The phylogenetic relationships of Balanophoraceae and related Santalales inferred from floral B homeotic genes and nuclear 18S rDNA sequences

Huei-Jiun Su and Jer-Ming Hu

Institute of Ecology and Evolutionary Biology, National Taiwan University

Abstract

Balanophoraceae are obligate root parasites, which comprise 17 genera with 44 species in the tropics and subtropics. Plants of Balanophoraceae are composed of underground tuberous structures that attach to the host, and only emerge above ground during reproduction. Unlike typical flowers of the core eudicots that have differentiated sepals and petals, the flowers of Balanophoraceae are highly reduced and the remnant floral organs are sometimes difficult to interpret the ontological origins. Due to the extreme reduction of morphological features in Balanophoraceae, the phylogeny of Balanophoraceae has been controversial. Although recent molecular data shows Balanophoraceae might be close to Santalales, the exact position of Balanophoraceae remains in question. Previous studies also showed these parasitic plants have accelerated DNA substitution rates, which have caused problems in phylogenetic reconstruction. To clarify the affinities of Balanophoraceae and other Santalales members, we compared phylogenetic trees resulted from analyses of nuclear 18S rDNA and the floral B-class homeotic genes. A total of 23 B-class gene homologues were identified for Balanophoraceae and Santalales species. The results of various phylogenetic analyses confirm the basal position of Balanophoraceae in the Santalales. In addition, the substitution rates of B-class genes do not have drastic changes among Balanophoraceae, Santalales and other eudicots, compared with 18S rDNA. These results suggest that floral homeotic genes could be potential tools for reconstructing difficult phylogenies such as holoparasites.

Introduction

Balanophoraceae are achlorophyllous obligate root parasites which comprising of 17 genera and 44 species in the tropics and subtropics (Kawakita and Kato, 2002). The plants are composed of underground tuberous structures attaching to the host and only emerge above ground during the reproductive stage. Flowering shoots of Balanophoraceae arise from tubers endogenously (Kuijt, 1969) and are leafless or with scaly leaves spirally arranged on the lower portion of the inflorescences (Kuijt and Dong, 1990). Due to loss of organs such as roots, stems and leaves as well as its highly reduced flower morphologies, the placement of Balanophoraceae in angiosperms has been problematic.

To clarify the relationships of Balanophoraceae and other Santalales members, we compared phylogenetic trees resulting from analyses of nuclear 18S rDNA and the floral B homeotic genes from selected taxa. The aims of this study are to (1) identify the floral B class genes from selected Balanophoraceae and Santalales species, hopefully to provide additional evidence to resolve their relationships. (2) examine if the rate acceleration is a common feature in the genome of Balanophoraceae. (3) provide functional perspectives of B class genes in regards of floral evolution in Balanophoraceae.

Materials and Methods

Taxonomic Sampling

Fourteen species of Santalales representing six families were used in this study, which are *Loranthus kaoi* (J. M. Chao) H. S. Kiu*, L. delivayi* Tiegh. and *Taxillus pseudochinensis* (Yamamoto) Danser of Loranthaceae, *Champereia manillana* Merr. of Opiliaceae, *Thesium chinense* Turcz. and *Santalum album* L. of Santalaceae, *Viscum articulatum* Burm. and *V. alniformosanae* Hayata of Viscaceae, *Olax imbricata* Roxb. of Olacaceae, *Schoepfia jasminodora* Sieb. & Zucc. of Schoepfiaceae and *Balanophora laxiflora* Hemsl., *B. fungosa* J.R. & G. Forst., *B. wrightii* Makino, and *B. harlandii* Hook. f. of Balanophoraceae.

Gene Sequence Determination

The B class MADS-box gene coding sequences were determined based on reverse transcriptase PCR from mRNA. Total RNAs were extracted using Plant Concert Reagent (Invitrogen, Carlsbad, CA, USA) from flower buds or mature flowers. Single strand cDNA was synthesized and reverse transcribed using SuperScript III RNase H^- reverse transcriptase kit with manufacturer's instructions (Invitrogen, Carlsbad, CA). Different pairs of degenerate primers, including the mostly used by recent studies (Stellari, Jaramillo, and Kramer, 2004), have

been tried in the survey. The cDNA then used as a template for amplifying the floral organ identity gene homologues. Degenerate primers for initial amplification of B-class genes were used according to previous publication (Stellari, Jaramillo, and Kramer, 2004). The PCR products were then purified with the QIAquick PCR purification kit (QIAGEN, Hilden, Germany) and cloned into pGEM-T Easy Vector System (Promega Corporation, Madison, WI, USA). The ligation mixtures were transformed into *Escherichia coli* DH5α-competent cells following the manufacturer's instruction (RBC Bioscience Corporation, Taipei, Taiwan). Thirty to eighty clones were sequenced for each species on an ABI Prism 3700 sequencer at Academia Sinica.

Results and Discussion

Sequence data

Twenty-five new B-class gene homologues and thirteen18S nrDNA sequences were identified from thirteen selected Santalales species (Table 1). All B-class gene homologues from Santalales displayed the conserved amino acids of M, I, K, C domains as well as the typical motifs in the C-terminal region (Fig. 1) (Kramer, Dorit, and Irish, 1998). Among all sampled Santalales taxa, only one homologue of eu*AP3*-like gene was found from *B. laxiflora.* Sequence information of four data sets used in phylogenetic analyses is provided in Table 2.

PI and AP3 lineages in Santalales

The trees generated from analyses of *PI* and *AP3* lineages (Figs. 3 and 4) generally concord with relationships of the angiosperm phylogenies inferred from multiple chloroplast, mitochondrial and nuclear rRNA gene data sets (Soltis, Soltis, and Chase, 1999; Soltis et al., 2000). All of the analyses of *PI* and *TM6* lineages agreed with a monophyly of Santalales except for the *PI* tree in maximum parsimonious analysis. In addition, all of the trees also revealed that *Viscum* (Viscaceae), *Thesium* (Santalaceae), *Santalum* (Santalaceae), *Champereia* (Opiliaceae), *Loranthus* (Loranthaceae) and *Schoepfia* (Schoepfiaceae) composed of a well-supported monophyletic group which *Balanophora* (Balanophoraceae) and *Olax* (Olacaceae) were paraphyletic to them. Both of the two lineages suggested *Balanophora* and *Olax* are primitive branches in the order, however, *Balanophora* was resolved as the first branch of Santalales clade in the BI analysis of *PI* lineages (1.0 PP), while the ML and BI analysis of *TM6* lineages supported that *Olax* is the first branch of Santalales (0.9 BP and 1.0 PP).

Previous analyses using nuclear 18S rDNA, chloroplast *rbcL*, *atpB*, *matK*, and mitochondrial *matR* sequences all indicated a strong monophyly of Santalales, with or without the inclusion of Balanophoraceae taxa in the analyses (Nickrent, 1998; Nickrent, 2001;

Barkman, 2007 ; Malécot & Nickrent, 2008). However, it is not clear whether Balanophoraceae is a sister group to Santalales or a derived group within Santalales (Nickrent, Der, and Anderson, 2005). Our results on both 18S rDNA and *AP3* sequences support the close relationships for Balanophoraceae and Olacaceae, and the two are the first branches sistered to other Santalales.

Rate heterogeneity comparison between floral B homeotic genes and 18S nrDNA gene

Relative rate tests between angiosperms groups revealed the evolutionary rates of 18S nrDNA sequences are homogeneous in all selected angiosperm representatives except for Santalales (Table 3). All sequences of Santalales 18S nrDNA show a drastic increase in nucleotide substitution rates, compared to other angiosperms. The rate is even higher in *Balanophora* compared to other Santalales taxa. In comparison, no significant rate differences exist among the Santalales members and other core eudicots for both *PI* and *AP3* sequences. However, a rate difference was observed between core eudicots and the rest of angiosperms (marked as 'Noncore' in Table 3). In comparison, substitution rates of 18S nrDNA sequences between *Balanophora* and other Santalales evolve about 3 times faster than the rates of *PI* and *TM6* sequences.

Despite numerous advantages of using 18S nrDNA sequences in resolving angiosperm phylogenies, the extremely high nucleotide substitution rates in some lineages have possessed high risk in erroneous phylogenetic reconstruction due to the long branch effect (Felsenstein, 1978). Our results show that these B class gene homologues evolved in a relatively constant rate among angiosperms, even in holoparasitic lineages.

Reference

FELSENSTEIN, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* 27: 401-410.

KAWAKITA, A., AND M. KATO. 2002. Floral biology and unique pollination system of root holoparasites Balanophora kuroiwai, and B-tobiracola (Balanophoraceae). *American Journal of Botany* 89: 1164-1170.

KRAMER, E. M., R. L. DORIT, AND V. F. IRISH. 1998. Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages. *Genetics* 149: 765-783.

KUIJT, J. 1969. Rafflesiaceae, Hydnoraceae, and Balanophoraceae, The Biology of Parasitic Flowering Plants, 118-133. University of California Press.

KUIJT, J., AND W. X. DONG. 1990. Surface-Features of the Leaves of Balanophoraceae - a

Family without Stomata. *Plant Systematics and Evolution* 170: 29-35.

NICKRENT, D. L., J. P. DER, AND F. E. ANDERSON. 2005. Discovery of the photosynthetic relatives of the "Maltese mushroom" *Cynomorium*. *Bmc Evolutionary Biology* 5.

SOLTIS, D. E., P. S. SOLTIS, M. W. CHASE, M. E. MORT, D. C. ALBACH, M. ZANIS, V. SAVOLANINEN, W. H. HAHN, S. B. HOOT, M. F. FAY, M. AXTELL, S. M. SWENSEN, L. M. PRINCE, W. J. KRESS, K. C. NIXON, AND J. S. FARRIS. 2000. Angiosperm phylogeny inferred from 18S rRNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* 133: 381-461.

SOLTIS, P. S., D. E. SOLTIS, AND M. W. CHASE. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402: 402-404.

STELLARI, G. M., M. A. JARAMILLO, AND E. M. KRAMER. 2004. Evolution of the APETALA3 and PISTILLATA lineages of MADS-box-containing genes in the basal angiosperms. *Molecular Biology and Evolution* 21: 506-519.

Table 1. List of species sampled for floral B-class homeotic genes and 18S nrDNA sequences

		Floral B homeotic gene			Nuclear rDNA gene
Family	Species	*PI*	*TM6*	*AP3*	18S
Balanophoraceae	*Balanophora laxiflora*	*BalPI*	*BalTM6*	*BalAP3*	*Bal18S*
	Balanophora fungosa	*BafPI*	*BafTM6*	-	*Baf18S*
	Balanophora harlandii				*Bah18S*
	Balanophora wrightii				*Baw18S*
Olacaceae	*Olax imbricata*	*OliPI*	*OliTM6*	-	*Oli18S*
Santalaceae	*Santalum album*	*SaaPI*	*SaaTM6*	-	*Saa18S*
	Thesium chinensis	*ThcPI*	*ThcTM6*	-	*Thc18S*
	Viscum articulatum	*ViarPI-1*	*ViarTM6-1*	-	*Viar18S*
		ViarPI-2	*ViarTM6-2*		
	Viscum alniformosanae	*VialPI*	*VialTM6*	-	*Vial18S*
Loranthaceae	*Loranthus kaoi*	*LokPI*	*LokTM6*	-	*Lok18S*
	Loranthus delivayi	*LodPI*	*LodTM6*		
	Taxillus pseudochinensis	-	*TaxTM6*	-	*Tap18S*
Opiliaceae	*Champereia manillana*	*ChmPI*	*ChmTM6*	-	*Chm18S*
Schoepfiaceae	*Schoepfia jasminodora*	*ScjPI*	*ScjTM6*	-	*Scj18S*

Table 2. Comparison of sequence variation and tree statistics in four data sets.

	18S -Santalales	18S- Angiosperms	*PI* -Angiosperms	*AP3*-Angiosperms
Sequence number	62	44	57	86
Outgroup number	3	1	3	5
Aligned length (bp)	1787	1676	774	888
GC content (%)	47.8	49.4	45.1	44.1
Number of variable characters (%)	182 (10%)	192 (11%)	105 (16%)	102 (11%)
Number of PI characters	446 (25%)	227 (14%)	507 (66%)	592 (67%)
Number of MPTs	235	358	115	36
Tree length	2267	1315	5087	8141
C.I. (consistency index)	0.398	0.426	0.254	0.198
R.I. (retention index)	0.604	0.367	0.431	0.494

Note. PI = parsimony informative; MPT = most parsimonious tree.