes and the combined data suggested that *Dipentodon* is sister to *Tapiscia* (Tapisciaceae) and that *Dipentodon* could be placed in euroside II of the APG system. The clade of *Dipentodon* and *Tapiscia* is closest to Malvales and Sapindales. Such finding does not support the previously suggested close relationship between *Dipentodon* and various other groups, including Celastraceae, Samydaceae, Flacourtiaceae, Hamamelidaceae, and Santalales.

Keywords: Bayesian inference; Chloroplast *rbc*L; *Dipentodon sinicus*; Maximum parsimony; Mitochondrial *mat*R; Molecular phylogeny; Nuclear ribosomal 18S.

Introduction

The genus Dipentodon Dunn consists of a single species D. sinicus Dunn, native to southern China and adjacent Burma and northeastern India (Merrill, 1941; Fischer, 1941; Li, 1986; Thorne, 1992; Bhattacharya and Johri, 1998). Its systematic position has been controversial since it was established and placed in the family Celastraceae in 1911 by S. T. Dunn. For example, Sprague (1925) moved it into the family Samydaceae based on the same variation range of floral base numbers in Dipentodon and Samydaceae. Many authors put the genus into the family Flacourtiaceae (including Samydaceae) (Fischer, 1941; Loesener, 1942; Metcalfe and Chalk, 1950; Lobreau, 1969). Record (1938) considered that Dipentodon is close to Hamamelidaceae based on the wood anatomic characters. Merrill (1941) proposed an independent family Dipentodontaceae Merr. and placed the family in Rosales between Hamamelidaceae and Rosaceae (Merrill, 1941; Hutchinson, 1959, 1973; Schultze-Motel, 1964; Dahlgren, 1980; Cronquist, 1981; Takhtajan, 1987, 1997). However, Cronquist (1981) put the Dipentodontaceae into the order Santalales based on similar characters of the gynoecial structure. In the update Angiosperm Phylogeny Group (APG) classification scheme of flowering plants, the phylogenetic position of Dipentodontaceae is still uncertain (Angiosperm Phylogeny Group, 1998; APG II, 2003).

Recently, DNA sequences of the chloroplast, nuclear ribosomal, and mitochondrial genes have been widely used in plant phylogenetics, especially in reconstructing angiosperm phylogeny (Hoot et al., 1995; Qiu et al., 1999; Soltis et al., 1998; Kuzoff and Gasser, 2000). More importantly, combining multiple genes from three genomes in plants has proved effective in reducing homoplasy generated by gene-, function-, and genome-specific molecular evolutionary phenomena (Qiu et al., 1999). However, phylogenetic information about *Dipentodon* is poorly known from molecular data. In the present paper, therefore, we conduct phylogenetic analyses of the *rbcL*, 18S, and *matR* sequences to determine the phylogenetic position of *Dipentodon* and its relationships with related groups.

^{*}Corresponding author. Tel: 86-20-84113677; Fax: 86-20-84113652; E-mail: lssssh@zsu.edu.cn

Materials and Methods

In-Group and Out-Groups

According to the APG classification scheme, 45 taxa for *rbcL*, 42 taxa for *matR*, and 31 taxa for 18S nrDNA, which represent the families Celastraceae, Samydaceae, Flacourtiaceae, and others from the core eudicots were selected as the in-group. *Platanus occidentalis* (Platanaceae) and *Akebia quinata* (Lardizabalanceae) were selected as the out-groups for the phylogenetic analyses. GenBank accessions and the taxa used in this study are listed in Table 1.

DNA Extraction and Sequencing

Total genomic DNA was extracted from fresh and silicagel-dried leaves using the CTAB procedure (Doyle and Doyle, 1987), and then purified with a DNA purification system (DPS) kit made by our laboratory. Aliquots of the total DNAs were used for sequencing all of the *matR* gene and part of *rbcL* and 18S rRNA genes. The PCR products of all samples were purified by using the QIAquick PCR Purification Kit (CN 28104, QIAGEN), and sequenced by using an ABI 377 Genetic Analyzer (Applied Biosystems, CA). All sequences have been deposited in GenBank (for accession numbers see Table 1).

Phylogenetic Analysis

The sequences used in this study were aligned with the Clustal-X program (Thompson et al., 1997) and modified manually. For phylogenetic analyses based on individual genes, the maximum parsimony (MP) method was used with PAUP* 4.0b5 (Swofford, 1999) (the heuristic search option with TBR branch-swapping and simple addition). Characters were assigned equal weights at all nucleotide positions (Fitch, 1971). Gaps were treated as missing data. Bootstrap analyses (Felsenstein, 1985) with 1000 replicates were performed to examine the relative level of support for individual clades on the phylogenetic trees. All phylogenetic trees were rooted using *Platanus occidentalis* and *Akebia quinata* as the out-groups.

A combined data set of the chloroplast, nuclear ribosomal, and mitochondrial sequences from 33 taxa was analyzed using the MP method implemented in the PAUP* 4.0b5 and the Baysian inference (BI) method implemented in the MrBayes 2.0 (Huelsenbeck and Ronquist, 2001; Huelsenbeck et al., 2001). Since there was no topological difference between trees of the two out-groups and with either one out-group, *P. occidentalis* was selected as the out-group for the combined MP and BI analyses. MrBayes uses a MCMC algorithm that runs four Markov chains simultaneously. The Markov chains were started from a



Figure 1. The most parsimonious tree (MPT) based on the combined data set. Numbers above branches represent the bootstrap values (%) for the clades with 1000 replicates. For tree parameters, see Table 2.

Table 1.	GenBank	accessions	and the ta	axa samp	led in	this study.	Classification	based c	on APG	(1998).
										< / >

Familiy	Service	GenBank accession No						
raininy	species -	matR	rbcL	185				
EUDICOTS								
Lardizabalaceae	Akebia quinata	AF197810	L12627	L31795				
Platanaceae	Platanus occidentalis	AF197793	L01943	U42794				
CORE EUDICOTS								
Hamamelidaceae	Hamamelis iaponica	*AF520087	L01922 (H mollis)	AF015654				
Iteaceae	Itea vunnanensis	*AF520099	AF190435 (I virginica)	U42545 (I virginica)				
Santalaceae	Osvris wightiana	*AF520155	$L11196(O \ lanceolata)$	$L_{24409}(O \ lanceolata)$				
POSIDS		111020100						
Geraniaceae	Garanium wilfordii	*AV121488	$I_{14608}(G_{10})$					
Stanbyleaceae	Stanbulga trifolia	*AF520105	A 1238406	A 1235078				
	Staphytea trijotta	AI 320103	AJ230400	AJ255778				
EUROSIDI		* 13/101500						
Begoniaceae	Begonia laciniata	*AY121500	U59814 (B. ulmifolia)	AF008950 (B. oxyloba)				
Celastraceae	Celastrus orbiculatus	*AY121493	$L_{12104} (E_{14})$	$\mathbf{V}_{1}(\mathbf{C}_{0})(\mathbf{F}_{1},\mathbf{L}_{1})$				
Celastraceae	Euonymus bungeanus	*AY121492	L13184 (E. alatus)	X16600 (E. alatus)				
Celastraceae	Maytenus hookeri		*AF499234					
Celastraceae	Iripterygium hypoglaucum	* 1 5520102	*AF499233					
Cephalotaceae	Cephalotus follicularis	*AF520193						
Cucurbitaceae	Luffa cylindrica	*AF5201/3	L21941 (L. quinquefida)	AF008957 (L. quinquefida)				
Flacourtiaceae	Flacourtia ramonicnii	*AF520186	*AF454/36 (F. montana)					
Flacourtiaceae	Homalium racemosa		AJ418822					
Flacourtiaceae	Poliothyrsis sinensis	* 1 5520072	AJ402991					
Juglandaceae	Juglans mandshurica	*AF5200/3	U00437 (J. nigra)	AF206943 (J. nigra)				
Oxalidaceae	Oxalis corymbosa	*AF520198	L01938 (O. dillenii)	AF206978 (<i>O. dillenii</i>)				
Passifloraceae	Passiflora edulis	*AF520188	L01940 (P. quadrangularis)	AF206981 (<i>P. standleyi</i>)				
Polygalaceae	Polygala tatarinowii	*AF520179	Z/01/6 (P. chamaebuxus)	U42/9/ (P. pauciflora)				
Rhamnaceae	Hovenia acerba	*AF520156	AJ39039 (H. dulcis)					
Rhoipteleaceae	Rhoiptelea chiliantha	*AF5200//	AF01/68/					
Rosaceae	Prinsepia uniflora	*AF520088	U06819 (<i>P. uniflora</i>)					
Salicaceae	Salix raddeana	*AF520191	AB012/98 (S. tetrasperma)	AF20/011 (S. reticulata)				
Ulmaceae	Cellis bungeana	*AF520086	L12638	1142810				
Umaceae	Zelkova serrala	*AF520089	D80317	042819				
EUROSID II								
Bixaceae	Bixa orellana	*AF520136	Y15139	AF206868				
Brassicaceae	Capparis membranifolia	*AF520146	M95754 (C. hastata)					
Bretschneideraceae	Bretschneidera sinensis	*AF520118	M95753					
Caricaceae	Carica papaya	*AF520141	M95671	U42514				
Combretaceae	Quisqualis indica	*AF520133	L01948	AF207004				
Lythraceae	Lythrum salicaria	*AF520110	L10218 (L. hyssopifolia)	AF206955				
Malvaceae	Bombax malabaricum	*AF520148	AF022118 (B. buonopozense)	U42507 (B. ceiba)				
Malvaceae	Tilia mandshurica	*AF520107	AF022127 (T. americana)	AF207042 (T. americana)				
Onagraceae	Epilobium hirsutum	*AF520108	L10217 (E. angustifolium)	AF206907 (E. angustifolium)				
Punicaceae	Punica granatum	*AY121502	L10223	U38311				
Sapındaceae	Acer mono	*AF520112	L01881 (A. saccharum)	U42494 (A. rubrum)				
Sapındaceae	Koelreuteria bipinnata	*AF520120	U39283 (K. paniculata)	AF206947 (K. sp.)				
Simaroubaceae	Ailanthus altissima	*AF520106	L12566	AF206842				
Tapisciaceae	Tapiscia sinensis	*AF520103	*AF447/50, AF206825 (2)	AF207034				
Tropaeolaceae	Tropaeolum majus	*AF520137	L14/06	L28750				
ASTERIDS								
Styraceae	Styrax grandiflorus	*AF520205	AF396160 (S. ramirezii)	U43296 (S. americana)				
EUASTERID I								
Acanthaceae	Blepharis hildebrantii	*AF520153	AF188127 (Elytraria crenata)					
Solanaceae	Withania somnifera	*AF520149	U08619 (Schizanthus pinnatus)	AF207016 (S. pinnatus)				
Symplocaceae	Symplocos setchuensis	*AF520111	Z80193 (S. ferruginea)	U43297 (S. paniculata)				
Theaceae	<i>Camellia</i> sp.	*AF520203	L12602 (C. japonica)	U42815 (C. japonica)				
FAMILIES OF LINC	ERTAIN POSITION							
Dinentodontaceae	Dipentodon sinicus	*AY121494	*AF375609	*AF375610				
Dipentodontacede	Dipeniouon sinicus	1111217/7	111 57 5007	111 5 / 5010				

*Represent the samples which were sequenced in this study.

random tree and run for 100,000 generations sampling every 50 generations for a total of 2,000 samples each run. The first 100 samples from each run were discarded as burn-in. The gamma distribution (Yang, 1994) and HKY model (Hasegawa et al., 1985) were used in the BI analysis.

Results and Discussion

After sequence alignment, four data sets were formed for phylogenetic analyses: 1) the chloroplast *rbc*L data set containing 48 taxa was 1,382 bp in length; 2) the nuclear ribosomal 18S data set containing 33 taxa was 1,734 bp; 3) the mitochondrial *mat*R data set containing 44 taxa was 1, 549 bp; and 4) the combined data set containing 33 taxa was 4,537 bp. The parameters of the most parsimonious trees (MPTs) obtained from the four data sets are presented in Table 2.

The strict consensus trees of the MPTs based on the individual and combined data sets showed a congruent topology. In the combined tree (Figure 1), *Dipentodon sinicus* is shown to be sister to *Tapiscia sinensis* (Tapisciaceae), with a relatively high bootstrap support value (95%).

The BI tree based on the combined data set is shown in Figure 2. The sister group relationship between *D*. *sinicus* and *T. sinensis* is strongly supported with a high

Table 2. Comparison of indices for the various trees analyzed using MP in this study, Consistency index (CI) excludes uninformative characters, RI = retention index, RC = rescaled consistency index.

Data set (OTUs)	No. variable characters (%)	No. informative characters (%)	No. of trees	Length of trees	CI	RI	RC	Bootstrap support (%)*
matR (44)	521 (33.6%)	230 (14.8%)	307	869	0.5566	0.6200	0.4495	<50
<i>rbc</i> L (48)	496 (34.6%)	323 (23.4%)	4	1707	0.3282	0.4813	0.1957	66
18S (33)	272 (15.6%)	146 (8.4%)	1	681	0.3720	0.4825	0.2400	<50
matR/rbcL/18S (32)	1104 (24.3%)	587 (12.6%)	2	2584	0.5329	0.4570	0.2436	76

*For the clade, Tapiscia+Dipentodon and the Sapindales, Malvaceae and Brassiclaes.



Figure 2. Phylogenetic tree determined by Bayesian Inference from the combined data set. Numbers above branches represent Posterior probabilities (PP). (HKY85 model: Ka=3.12).

posterior probability (PP = 0.98). The two species are then sister to the clade consisting of three orders: Malvales, Sapindales, and Brassicales with the same posterior probability support (PP = 0.98). The clade, which can be identified as the eurosid II in APG, is also shown to be sister to the clade consisting of the eurosid I clade and the genus *Staphylea*. Within the eurosid I clade, the sister group relationship between the families Flacourtiaceae and Celastraceae is well supported (PP = 0.97).

Based on the small number of genera sampled, our study suggests that the genus *Tapiscia* is most closely related to *Dipentodon*. The genus *Tapiscia* was established by Oliver (1890) and placed in the family Sapindaceae. However, it has been placed in the family Staphyleaceae by many authors (Diels, 1909; Bean, 1909; Schneider, 1912; Cronquist, 1981) and recognized as a distinct family by Takhtajan (1987). In the APG system, the family Tapisciaceae has not been assigned to any order but at the base of the "eurosid II" (APG, 1998). Simmons et al. (1998) indicated that the data from *rbcL* and ITS sequences suggested a different ordinal placement for *Huertea* and *Tapscia* from other members of Staphyleaceae.

Our study also shows that the clade *Dipentodon* and *Tapiscia* and that of Malvales and Sapindales form a sister group. This does not support a close relationship of *Dipentodon* to any of the families to which it traditionally been considered closely related. Tapisciaceae are also shown to have a distant relationship to Staphyleaceae. However, there are relatively low bootstrap values at some deeper nodes. In general, the results obtained from the three different genes mostly agree with the treatments about those families in APG system (APG, 1998).

Acknowledgments. This study is supported by grants from the National Science Funds for Distinguished Young Scholars (39825104), the Major Research Project of the Chinese Academy of Sciences (KSCXZ-SW-101A), National Natural Science Foundation of China (30070053, 30130030, 30170071), Natural Science Foundation of Guangdong Province (001223), National Ministry of Education Foundation for Key Member Teachers and for Post-doctor program (20010558013) and Qiu Shi Science & Technologies Foundation. We would like to thank Jason Clower of Harvard University for critical reading of the manuscript.

Literature Cited

- Angiosperm Phylogeny Group [APG]. 1998. An ordinal classification for the families of flowering plants. Ann. Missouri Bot. Gard. 85: 531-553.
- Angiosperm Phylogeny Group [APG]. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141: 399-436.
- Bhattacharya, B. and B.M. Johri. 1998. Flowering Plants: Taxonomy and Phylogeny. Berlin.
- Bean, W.J. 1909. Garden Notes on new trees and shrubs. Bull. Misc. Inform. Kew **1909:** 355-356.
- Cronquist, A. 1981. An Integrated System of Cassification of

Flowering Plants. Columbia University Press, New York, 681 pp.

- Dahlgren, R.M.T. 1980. A revised classification of the angiosperms. Bot. J. Linn. Soc. 80: 91-124.
- Diels, L. 1909. Staphyleaceae. Bot. Jahrb. Syst. 29: 448.
- Doyle, J.J. and J.S. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19:** 11-15.
- Dunn, A. 1911. *Dipentodon*, a new genus of uncertain systematic position. Bull. Misc. Inform. Kew **1911**: 310-313.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:** 783-791.
- Fischer, C.E.C. 1941. Contributions to the flora of Burma: XVIII. Kew Bull. **1940:** 282-294.
- Fitch, W.M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. **20**: 406-416.
- Hoot, S.B., S. Magallon, and P.R. Crane. 1999. Phylogeny of basal Eudicots based on three molecular Data sets: *atpB*, *rbcL*, and 18S nuclear ribosomal DNA sequences. Ann. Missouri Bot. Gard. **86:** 1-32.
- Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitchondrial DNA. J. Mol. Evol. **22:** 160-174.
- Huelsenbeck, J.P. and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics **17:** 754-755.
- Huelsenbeck, J.P., F. Ronquist, R. Nielsen, and J.P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294: 2310-2314.
- Hutchinson, J. 1959. The Families of Flowering Plants, vol. 1. Dicotyledons. 2nd edn. Oxford, 329 pp.
- Hutchinson, J. 1973. The Families of Flowering Plants Arranged According to a New System Based on Their Probable Phylogeny, 3rd edn. London, 415 pp.
- Kuzoff, R.K. and C.S. Gasser. 2000. Recent progress in angiosperm phylogeny reconstruction. Trends Plant. Sci. 5: 330-336.
- Li, Y.R. 1986. Celastraceae. in CY Wu eds. Flora Xizangica. vol 3. Beijing, pp. 118-138.
- Lobreau, D. 1969. Les limites de 1"ordre" des Celastrales d'apres le pollen. Pollen & Spores **11:** 499-555.
- Loesener, T. 1942. Celastraceae. In A. Engler and K. Prantl (eds.), Die Natürlichen Pflanzenfamilien 20b. Duncker & Humblot, Berlin, pp. 87-197.
- Merrill, E.D. 1941. Dipentodontaceae. Plants collected by Captain F Kindon-Ward on the Vernay-Cutting Expedition. 1938-39. Brittonia **4:** 69-73.
- Metcalfe, C.R. and L. Chalk. 1950. Anatomy of the Dicotyledons. Oxford at the Clarendon Press. 126, 395 pp.
- Oliver, D. 1890. *Tapiscia sinensis* in JD Hooker eds Hooker's Icones Plantarum.Vol. X., pt. 1928, 3rd.
- Qiu, Y.-L., J. Lee, F. Bernasconi-Quadroni, D.E. Soltis, P.S. Soltis, M. Zanis, E.A. Zimmer, Z. Chen, V. Savolainen, and M.W. Chase. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature **402**: 404-407.
- Record, S.J. 1938. The American woods of the orders Celastrales, Olacales, and Santalales. Trop. Woods. **53**: 11-38.
- Schneider, C.K. 1912. Tapiscia. Illustr. Handb. Laubholzk. 2:

1026, f. 607.

- Schultze-Motel, W. 1964. Santalales. Engler, Adolf: Syllabus der Pflanzenfamilien, mit besonderer Berücksichtigung der Nutzpflanzen nebst einer Ubersicht über die Florenreiche und die Florengebiete der Erde. 12. Berlin-Nikolassee.
- Simmons, S.L., J.L. Panero, and M.W. Chase. 1998. Molecular systematics of Staphyleaceae: a re-evaluation of its composition, ordinal placement, intra- and infrageneric relationships as inferred from *rbcL* and ITS sequence data. Abstract. Annual Meeting of the Botanical Society of America 2-6 August. Baltimore, MD, pp. 143.
- Soltis, D.E., P. S. Soltis, M.E. Mort, M.W. Chase, V. Savolainen, S.B. Hoot, and C.M. Morton. 1998. Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for angiosperms. Syst. Biol. 47: 32-42.

Sprague, T.A. 1925. Samydaceae. J. Bot. Brit. Forein. 63: 9-13.

- Swofford, D.L. 1999. PAUP* 4.0. Phylogenetic Analysis Using Parsimony (and Other Methods). Sinauer, Sunderland, Mass.
- Takhtajan, A. 1987. Tapisciaceae. Systema Magnoliofphytorum (in Russian): 171. 31 Mar-1 Jun. Leningrad, Nauka.
- Takhtajan, A. 1997. Diversity and Classification of Flowering Plants. New York, 204 pp.
- Thompson, J.D., T.J. Gibson, and F. Plewniak. 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. **25**: 4876-4882.
- Thorne, R.F. 1992. Classification and geography of the flowering plants. Bot. Rev. **58:** 225-348.
- Yang, Z. 1994. Estimating the pattern of nucleotide substitution. J. Mol. Evol. **39:** 105-111.

十齒花(Dipentodon sinicus)的系統位置—來自葉綠體 rbcL、 核核糖體 18S 和線粒體 matR 基因的 DNA 序列證據

彭雅林1 陳之端2 龔 洵3 鍾 揚4 施蘇華1

1廣州中山大學基因工程教育部重點實驗室 乳京中國科學院植物研究所系統與進化植物學開放研究實驗室 3昆明中國科學院昆明植物研究所 4上海復旦大學生物多樣性與生態工程教育部重點實驗室

十齒花科(Dipentodontaceae)在目前流行的有花植物分類系統(APG)中屬於位置未定的類群。本文 採用 PCR 和自動測序的方法獲得十齒花(Dipentodon sinicus)及若干相關類群的葉綠體 rbcL、核核糖體 18S 和線粒體 matR 基因的 DNA 序列。應用最大簡約法和貝葉斯推斷法,對自測序列和 GenBank 下載序 列組成的資料矩陣進行單基因和三基因聯合的分支分析及統計檢驗。結果表明:十齒花與銀鵲樹構成姐妹 群,二者構成的分支與錦葵目和無患子目關係較近,不支持前人認為與衛矛科、天料目科、大風子科、金 縷梅科以及檀香目等親緣關係較近的觀點。

關鍵詞:十齒花;分子系統發育;最大簡約法;貝葉斯推斷法;葉綠體 *rbc*L基因;核核糖體 18S基因; 線粒體 *mat*R 基因。