Annals of Botany **110**: 1137–1147, 2012
doi:10.1093/aob/mcs197, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

Rate heterogeneity in six protein-coding genes from the holoparasite *Balanophora* (Balanophoraceae) and other taxa of Santalales

Huei-Jiun Su and Jer-Ming Hu*

Institute of Ecology and Evolutionary Biology, National Taiwan University, Taiwan
** For correspondence. E-mail jmhu@ntu.edu.tw*

Received: 7 February 2012 Returned for revision: 28 February 2012 Accepted: 23 July 2012 Published electronically: 21 September 2012

• *Background and Aims* The holoparasitic flowering plant *Balanophora* displays extreme floral reduction and was previously found to have enormous rate acceleration in the nuclear 18S rDNA region. So far, it remains unclear whether non-ribosomal, protein-coding genes of *Balanophora* also evolve in an accelerated fashion and whether the genes with high substitution rates retain their functionality. To tackle these issues, six different genes were sequenced from two *Balanophora* species and their rate variation and expression patterns were examined.
• *Methods* Sequences including nuclear *PI*, eu*AP3*, *TM6*, *LFY* and *RPB2* and mitochondrial *matR* were determined from two *Balanophora* spp. and compared with selected hemiparasitic species of Santalales and autotrophic core eudicots. Gene expression was detected for the six protein-coding genes and the expression patterns of the three B-class genes (*PI*, *AP3* and *TM6*) were further examined across different organs of *B. laxiflora* using RT-PCR.
• *Key Results* Balanophora mitochondrial *matR* is highly accelerated in both nonsynonymous (d_N) and synonymous (d_S) substitution rates, whereas the rate variation of nuclear genes *LFY*, *PI*, eu*AP3*, *TM6* and *RPB2* are less dramatic. Significant d_S increases were detected in *Balanophora PI*, *TM6, RPB2* and d_N accelerations in eu*AP3*. All of the protein-coding genes are expressed in inflorescences, indicative of their functionality. *PI* is restrictively expressed in tepals, synandria and floral bracts, whereas *AP3* and *TM6* are widely expressed in both male and female inflorescences.
• *Conclusions* Despite the observation that rates of sequence evolution are generally higher in *Balanophora* than in hemiparasitic species of Santalales and autotrophic core eudicots, the five nuclear protein-coding genes are functional and are evolving at a much slower rate than 18S rDNA. The mechanism or mechanisms responsible for rapid sequence evolution and concomitant rate acceleration for 18S rDNA and *matR* are currently not well understood and require further study in *Balanophora* and other holoparasites.

Key words: *Balanophora*, Balanophoraceae, B-class genes, *LFY*, *RPB2*, mitochondrial *matR*, substitution rate, phylogeny, Santalales, parasitic plants..

INTRODUCTION

Rates of nucleotide substitution are known to vary among species and genes, and the factors affecting such variation include population size, life history, metabolic rate and DNA repair efficiency (Britten, 1986; Nickrent and Starr, 1994; Gaut *et al.*, 1996; Laroche *et al.*, 1997; Ohta, 2000; Smith and Donoghue, 2008). Several studies have demonstrated that the substitution rates in nuclear small-subunit ribosomal DNA (18S rDNA) for many holoparasitic and mycoheterotrophic plants are usually high compared with other angiosperms (Nickrent and Starr, 1994; Lemaire *et al.*, 2010). Such rate accelerations are found in the other two intracellular genomes of many holoparasites, including plastid 16S rDNA (Nickrent *et al.*, 1997) and *rps2* (dePamphilis *et al.*, 1997) and mitochondrial small-subunit rDNA (Duff and Nickrent, 1997) and *matR* (Barkman *et al.*, 2004). Among the cases with accelerated substitution rates, the holoparasitic plant *Balanophora* has one of the highest substitution rates in 18S rDNA among angiosperms (Nickrent and Starr, 1994; Nickrent *et al.*, 2005).

Balanophora (Balanophoraceae) is a root parasite with unisexual flowers, the plants being either dioecious or monoecious. The floral organs of *Balanophora* are highly modified and reduced (Fig. 1A–D). The male flower (Fig. 1D) of *Balanophora* is composed of a single tepal whorl surrounding the synandrium formed by fused anthers. The female flower (Fig. 1C) completely lacks a perianth and consists only of a single pistil, and the female flowers (pistils) are clustered around numerous club-shaped structures called spadicles (Hansen, 1972) or claviform bodies (Eberwein *et al.*, 2009).

Based on phylogenetic studies using nuclear 18S rDNA and mitochondrial *matR* regions, the family Balanophoraceae is allied with Santalales (Nickrent *et al.*, 2005; Barkman *et al.*, 2007) and is included in that order in the Angiosperm Phylogeny Group classification (APG III, 2009), but its exact position remains unresolved. Considering that the extreme rate acceleration of 18S rDNA in *Balanophora* could impose problems for the inference of phylogenetic relationships (see Nickrent *et al.*, 2005), additional gene markers with less rate heterogeneity would be useful in further clarifying phylogenetic relationships of Balanophoraceae. The MADS-box genes

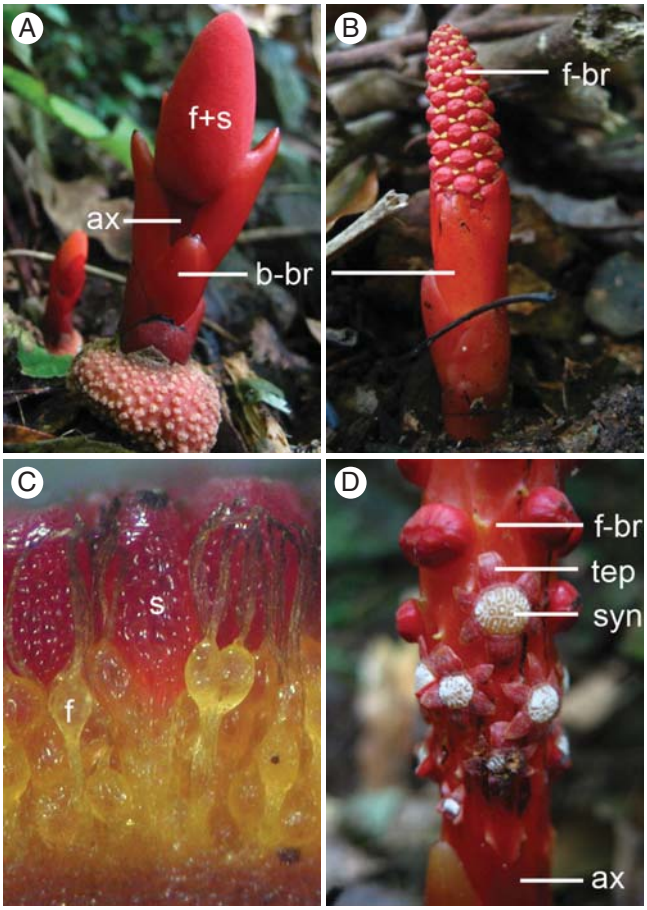

FIG. 1. Photographs of *Balanophora laxiflora*: (A, C) female flowers and inflorescences; (B, D) male inflorescence and flowers. Abbreviations: f, female flower; s, spadicle; ax, inflorescence axis; b-br, basal bract; f-br, floral bract; tep, tepal; syn; synandrium.

that have conserved functions for floral homoeotic patterning in angiosperms are good candidates for consideration.

The floral homoeotic genes for floral patterning under the current ABCDE model have been extensively studied in the two model species, *Arabidopsis thaliana* and *Antirrhinum majus* (Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994; Theissen and Saedler, 2001). According to the model, five major classes of floral homoeotic genes work in different combinations to specify the identities of sepals, petals, stamens and carpels. Among these, B-class genes are known to be responsible for specifying the identities of petals and stamens. Two major duplication events occurred during the evolution of B-class genes in seed plants (Kramer *et al.*, 1998). The first duplication event was before the origin of the angiosperms and this gave rise to the *PISTILLATA/GLOBOSA* (*PI/GLO*) and *APELATA3/DEFICIENS* (*AP3/DEF*) gene lineages (Kramer *et al.*, 1998; Kramer and Irish, 2000; Theissen *et al.*, 2000). The second major duplication event was in the *AP3* lineage after the divergence of Trochodendraceae in extant core eudicots, giving rise to the eu*AP3* and *TM6* gene sub-lineages (Kramer *et al.*, 2006).

A gene duplication event also occurred in the evolution of RNA polymerase II (*RPB2*) gene which is responsible for transcription. A major duplication occurred shortly before the core eudicot divergence that leads to the *RPB2*-i and *RPB2*-d gene lineages (Oxelman *et al.*, 2004; Luo *et al.*, 2007). Because the i copies have been lost several times during the diversification of core eudicots, some lineages may only retain the d copies such as *Arceuthobium* (Santalaceae *s.l.* or Viscaceae) of Santalales and *Liquidambar* (Altingiaceae) of Saxifragales (Luo *et al.*, 2007).

Compared with plastid and mitochondrial genes, fewer studies exist that have analysed evolutionary rates of nuclear protein-coding genes in parasitic plants. Extreme rate acceleration of 18S rDNA has been found in many heterotrophic taxa including *Rafflesia keithii* (Rafflesiaceae) and *Balanophora* spp. (Balanophoraceae), but not all holoparasitic plants, however, display such rate accelerations (Nickrent and Starr, 1994; Nickrent *et al.*, 1998; Young and dePamphilis, 2005; Lemaire *et al.*, 2010). One hypothesis to be tested for the holoparasitic *Balanophora* is to see if it is a general feature that all genes in the genome evolve more quickly than those from hemiparasitic relatives and autotrophic core eudicots.

The highly reduced and modified floral morphology of *Balanophora* provides an opportunity to examine the conservation and divergence of B-class genes. The B-class genes are expected to show a certain degree of conservation due to the functional constraint. Similarly for the *LEAFY* homologues, a single-copy gene is responsible for the initiation and patterning of flowers (Blázquez *et al.*, 1997; Moyroud *et al.*, 2009) and *RPB2* homologues from *Balanophora* and other Santalales species. Comparison of the substitution rates of the B-class genes and other nuclear markers could help elucidate whether the rate acceleration is a common feature in the nuclear genome of *Balanophora*, compared with other santalalean relatives and non-parasitic core eudicots. Moreover, we examined the expression patterns of floral B-class genes in *B. laxiflora* to determine whether there are underlying genetic patterns that correlate with the unusual floral phenotypes.

MATERIALS AND METHODS

Taxonomic sampling

Eleven members of Santalales representing eight (or six if Santalaceae *s.l.* is used) families were used in this study: *Loranthus kaoi* and *L. delavayi* (Loranthaceae), *Champereia manillana* (Opiliaceae), *Thesium chinense* (Santalaceae *s.l.* or Thesiaceae), *Santalum album* (Santalaceae *s.s.*), *Viscum articulatum* and *V. alniformosanae* (Santalaceae *s.l.* or Viscaceae), *Olax imbricata* Roxb. (Olacaceae), *Schoepfia jasminodora* (Schoepfiaceae) and *Balanophora laxiflora* and *B. fungosa* (Balanophoraceae). All plant materials were collected in Taiwan.

Gene sequence determination

The 18S rDNA and mt *matR* gene sequences were amplified by direct PCR of genomic DNA from the above taxa. Total genomic DNAs were extracted according to standard CTAB extraction methods (Doyle and Doyle, 1987) from fresh leaves or flowers. The PCR conditions and primers for

amplification and sequencing of 18S rDNA followed Nickrent (1994) and the primers used for the *matR* region are listed in Supplementary Data Table S1.

For *LFY*, *RPB2* and B-class MADS-box genes (*PI*, eu*AP3*, *TM6*), coding sequences were determined by analysis of reverse transcriptase PCR (RT-PCR) products generated from mRNA. Total RNA was prepared from young inflorescences or flowers using either Concert Plant reagent (Invitrogen, Carlsbad, CA, USA) or a method developed for *Pinus* (Chang *et al.*, 1993). The cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen). Degenerate primers for initial amplification of B-class, *LFY* and *RPB2* gene homologues were based on previous studies (Frohlich and Meyerowitz, 1997; Oxelman *et al.*, 2004; Stellari *et al.*, 2004). PCR reactions were performed with 1 μL of Advantage 2 polymerase mix (Clontech, Palo Alto, CA, USA), 5 μL 10× Advantage 2 PCR buffer, 200 μM of each dNTP, 200 μM of each primer and 1 μL cDNA template in a final volume of 50 μL for the reaction solution. The PCR amplification programme was as follows: 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 s, 55 °C for 45 s, 72 °C for 50 s and a final extension at 72 °C for 3 min. The PCR products were then cloned into the pGEM-T Easy Vector system (Promega, Madison, WI, USA) and 30–50 clones were sequenced for each taxon. Complete cDNA sequences of B-class gene homologues of *B. laxiflora* were obtained using 5'RACE kit (Invitrogen) and new gene specific primers were then designed for RT-PCR analysis; these primers are listed in Supplementary Data Table S1.

To determine the gene structure of *Balanophora* B-class genes (*BalAP3*, *BalTM6*, *BafAP3* and *BafTM6*), PCR was performed with genomic DNA of *B. laxiflora* and *B. fungosa*. Gene specific primers were designed (see Supplementary Data Table S1) and the PCR amplifications were carried out using Blend Tag (Toyobo, Osaka, Japan) following the manufacturer's instructions. Intron position was determined by aligning the genomic fragments with the corresponding cDNA sequences.

Phylogenetic analyses

Phylogenetic analyses were conducted with Bayesian inference (BI) using MRBAYES 3·04b (Huelsenbeck and Ronquist, 2001) and with maximum likelihood using Genetic Algorithm for Rapid Likelihood Inference (GARLI) version 0·951 (Zwickl, 2006). The model of DNA substitution for later incorporation was estimated using Modeltest 3·06 (Posada and Buckley, 2004), and the GTR + I + G model was selected as the fittest for the 18S rDNA, *PI*, *AP3/TM6* and *RPB2* datasets according to the Akaike information criterion. For the mt *matR* dataset, the fittest TIM + G model suggested by Modeltest is not implemented in MRBAYES, and therefore we chose the closely related GTR + G model instead. Sequence(s) of *Akebia* (Lardizabalaceae) was chosen as outgroup in all analyses [following Shan *et al.* (2006) and Jaramillo and Kramer (2007)]. Trees were sampled every 500 generations from 2·5 million generations of Markov chain Monte Carlo searches in BI analysis. The first 500 trees were discarded and the remaining trees were used to calculate posterior probabilities. Maximum likelihood bootstrap percentages (MLBP) (Felsenstein, 1985) were evaluated from 200 iterations using GARLI version 0·951 (Zwickl, 2006).

Branch length estimates and rate heterogeneity test

Branch lengths representing the number of non-synonymous (d_N) and synonymous (d_S) substitutions per site were estimated for protein-coding genes in our study using HyPhy package (Pond *et al.*, 2005). The Muse-Gaut (MG94) model of codon substitution was adopted to estimate d_N and d_S rates independently in each lineage. Likelihood ratio tests (Yang, 1998) were implemented to determine the significance of rate heterogeneity for d_N and d_S of protein-coding genes among holoparasitic *Balanophora*, hemiparasitic Santalales and autotrophic core eudicots. Both *B. laxiflora* and *B. fungosa* sequences were included to represent '*Balanophora*' in the analysis. Each of the tests contained two hypotheses for the categories compared: (1) that they share the same rate parameters (null hypothesis) or (2) that they have different parameters (alternate hypothesis). If the likelihood log estimate of alterative hypothesis is significantly higher than that of the null hypothesis ($P < 0·05$), it is assumed that the two categories have significantly different evolutionary rates. Rate correlations between d_N and d_S were analysed using Spearman's rank correlation test under the statistical program R (R Development Core Team, 2008).

Gene expression analysis

Different fresh floral tissues of *B. laxiflora* were dissected and frozen in liquid nitrogen for gene expression assays. Tepals, synandria and pollen were collected from open male flowers, and floral bracts were collected from floral buds and mature flowers. The bracts that occur basally on the inflorescence and parts of the inflorescence axis were collected from young male and female inflorescences. Female flowers and spadicles were prepared in a mixture since they are difficult to separate. Total RNA was extracted from the dissected tissues separately and was reverse transcribed as described above. Gene-specific primers of *B. laxiflora* used in this study are listed in Supplementary Data Table S1. Fifty nanograms of RNA was used in each PCR reaction and the amplification was performed with 94 °C for 2 min, followed by 34 cycles of 94 °C for 30 s, 55 °C for 40 s and 72 °C for 50 s and a final extension at 72 °C for 3 min using Platinum Taq DNA Polymerase (Invitrogen). The amplified fragments were further confirmed by direct sequencing.

Scanning electron microscopy

Male inflorescences of *B. laxiflora* were cut into small pieces and preserved in 70 % ethanol and then dehydrated in an ethanol series to 100 % ethanol. The samples were then critical-point dried, mounted onto pin stubs, coated with gold and examined using a Hitachi S-520 scanning electron microscope.

RESULTS

Sequence data

In this study we identified 25 B-class, two *LFY*, nine 18S rDNA, nine *RPB2* (d-copy; Supplementary Data Fig. S1) and nine mitochondrial *matR* gene homologues from two *Balanophora* and nine hemiparasitic species (Supplementary Data Table S2). All these taxa have one *PI*-like gene and one *TM6*-like gene except for *V. articulatum*, which has two *PI* homologues and two *TM6* homologues. However, only two eu*AP3* genes were obtained in our screening for Balanophoraceae and other Santalales, i.e. *BalAP3* from *B. laxiflora* and *BafAP3* from *B. fungosa*. All the B-class gene homologues obtained display the conserved amino acid domains M, I, K and C and the typical motifs of the C-terminal region for *PI*/*AP3* subgroups (Kramer *et al.*, 1998; see Supplementary Data Fig. S2).

Genomic DNA sequences of *BalAP3* and *BafAP3* were determined by direct PCR of the genomic DNA using several pairs of primers (see Supplementary Data Table S1) amplifying partial genomic sequences. All exon–intron structure and sequences can be identified, except for intron 4 sequences in all four B-class genes of *Balanophora*, because we could not obtain a PCR product using flanking primers of intron 4. After comparing cDNA sequences with other known eu*AP3* homologues, an additional 39-bp new exon was found at the C-domain region (which we denote as 6a) between the ordinary exons 6 and 7 of B-class genes in both *B. laxiflora* and *B. fungosa* (Supplementary Data Fig. S3). An approx. 0·5-kb region between exons 2 and 3 of *LFY* and an approx. 1·7-kb region between exons 12 and 24 of *RPB2* genes were amplified. The two *LFY* homologues from *Balanophora* and the nine *RPB2* homologues from *Balanophora* and other Santalales are quite conserved and can be easily aligned to other eudicot homologues. All of the sequences obtained and gene accessions used in the phylogenetic analyses are provided in Supplementary Data Table S2.

Molecular phylogenetics

The BI phylogenetic trees resulting from analysis of 18S rDNA, mt *matR*, *PI*, *AP3* and *RPB2* sequences are shown in Fig. 2. These are generally congruent with the relationships of angiosperm phylogenies inferred from previously published data based on multiple plastid, mitochondrial and nuclear rDNA gene datasets (Soltis *et al.*, 1999; Soltis *et al.*, 2000). In general, the *PI* and *AP3* trees received higher nodal support compared with the 18S rDNA and mt *matR* trees. All of the trees consistently showed that Santalales including Balanophoraceae formed a monophyletic group, which received high BI support (1·0 on the *PI*, *AP3* and *RPB2* tree) and low to moderate maximum likelihood support (MLBP = 66 % on *RPB2*, 55 % on *PI*, 87 % on the *AP3* tree) in *RPB2* and B-class gene trees (Fig. 2C–E), and still lower for 18S rDNA (BI = 0·87, MLBP < 50 %) and mt *matR* (BI = 0·89, MLBP < 50 %) gene trees (Fig. 2A, B). In contrast, the relationship of *Balanophora* to other members of Santalales is not resolved in any of the five gene trees. With 18S rDNA, *matR* and *RPB2*, *Balanophora* is related to *Olax*, *Schoepfia* and/or *Loranthus* (Fig. 2A–C). With *PI* (Fig. 2D), *Balanophora* is sister to *Olax* and this clade is sister to the remaining Santalales, whereas with *AP3*/*TM6* it is sister to all other Santalales except *Olax*. The relationships of the other santalalean taxa also differed among the four gene trees. Some of the nodes received relatively high support (BI > 0·95, MLBP > 90 %, marked as black dots on Fig. 2), but none of those nodes in Santalales showed incongruent phylogenetic relationships among the four gene trees.

Rates of molecular evolution

Extensive rate variation was detected among the six *Balanophora* gene sequences (18S rDNA, *matR*, *LFY*, *PI*, *AP3* and *RPB2*) (Fig. 3). *Balanophora* 18S rDNA showed an extraordinarily high substitution rate compared with hemiparasitic Santalales and non-parasitic species. Long branches for *Balanophora* sequences were also detected on mitochondrial *matR* and some of the other gene trees constructed from sites with synonymous substitutions. Therefore we performed likelihood ratio tests to examine rate heterogeneity in non-synonymous (d_N) and synonymous (d_S) substitutions for the three B-class genes (*PI*, eu*AP3*, *TM6*), as well as *LFY*, *RPB2* and mt *matR*, among holoparasitic *Balanophora*, hemiparasitic Santalales and autotrophic core eudicots.

The *matR*, *PI*, *TM6* and *RPB2* sequences of other Santalales showed significantly lower rates of d_N and d_S in all the comparisons to *Balanophora* or other core eudicots (Table 1). In contrast, *Balanophora* sequences generally showed an increase in d_N or d_S; however, only *Balanophora matR* showed significant rate accelerations (>10-fold). On the other hand, the B-class genes and *LFY* of *Balanophora* showed varying substitution rates compared with the core eudicots. Among these genes, significant rate accelerations were only found in d_S of *Balanophora PI, TM6* and *RPB2* genes, and d_N of *Balanophora* eu*AP3*. In other comparisons, *Balanophora* B-class gene sequences have only slight increases in substitution rates. It is worth noting that the d_N and d_S values of *Balanophora* B-class and *RPB2* genes are all less than two times higher than the core eudicot sequences, far less than seen with mt *matR*. Furthermore, *Balanophora LFY* sequences not only did not show a significant rate increase in d_N, they actually showed significantly lower d_S values compared with the core eudicots (0·290 vs. 0·732, Table 1). The analyses showed that d_N and d_S are significantly correlated within eu*AP3*, *TM6*, *LFY, RPB2* and *matR* among the groups, but the rates are poorly correlated within *PI* sequences.

Expression in the six protein-coding genes of Balanophora laxiflora

To find out whether the six protein-coding genes identified in *B. laxiflora* are actually expressed, RT-PCR was carried out using gene-specific primers on RNA samples. The results showed that all six protein-coding gene products were present in inflorescences (Fig. 4). The three B-class homoeotic genes displayed differential expression patterns in different parts including both reproductive and vegetative organs (Fig. 4B). *BalAP3* is expressed at high levels in male and female inflorescences, and the expression can be detected in all parts examined. Similarly, *BalTM6* is also expressed in

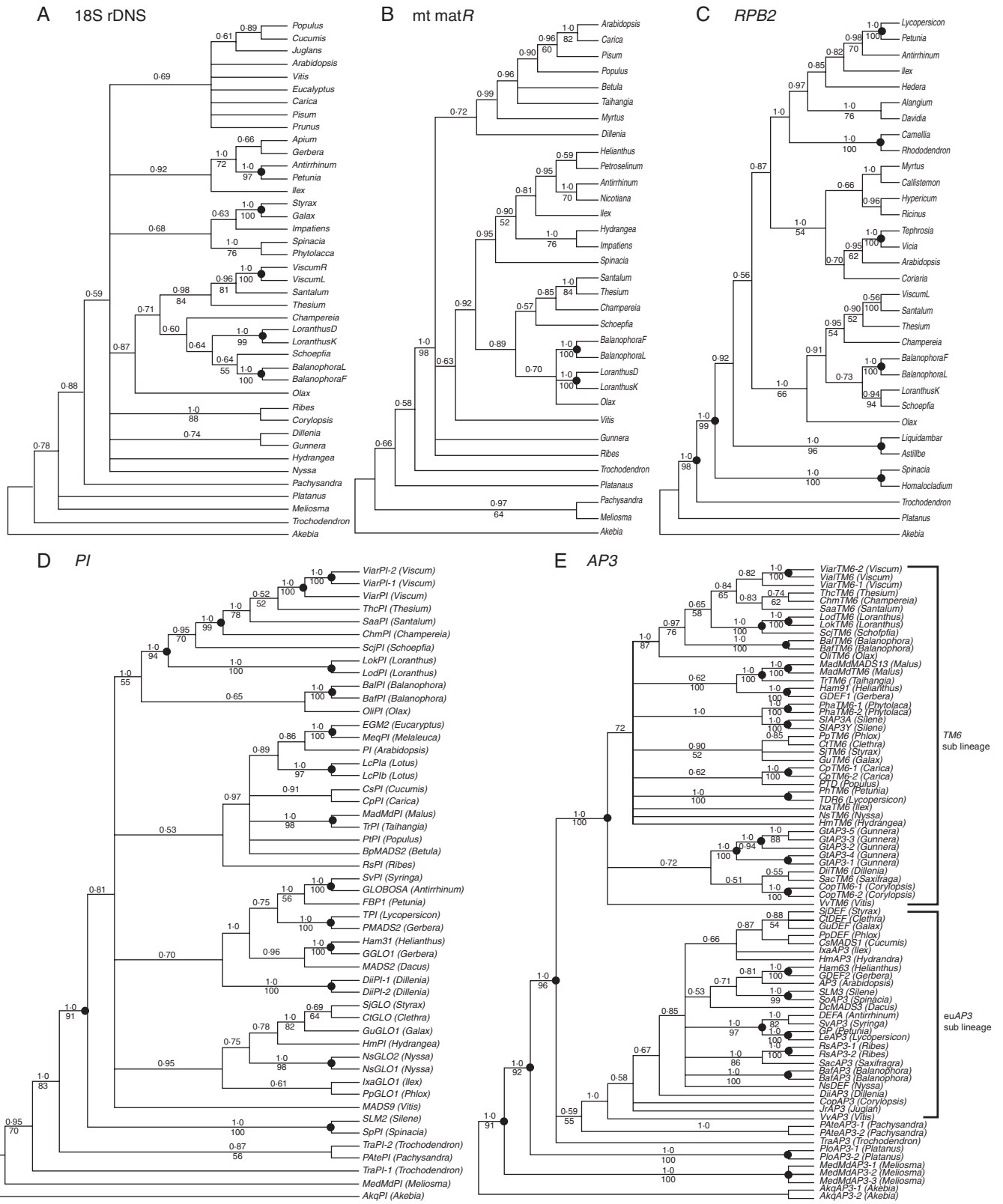

FIG. 2. Phylogenetic relationships of Santalales and *Balanophora* spp. inferred from Bayesian analysis (BI) of 18S rDNA, *matR*, *RPB2*, *PI* and *AP3* gene sequences. Numbers above the branches are Bayesian posterior probabilities, and values below are maximum likelihood branch support determined by GARLI. Mitochondrial *matR* trees are based on a 1968 nucleotide alignment, and nuclear gene trees are based on 1739 (18S rDNA), 765 (*PI*), 864 (*AP3*), 549 (*LFY*), 1719 (*RPB2*) nucleotide alignments, respectively. See Supplementary Data Fig. S1 for *RPB2* gene tree showing the obtained *RPB2* sequences belonging to the d clade.

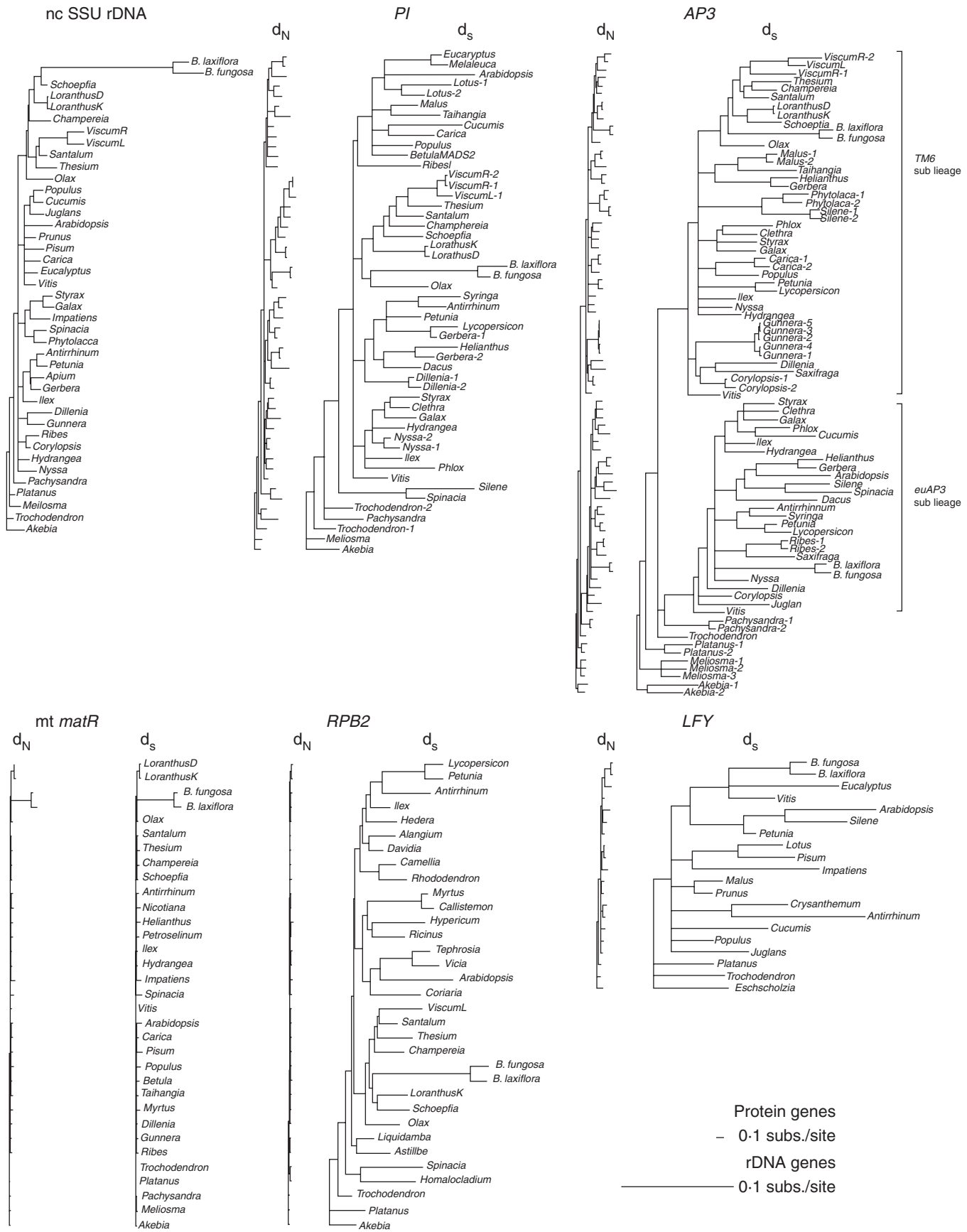

FIG. 3. Rate divergence of nuclear and mitochondrial genes in *Balanophora*. BI trees based on d_N and d_S for protein-coding genes or rDNA gene. All protein-coding gene trees are constrained to the same scale, whereas the rDNA tree is on a different scale (see notes in figure).

TABLE 1 *The estimated d_N, d_S values and* P*-values of the likelihood ratio tests between taxa*

	mt *matR*	*PI*	*TM6*	*AP3*	*LFY*	*RPB2*
d_N values						
Core eudicots	0·015	0·078	0·072	0·083	0·037	0·009
Balanophora	0·134	0·091	0·088	0·118	0·039	0·010
Other Santalales*	0·011	0·059	0·056	–	–	0·005
P-values of d_N comparison						
Core eudicots vs. other Santalales*	<0·01	<0·001	<0·001	–	–	<0·001
Balanophora vs. other Santalales*	<0·001	<0·001	<0·001	–	–	<0·001
Balanophora vs. core eudicots	<0·001	0·14 (n.s.)	0·05 (n.s.)	<0·001	0·85 (n.s.)	0·61 (n.s.)
d_S values						
Core eudicots	0·021	0·378	0·315	0·369	0·732	0·293
Balanophora	0·220	0·563	0·560	0·450	0·290	0·572
Other Santalales*	0·013	0·237	0·308	–	–	0·266
P-values of d_S comparison						
Core eudicots vs. other Santalales*	<0·01	<0·001	<0·001	–	–	<0·05
Balanophora vs. other Santalales*	<0·001	<0·001	<0·001	–	–	<0·001
Balanophora vs. core eudicots	<0·001	<0·001	<0·001	0·10 (n.s.)	<0·001	<0·001
d_N/d_S						
Core eudicots	0·714	0·206	0·228	0·225	0·050	0·030
Balanophora	0·609	0·161	0·157	0·262	0·135	0·017
Other Santalales*	0·846	0·248	0·181	–	–	0·018
Correlation coefficient of d_N and d_S	0·881	−0·272	0·739	0·653	0·614	0·672
P-values of d_N and d_S correlation	<0·01	0·06 (n.s.)	<0·01	<0·01	<0·01	<0·01

n.s., not significant; –, data not available.
*Santalales here includes hemiparasitic taxa and excludes Balanophoraceae.

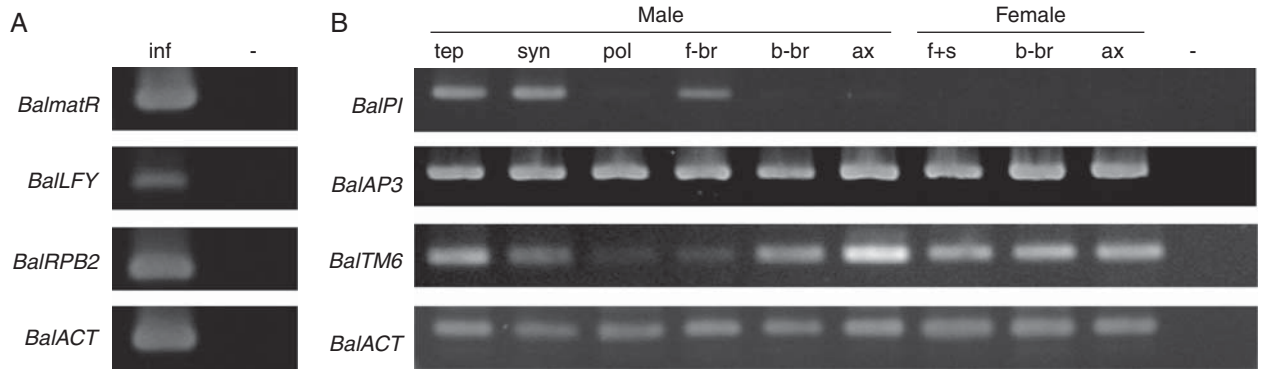

FIG. 4. Expression of the six protein-coding genes in *Balanophora laxiflora*: (A) expression of mitochondrial *matR*, *RPB2* and *LFY* from male inflorescences; (B) expression of floral B-class genes from different parts of male and female inflorescences. Actin primers were used as an internal control. Abbreviations: inf, inflorescence; tep, tepal; syn, synandrium; pol, pollen; f-br, floral bract; b-br, basal bract; ax, inflorescence axis; f + s, female flowers and spadicles; –, negative control.

male and female inflorescences, but the expression is low in the pollen and floral bracts compared with *BalAP3*. In contrast, *BalPI* is not expressed in the female inflorescence, and it is mainly restricted to the tepals, synandria and the floral bracts of male inflorescences.

Surface morphology of bracts and tepals in Balanophora laxiflora

Surface morphology was examined in tepals, floral bracts and basal bracts in *B. laxiflora*. There are up to 14 bracts (sometimes referred to as 'leaves') at the bases of the male and female inflorescences of *B. laxiflora* (Fig. 1A, B). Each of the male flowers has two small subtending floral bracts (Fig. 1B, D). The epidermal cells of these floral bracts are oval in shape and the surface is papillate covered with short ridges (Fig. 5A). Although the basal bracts are showy, their epidermal cells are flattened and elongated with smooth surfaces on both adaxial and abaxial sides (Fig. 5E, F). In contrast, the epidermal cells of mature tepals are oval in shape and their surface is smooth on the adaxial side (Fig. 5D). They also show irregular cuticular sculpturing on their abaxial sides (Fig. 5B, C), similar to what was observed on the floral bracts (Fig. 5A). No stomata were observed on either side of the tepals, floral bracts or basal bracts.

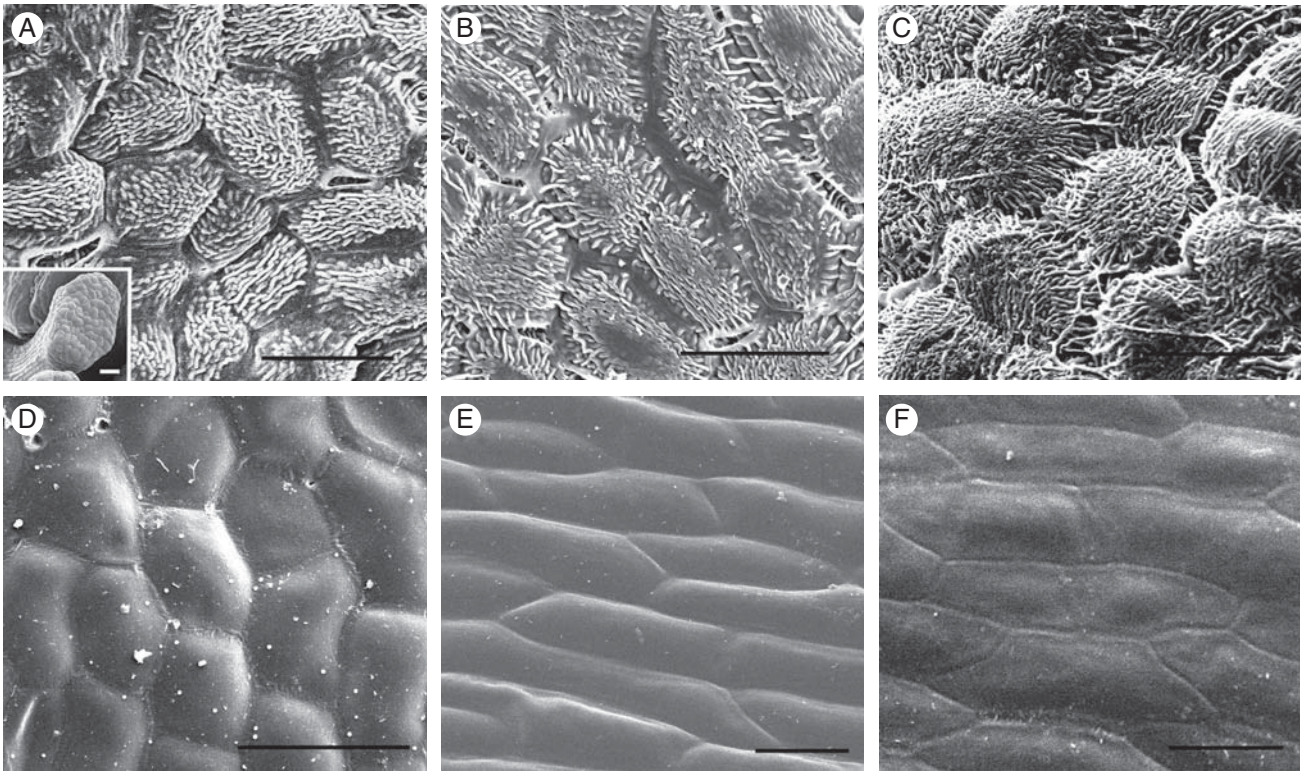

FIG. 5. Scanning electron micrographs of tepals and bracts of *Balanophora laxiflora*: (A) epidermal cells of the floral bract; (B) abaxial epidermal cells of the young tepals; (C) abaxial epidermal cells of the tepals at maturity; (D) adaxial epidermal cells of the tepals at maturity; (E) adaxial epidermal cells of the basal bract at maturity; (F) abaxial epidermal cells of the basal bract at maturity. Scale bars = 50 μm.

DISCUSSION

Rate variation in Balanophora *species and other Santalales*

It has previously been shown that *Balanophora* and other taxa of Balanophoraceae have significantly increased rates of nucleotide substitution in the 18S rDNA (Nickrent *et al.*, 2005) and *matR* (Barkman *et al.*, 2007) genes. In this study, we demonstrated that extreme rate acceleration is not ubiquitous in the *Balanophora* genome. The six B-class genes of *Balanophora* identified, three each from *B. laxiflora* and *B. fungosa*, show more modest nucleotide substitution rates as compared with rates for 18S rDNA and *matR*. Non-synonymous (d_N) and synonymous (d_S) substitutions per site were both estimated for protein-coding genes. In general, a lower $d_N : d_S$ ratio indicates a conserved sequence or an increased rate of substitutions at silent sites, whereas a higher $d_N : d_S$ ratio indicates a relaxation of purifying selection. Therefore the d_S value is an indication of non-selective substitution rate, i.e. similar to the background substitution of a genome (Yang, 1998). In our results, both *PI* and *TM6* display significant d_S rate accelerations compared with d_N rate, whereas the eu*AP3* paralogues are significantly elevated in d_N rate (but not d_S rate), suggesting that *Balanophora* eu*AP3* has evolved under a relaxed purifying selection, but they do not show congruent patterns among B-class genes.

In general, positive correlations between d_N and d_S are found in nuclear genes of mammals, prokaryotes and *Drosophila* (Graur, 1985; Li *et al.*, 1985; Wolfe and Sharp, 1993; Akashi, 1994). In plants, d_N and d_S of the nuclear paralogues of *Arabidopsis thaliana* (Zhang *et al.*, 2002) and the plastid genes of Geraniaceae (Guisinger *et al.*, 2008) are also reported to be positively correlated. Our result also found significant correlations between d_N and d_S in the four protein-coding genes (except for *PI*, see Table 1), suggesting that the intensity of selective constraints on both d_N and d_S are similar in these gene regions. The reason for the incongruence of *PI* data requires further investigation.

Although the rates of sequence evolution detected from d_N and d_S in *Balanophora* are generally higher than those of other plants, the $d_N : d_S$ ratios of the six protein-coding genes examined for the three groups (*Balanophora*, hemiparasitic Santalales and core eudicots) are all $<0{\cdot}3$, with the exception of mitochondrial *matR* (Table 1), suggesting the B-class, *LFY* and *RPB2* genes are being subjected to strong purifying selection. However, functional constraints alone cannot explain the rate heterogeneity in *Balanophora* since rDNA should be also under strong selection against nucleotide changes. The 18S rDNA of *Balanophora* has at least a 5-fold higher substitution rate compared with other flowering plants and it is even more extreme (10-fold higher) for its *matR* gene.

Unusual acceleration of substitution rate in genes or genomes of certain species or lineages have been documented in several plants, mostly for mitochondria and plastids. Mitochondrial-specific increases in substitution rate were found in some *Plantago* (Cho *et al.*, 2004), *Pelargonium hortorum* (Parkinson *et al.*, 2005) and several *Silene* spp. (Sloan

et al., 2009). The nuclear and plastid genomes of these taxa are largely unaffected and possess typical substitution rates. On the other hand, plastid genes in parasitic plants sometimes show rate variations due to different degrees of selective constraints on photosynthetic genes. For example, the synonymous rate of *Cuscuta obtusiflora* plastid genes is five to eight times higher than the rates in *C. exaltata* because of a higher functional constraint for photosynthesis in the former (McNeal *et al.*, 2007). Such rate increases are usually found to be lineage-specific and probably affecting the whole plastid genome, as in the case of Orobanchaceae (Young and dePamphilis, 2005).

The pattern of rate variation in 18S rDNA is more complicated since there is no clear relationship between elevated substitution rates and functional constraints. Although the extreme rate acceleration cases of 18S rDNA are all found in heterotrophic taxa such as *Rafflesia keithii* (Rafflesiaceae) and *Balanophora* spp., not all holoparasitic plants display such rate accelerations (Nickrent and Starr, 1994; Nickrent *et al.*, 1998; Young and dePamphilis, 2005; Lemaire *et al.*, 2010). *Cynomorium coccineum* (Cynomoriaceae) and *Orobanche fasciculata* (Orobanchaceae) do not show significantly accelerated substitution rates in 18S rDNA (dePamphilis *et al.*, 1997; Wolfe and dePamphilis, 1997; Lemaire *et al.*, 2010). Similarly, in mycoheterotropic plants, rate accelerations in 18S rDNA can be found in *Rhizanthella gardneri* (Orchidaceae) and *Thismia aseroe* (Thismiaceae); others like *Monotropa uniflora* (Ericaceae) and *Corallorhiza maculata* (Orchidaceae) lack such rate increases (Lemaire *et al.*, 2010).

It was suggested that the functionality of 18S rDNA with accelerated substitution rates is retained in those heterotrophic plants, given the pattern of compensatory mutations, and that few mutations occur in the functionally and structurally important regions of the ribosomal molecule (Nickrent and Starr, 1994; dePamphilis, 1995; Lemaire *et al.*, 2010). Therefore the increased substitution rate in 18S rDNA might reflect an overall elevated mutation rate in the nuclear genome in these plants. Several possible causes of the elevated substitution rates have been proposed, such as defective DNA repair efficiency, rapid generational time, higher speciation rates and small effective population size (Nickrent and Starr, 1994; Nickrent *et al.*, 1998; Lemaire *et al.*, 2010). However, such a hypothesis would suggest an overall elevated substitution rate in the entire genome for that species.

Here we provided evidence of rate heterogeneity in nuclear genomes of the two *Balanophora* spp. The substitution rates in 18S rDNA in *B. laxiflora* and *B. fungosa* are five times higher than other flowering plants, whereas the synonymous substitutions in *PI*, *AP3*/*TM6*, *LFY* and *RPB2* show a comparable rate (0·5–2·0 times higher) to the autotrophic core eudicots. These results suggest that the molecular evolutionary factors affecting substitution rates in 18S rDNA might be different from nuclear protein-coding genes, at least for the B-class, *LFY* and *RPB2* genes demonstrated in this study. Three models of rate variation have been proposed for explaining nucleotide substitution patterns: (1) gene-specific, (2) organism-specific and (3) genome-specific (Nickrent *et al.*, 1998). Our preliminary data on the six protein-coding genes indicate that the rate variation in *Balanophora* probably evolved in a gene-specific manner. The mechanism for such rate heterogeneity in *Balanophora* is unknown and a thorough screening for other nuclear genes will reveal if it is ubiquitous in the *Balanophora* genome. It also remains to be determined if a rate discrepancy exists between 18S rDNA and other nuclear genes for the heterotrophic plants that are known to have highly elevated 18S rDNA substitutions.

B-class genes in Santalales and the novel exon in Balanophora AP3

Despite the fact that many *PI* and *TM6* homologues were successfully identified from the selected members of Santalales including *Balanophora*, only two eu*AP3* homologues were found in the two *Balanophora* spp. examined, and no eu*AP3* homologues were detected in other members of Santalales. Different pairs of degenerate primers, including those most commonly used in previous studies (Kramer *et al.*, 1998; Stellari *et al.*, 2004) and primers designed specifically for the *Balanophora* eu*AP3* genes (Bal*AP3* and Baf*AP3*), were employed under several PCR conditions. The lack of eu*AP3* sequences from these taxa of Santalales indicates that their homologues are either too divergent to be detected by PCR-based screening or their genes might express at a low levels, thus escaping detection using RT-PCR.

Most of the B-class genes in angiosperms consist of seven exons and six introns. Our alignment of eu*AP3* homologues showed that *AP3* from both *Balanophora* species have relatively conserved M, I and K domains. Exons 1–5 in the *Balanophora* B-class genes have the same length as those in *Vitis*, but the length of the C-domain is quite variable (Supplementary Data Fig. S3). An additional *AP3* exon in both *Balanophora* species is found between the C1 and C2 domains, which are known to involve transcriptional activation (Riechmann and Meyerowitz, 1997). However, we were unable to find other cases of similarly modified C-domains in other plants, and thus the issue of its origin is unclear at this time. Moreover, an intensive screening of Santalales and/or other members of Balanophoraceae is required to understand better the evolution of these B-class genes and to make inferences on their function.

Differential expression of B-class genes in Balanophora laxiflora

The expression patterns of *Balanophora AP3* and *TM6* suggest a largely redundant role for these two genes since they are co-expressed at comparable levels in all organs, except that *TM6* expression is lower in pollen and floral bracts (Fig. 4B). Given the fact that *AP3* has distinct C-terminal sequences, it is reasonable to propose a certain degree of functional differentiation from *TM6* in *Balanophora*. In contrast, *PI* might have a conserved B-class gene function since it has strong expressions in tepals and synandria, but not in any reproductive organs of the female *Balanophora* plant. This result is congruent with the floral ABCDE model prediction, since no perianth and androecium structures can be found in the female plant of *B. laxiflora*. The tepals in male *B. laxiflora* have cobblestone-like epidermal cells that are similar to those seen in the petals of other plants which have short cuticular ridges on the surface (Kay and Daoud, 1981). Similar cuticular modification can also be found in the epidermal cells of floral bracts in the male *B. laxiflora*.

It remains to be determined if the expression of *PI* in the floral bracts of *Balanophora* is correlated with epidermal cell morphology in these organs (tepals, synandria and floral bracts).

Phylogenetic utility of nuclear RPB2 *and floral B-class genes in Balanophoraceae*

Although 18S rDNA and *matR* are widely used markers in reconstructing angiosperm phylogeny, the extreme rate acceleration of nucleotide substitution in these genes in *Balanophora* presents problems for accurate phylogenetic reconstruction. Our results clearly showed that *Balanophora* is a member of Santalales based on both *RPB2* and B-class gene phylogenies, even though the exact relationships are still problematic owing to low support for some nodes. We demonstrated that the evolutionary rates of the nuclear protein-coding genes are relatively conserved compared with 18S rDNA and *matR*; thus they represent promising markers for further phylogenetic analyses. Future work will be to increase the sampling of the highly diverse Olacaceae *s.l.* (Malecot and Nickrent, 2008).

Conclusions

For the holoparasite *Balanophora*, nucleotide substitution rates in the B-class genes, *LFY* and *RPB2* homologues show that these nuclear protein-coding genes are relatively conserved compared with 18S rDNA and mitochondrial *matR*. These sequences are thus potential candidates for future phylogenetic analyses of Balanophoraceae. There is rate heterogeneity between 18S rDNA, mitochondrial *matR* and the five nuclear protein-coding genes, hinting at an unusual and currently unexplored mechanism of sequence evolution for the mitochondrial and nuclear ribosomal genes of *Balanophora*.

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxfordjournals.org and consist of the following. Table S1: primers used in this study. Table S2: listing of the GenBank accession numbers for all sequences used in this study. Figure S1: *RPB2* gene tree showing the obtained *RPB2* sequences belonging to the d clade. Figure S2: conserved C-terminal of Santalales B-class gene homologues. Figure S3: genomic structure of *Balanophora AP3* and *TM6* homologues.

ACKNOWLEDGEMENTS

We appreciate Daniel Nickrent and another anonymous reviewer for making valuable comments on the manuscript and Shao-Lun Liu for helpful discussions. We appreciate the help provided by Cheng-Chiang Wu, Jiunn-Yih Huang, Shu-hui Wu and Shiu-Chung Wu in obtaining specimens. We are also grateful for the technical assistance with SEM by the staff of Technology Commons, College of Life Science, NTU. This study was supported by grants from the National Science Council, Taiwan (NSC 96-2621-B-002-008-MY3 and 99-2918-I-002-020).

LITERATURE CITED

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161**: 105–121.

Akashi H. 1994. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. *Genetics* **136**: 927–935.

Barkman TJ, Lim SH, Salleh KM, Nais J. 2004. Mitochondrial DNA sequences reveal the photosynthetic relatives of *Rafflesia*, the world's largest flower. *Proceedings of the National Academy of Sciences of the USA* **101**: 787–792.

Barkman TJ, McNeal JR, Lim SH, *et al*. 2007. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evolutionary Biology* **7**: 248. http://dx.doi.org/10.1186/1471-2148-7-248.

Blázquez MA, Soowal LN, Lee I, Weigel D. 1997. *LEAFY* expression and flower initiation in *Arabidopsis*. *Development* **124**: 3835–3844.

Britten RJ. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231**: 1393–1398.

Chang S, Puryear J, Chirney J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* **11**: 113–116.

Cho Y, Mower JP, Qiu Y-L, Palmer JD. 2004. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proceedings of the National Academy of Sciences of the USA* **101**: 17741–17746.

Coen ES, Meyerowitz EM. 1991. The war of the whorls: genetic interactions controlling flower development. *Nature* **353**: 31–37.

dePamphilis CW. 1995. Genes and genomes. In: Press MC, Graves JD. eds. *Parasitic plants*. London: Chapman & Hall, 177–205.

dePamphilis CW, Young ND, Wolfe AD. 1997. Evolution of plastid gene *rps2* in a lineage of hemiparasitic and holoparasitic plants: many losses of photosynthesis and complex patterns of rate variation. *Proceedings of the National Academy of Sciences of the USA* **94**: 7367–7372.

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

Duff RJ, Nickrent DL. 1997. Characterization of mitochondrial small-subunit ribosomal RNAs from holoparasitic plants. *Journal of Molecular Evolution* **45**: 631–639.

Eberwein R, Nickrent DL, Weber A. 2009. Development and morphology of flowers and inflorescences in *Balanophora papuana* and *B. elongata* (Balanophoraceae). *American Journal of Botany* **96**: 1055–1067.

Felsenstein J. 1985. Confidence limits on phylogenies with a molecular clock. *Systematic Zoology* **34**: 152–161.

Frohlich MW, Meyerowitz EM. 1997. The search for flower homeotic gene homologs in basal angiosperms and Gnetales: a potential new source of data on the evolutionary origin of flowers. *International Journal of Plant Sciences* **158**: S131–S142.

Gaut BS, Morton BR, McCaig BC, Clegg MT. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear *Adh* parallel rate differences at the plastid gene *rbcL*. *Proceedings of National Academy of Sciences of the USA* **93**: 10274–10279.

Graur D. 1985. Amino acid composition and the evolutionary rates of protein-coding genes. *Journal of Molecular Evolution* **22**: 53–62.

Guisinger MM, Kuehl JNV, Boore JL, Jansen RK. 2008. Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. *Proceedings of the National Academy of Sciences of the USA* **105**.: 18424–18429.

Hansen B. 1972. The genus *Balanophora* J. R. & G. Forster, a taxonomic monograph. *Dansk Botanisk Arkiv* **28**: 1–188.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.

Jaramillo MA, Kramer EM. 2007. Molecular evolution of the petal and stamen identity genes, *APETALA3* and *PISTILLATA*, after petal loss in the Piperales. *Molecular Phylogenetics and Evolution* **44**: 598–609.

Kay QON, Daoud HS. 1981. Pigment distribution, light reflection and cell structure in petals. *Botanical Journal of the Linnean Society* **83**: 57–84.

Kramer EM, Irish VF. 2000. Evolution of the petal and stamen developmental programs: evidence from comparative studies of the lower eudicots and basal angiosperms. *International Journal of Plant Sciences* **161**: S29–S40.

Kramer EM, Dorit RL, Irish VF. 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.

Kramer EM, Su HJ, Wu CC, Hu JM. 2006. A simplified explanation for the frameshift mutation that created a novel C-terminal motif in the *APETALA3* gene lineage. *BMC Evolutionary Biology* **6**: 30. http://dx.doi.org/10.1186/1471-2148-6-30.

Laroche J, Li P, Maggia L, Bousquet J. 1997. Molecular evolution of angiosperm mitochondrial introns and exons. *Proceedings of the National Academy of Sciences of the USA* **94**: 5722–5727.

Lemaire B, Huysmans S, Smets E, Merckx V. 2010. Rate accelerations in nuclear 18S rDNA of mycoheterotrophic and parasitic angiosperms. *Journal of Plant Research* **124**: 561–576.

Li WH, Wu CI, Luo CC. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Molecular Biology and Evolution* **2**: 150–174.

Luo J, Yoshikawa N, Hodson MC, Hall BD. 2007. Duplication and paralog sorting of *RPB2* and *RPB1* genes in core eudicots. *Molecular Phylogenetics and Evolution* **44**: 850–862.

McNeal JR, Kuehl JV, Boore JL, de Pamphilis CW. 2007. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. *BMC Plant Biology* **7**: 57. http://dx.doi.org/10.1186/1471-2229-7-57.

Malecot V, Nickrent DL. 2008. Molecular phylogenetic relationships of Olacaceae and related Santalales. *Systematic Botany* **33**: 97–106.

Moyroud E, Tichtinsky G, Parcy F. 2009. The *LEAFY* floral regulators in angiosperms: conserved proteins with diverse roles. *Journal of Plant Biology* **52**: 177–185.

Nickrent DL, Starr EM. 1994. High rates of nucleotide substitution in nuclear small-subunit (18S) rDNA from holoparasitic flowering plants. *Journal of Molecular Evolution* **39**: 62–70.

Nickrent DL, Duff RJ, Colwell AE, *et al.*1998. Molecular phylogenetic and evolutionary studies of parasitic plants. In: Soltis DE, Soltis PS, Doyle JJ. eds. *Molecular systematics of plants. II. DNA sequencing*. Boston, MA: Kluwer Academic Publishers, 211–241.

Nickrent DL, Der JP, Anderson FE. 2005. Discovery of the photosynthetic relatives of the 'Maltese mushroom' *Cynomorium*. *BMC Evolutionary Biology* **5**: 38. http://dx.doi.org/10.1186/1471-2148-5-38.

Nickrent DL, Duff RJ, Konings DAM. 1997. Structural analyses of plastid-derived 16S rRNAs in holoparasitic angiosperms. *Plant Molecular Biology* **34**: 731–743.

Ohta T. 2000. Mechanisms of molecular evolution. *Philosophical Transactions of the Royal Society of London – Series B: Biological Sciences* **355**: 1623–1626.

Oxelman B, Yoshikawa N, McConaughy BL, Luo J, Denton AL, Hall BD. 2004. *RPB2* gene phylogeny in flowering plants, with particular emphasis on asterids. *Molecular Phylogenetics and Evolution* **32**: 462–479.

Parkinson CL, Mower JP, Qiu Y-L, *et al.*2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evolutionary Biology* **5**: 73. http://dx.doi.org/10.1186/1471-2148-5-73.

Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**: 676–679.

Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**: 793–808.

R Development Core Team. 2008. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. www.r-project.org.

Riechmann JL, Meyerowitz EM. 1997. MADS domain proteins in plant development. *Biological Chemistry* **378**: 1079–1118.

Shan HY, Su KM, Lu WL, Kong HZ, Chen ZD, Meng Z. 2006. Conservation and divergence of candidate class B genes in *Akebia trifoliata* (Lardizabalaceae). *Development Genes and Evolution* **216**: 785–795.

Sloan DB, Oxelman B, Rautenberg A, Taylor DR. 2009. Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. *BMC Evolutionary Biology* **10**: 12. http://dx.doi.org/10.1186/1471-2148-10-12.

Smith SA, Donoghue MJ. 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science* **322**: 86–89.

Soltis DE, Soltis PS, Chase MW, *et al.*2000. Angiosperm phylogeny inferred from 18S rRNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* **133**: 381–461.

Soltis PS, Soltis DE, Chase MW. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* **402**: 402–404.

Stellari GM, Jaramillo MA, Kramer EM. 2004. Evolution of the *APETALA3* and *PISTILLATA* lineages of MADS-box-containing genes in the basal angiosperms. *Molecular Biology and Evolution* **21**: 506–519.

Theissen G, Saedler H. 2001. Floral quartets. *Nature* **409**: 469–471.

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Münster T. 2000. A short history of MADS-box genes in plants. *Plant Molecular Biology* **42**: 115–149.

Weigel D, Meyerowitz EM. 1994. The ABCs of floral homeotic genes. *Cell* **78**: 203–209.

Wolfe AD, de Pamphilis CW. 1997. Alternate paths of evolution for the photosynthetic gene *rbcL* in four nonphotosynthetic species of *Orobanche*. *Plant Molecular Biology* **33**: 965–977.

Wolfe KH, Sharp PM. 1993. Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. *Journal of Molecular Evolution* **37**: 441–456.

Yang ZH. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular Biology and Evolution* **15**: 568–573.

Young ND, dePamphilis CW. 2005. Rate variation in parasitic plants: correlated and uncorrelated patterns among plastid genes of different function. *BMC Evolutionary Biology* **5**: 16. http://dx.doi.org/10.1186/1471-2148-5-16.

Zhang LQ, Vision TJ, Gaut BS. 2002. Patterns of nucleotide substitution among simultaneously duplicated gene pairs in *Arabidopsis thaliana*. *Molecular Biology and Evolution* **19**: 1464–1473.

Zwickl DJ. 2006. *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. PhD Dissertation, University of Texas, Austin, TX, USA.