

Seishiro Aoki · Koichi Uehara · Masao Imafuku
Mitsuyasu Hasebe · Motomi Ito

Phylogeny and divergence of basal angiosperms inferred from *APETALA3*- and *PISTILLATA*-like MADS-box genes

Received: December 29, 2003 / Accepted: March 10, 2004 / Published online: May 12, 2004

Abstract The B-class MADS-box genes composed of *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) lineages play an important role in petal and stamen identity in previously studied flowering plants. We investigated the diversification of the *AP3*-like and *PI*-like MADS-box genes of eight species in five basal angiosperm families: *Amborella trichopoda* (Amborellaceae); *Brasenia schreberi* and *Cabomba caroliniana* (Cabombaceae); *Euryale ferox*, *Nuphar japonicum*, and *Nymphaea tetragona* (Nymphaeaceae); *Illicium anisatum* (Illiciaceae); and *Kadsura japonica* (Schisandraceae). Sequence analysis showed that a four amino acid deletion in the K domain, which was found in all previously reported angiosperm *PI* genes, exists in a *PI* homologue of Schisandraceae, but not in six *PI* homologues of the Amborellaceae, Cabombaceae, and Nymphaeaceae, suggesting that the Amborellaceae, Cabombaceae, and Nymphaeaceae are basalmost lineages in angiosperms. The results of molecular phylogenetic analyses were not inconsistent with this hypothesis. The *AP3* and *PI* homologues from *Amborella* share a sequence of five amino acids in the 5' region of exon 7. Using the linearized tree and likelihood methods, the divergence time between the *AP3* and *PI* lineages was estimated as somewhere between immediately after to several tens of millions of years after the split between angiosperms and extant gymnosperms. Estimates of the age of the most

recent common ancestor of all extant angiosperms range from ~140–210 Ma, depending on the trees used and assumptions made.

Key words *Amborella* · Floral homeotic genes · *Illicium* · *Kadsura* · MADS · Nymphaeales

Introduction

There are at least 250,000 species of angiosperms, and the morphologies of their reproductive organs, flowers, vary extensively (Thorne 1992; Crane et al. 1995). Great efforts have been made to understand the extent and origin of the diversity of flowering plants, and progress has been made by applying molecular analysis to the fields of phylogenetics and developmental biology. Studies have shown the monophyletic relationship of extant angiosperms, but the early evolution of angiosperms, as well as their closest relatives among gymnosperms, is still unclear because of a lack of appropriate fossils and ambiguous phylogenetic placement of basal angiosperm lineages (Judd et al. 2002). Five groups of angiosperms have been identified as the earliest flowering plants: Amborellaceae, Nymphaeales, Illiciales, Trimeniaceae, and *Austrobaileya* (Mathews and Donoghue 1999; Parkinson et al. 1999; Qiu et al. 1999; Soltis et al. 1999; Borsch et al. 2003). These groups of angiosperms are termed the ANITA grade. Although these groups are likely basal, further investigation is still necessary to determine the exact order of relationships among them (Barkman et al. 2000; Graham and Olmstead 2000; Mathews and Donoghue 2000; Zanis et al. 2002; Goremykin et al. 2003).

Investigation of the diversity of angiosperms often makes use of the structures of flowers. To investigate floral structure and development, molecular genetic studies used dicotyledonous plants, such as *Arabidopsis thaliana*, *Antirrhinum majus*, and *Petunia hybrida*. These plants have the general floral structure of eudicotyledonous plants, with four organ types (sepal, petal, stamen, and gynoecium) arranged in four whorls. Molecular genetic analyses with

S. Aoki (✉)
Department of Life Sciences, Graduate School of Arts and Sciences,
University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902,
Japan
Tel. +81-3-54546638; Fax +81-3-54546638
e-mail: saoki@bio.c.u-tokyo.ac.jp

K. Uehara
Faculty of Horticulture, Chiba University, Chiba, Japan

M. Imafuku · M. Ito
Graduate School of Arts and Sciences, University of Tokyo, Tokyo,
Japan

M. Hasebe
National Institute for Basic Biology, Okazaki, Japan

M. Hasebe
Department of Molecular Biomechanics, The Graduate University for
Advanced Studies, Okazaki, Japan

these plants have focused on a set of homeotic floral-organ mutants, leading to a model in which five classes of loci, known as the A, B, C, D, and E classes, interact in a combinatorial manner to specify organ identities (Bowman et al. 1991; Coen and Meyerowitz 1991; Meyerowitz et al. 1991; Theissen 2001; Theissen and Saedler 2001). In mutants of the B-class loci, petals are transformed into sepals and stamens into carpels, indicating that the products of the B-class genes are required for the specification of petal and stamen identities (Sommer et al. 1990; Jack et al. 1992; Tröbner et al. 1992; Goto and Meyerowitz 1994). Like many floral-organ identity genes, the *Arabidopsis* B-class genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) belong to the MADS-box gene family, which encodes transcription factors with four domains (Fig. 1; Ma et al. 1991; Richmann and Meyerowitz 1997; Irish 2003). The region in these proteins that contains the highly conserved MADS-box sequence is termed the MADS domain. The second of the four domains, the K domain, is similar to the coiled-coil segment of keratin. The intervening region between the MADS and K domains is termed the I domain, and the C-terminal region is also called the C domain. MADS-box genes that contain these four regions are specific to plants and are designated as the MIKC-type MADS-box genes.

Phylogenetic analysis of the MADS-box gene family has revealed that this family is composed of several defined gene clades (Theissen et al. 1996; Munster et al. 1997; Hasebe et al. 1998; Becker and Theissen 2003). The establishment of these gene clades is thought to have been an important event in the evolution of flowers (Theissen et al. 1996, 2000; Bowman 1997; Kramer et al. 1998; Becker et al. 2002). The clade of B-class genes of angiosperms comprises two lineages: the *AP3* and the *PI* homologues. Numerous genes of the *AP3* and *PI* lineages have been isolated from eudicots and magnoliids (Kramer et al. 1998; Kramer and Irish 2000; Kramer et al. 2003). Recently, putative orthologs of angiosperm B-class genes have been isolated from several gymnosperms (Mouradov et al. 1999; Sundström et al. 1999; Winter et al. 1999; Fukui et al. 2001). A group of MADS-box genes has also been identified as the sister lineage of the B-class genes and termed the Bsister (Bs) clade (Becker et al. 2002). Phylogenetic studies with these genes have provided evidence that the evolution of these lineages was complex, involving many duplication events followed

by considerable sequence divergence (Kramer and Irish 2000; Nam et al. 2003). It has been suggested that additional sequence data from basal angiosperms would clarify the relationships among these genes (Kramer and Irish 2000).

We identified putative B-class MADS-box genes from eight species of the basal angiosperm groups Amborellaceae, Nymphaeales, Illiciaceae, and Schisandraceae. The molecular evolutionary analysis that we performed using these genes will advance the understanding not only of the evolution of floral homeotic genes but also of the phylogenetic relationships among the major lineages of basal angiosperms. Furthermore, we use several methods of analysis to estimate the divergence times between the *AP3* and *PI* lineages as well as the time of first divergence of the extant angiosperms.

Materials and methods

Sources of plant materials

The plant materials used in this study are shown in Table 1. Developing floral buds and flowers were collected throughout the growing season from the field or botanical gardens. The taxonomic designations in this study follow the system proposed by APG (Angiosperm Phylogeny Group) II (2003) and two other studies (Soltis et al. 1999; Endress 2001).

Isolation of *AP3*-like and *PI*-like MADS-box genes

Total RNA from flowers, floral buds, and floral organs was prepared using an improved version of the CTAB method (Shindo et al. 1999). Single-stranded complementary DNA (cDNA) was synthesized using the 3' Rapid Amplification of cDNA Ends (RACE) System (Gibco BRL). Primers for degenerate polymerase chain reactions (PCR) were designed to match the MADS-box sequences, as follows: MADS head primer, 5'-CAUCAUCAUATGGGIMGIGGIAARATHGARAT-3'; *AP3*MADS3 primer, 5'-CAUCAUCAUACIAAYMGICARGTACNTA-3'. PCR analysis was performed in 25- μ l reactions containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 1 μ M

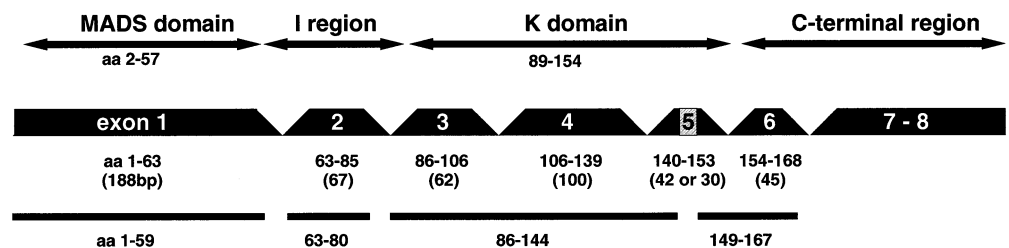


Fig. 1. Diagram of the exon/intron organization of the B-class genes. Intron positions are indicated by *triangular cutouts*. Numbers of amino acid residues from the methionine initiation codon are shown *below the diagram*. Exon lengths in nucleotides are shown in *parentheses* (see also Table 3). The MADS, K, I, and C domain regions are indicated

above the diagram. The ranges of MADS and K domains are consistent with those reported by Ma et al. (1990). *Lines with numbers shown below the diagram* indicate the region used for the phylogenetic analysis

Table 1. Plant materials and homologues^a of B-class genes identified in this study

Family	Species	Source	AP3	PI
Amborellaceae	<i>Amborella trichopoda</i> Baill.	Cultivated in Botanical Gardens, University of Tokyo, Koishikawa	<i>AmAP3</i> (2)	<i>AmPI</i> (3)
Cabombaceae	<i>Brasenia schreberi</i> Gmel.	Yozawa-Monya pond, Ibaraki pref.	<i>BsAP3</i> (6)	<i>BsPI</i> (4)
Cabombaceae	<i>Cabomba caroliniana</i> A.	Otonuma pond, Ibaraki pref.	<i>CcAP3</i> (2)	<i>CcPI</i> (4)
Nymphaeaceae	<i>Euryale ferox</i> Salisb.	Sagata Lake, Niigata pref.	<i>EfAP3</i> (3)	<i>EfPI</i> (2)
Nymphaeaceae	<i>Nuphar japonicum</i> DC.	Cultivated in Botanical Gardens, University of Tokyo, Nikko	<i>NjAP3-1</i> (2) <i>NjAP3-2</i> (4)	<i>NjPI-1</i> (3) <i>NjPI-2</i> (7)
Nymphaeaceae	<i>Nymphaea tetragona</i> Georgi.	Koikenuma pond, Gunma pref.	<i>NtAP3-1</i> (4) <i>NtAP3-2</i> (1)	<i>NtPI</i> (3)
Illiciaceae	<i>Illicium anisatum</i> L.	Cultivated in Botanical Gardens, University of Tokyo, Koishikawa	<i>IaAP3-1</i> (2) <i>IaAP3-2</i> (2) <i>IaAP3-3</i> (2)	
Schisandraceae	<i>Kadsura japonica</i> Benth.	Cultivated in Botanical Gardens, University of Tokyo, Koishikawa	<i>KjAP3</i> (2)	<i>KjPI</i> (3)

^aThe numbers in parentheses denote the number of independent clones sequenced for that gene

each of the 5' and 3' primers, 200 μM dNTPs, and 2 units of AmpliTaq Gold polymerase (Perkin-Elmer). Amplifications were performed using the GeneAmp PCR System 9700 (PE Applied Biosystems) with the following program: preheating at 96 °C for 12 min; 45 cycles of 95 °C for 20 s (denaturation), 52 °C for 20 s (annealing), and 72 °C for 1.5 min (extension); and post-heating at 72 °C for 7 min. The annealing temperature was modified appropriately for different primers and plant materials. PCR products of about 0.8–1 kbp were purified using gel electrophoresis and inserted into the pAMP vector (Promega). Both strands of the inserted DNA fragments were sequenced with the T7 and SP6 primers. Primers for 5' RACE were then designed according to the sequences obtained from the 3' RACE amplification, and isolation and sequencing of the fragments was performed using the 5' RACE System (Gibco BRL). The number of independent clones sequenced for each gene is shown in Table 1. We obtained several clones from different PCR reactions for each gene except *NtAP3-2*. Nucleotide sequence of *NtAP3-2* is determined from a clone and may contain PCR errors. The sequences of cDNAs isolated in this study were submitted to the DDBJ/EMBL/GenBank sequence database (Appendix).

Genomic DNA of plant materials, mainly leaves, was isolated using the DNeasy Plant Mini Kit (Qiagen). Partial genomic DNA fragments were isolated using the shuttle-PCR method with suitable primers and the following program: preheating at 96 °C for 12 min; 45 cycles of 95 °C for 20 s (denaturation) and 58 °C for 2 min (annealing and extension); and post-heating at 72 °C for 7 min. The resulting products were inserted into the pGEM-T vector (Promega) and sequenced. The genomic sequences identified in this study have been deposited into the DDBJ/EMBL/GenBank sequence database with accession numbers AB154840, AB164842, and AB154847.

Phylogenetic analysis

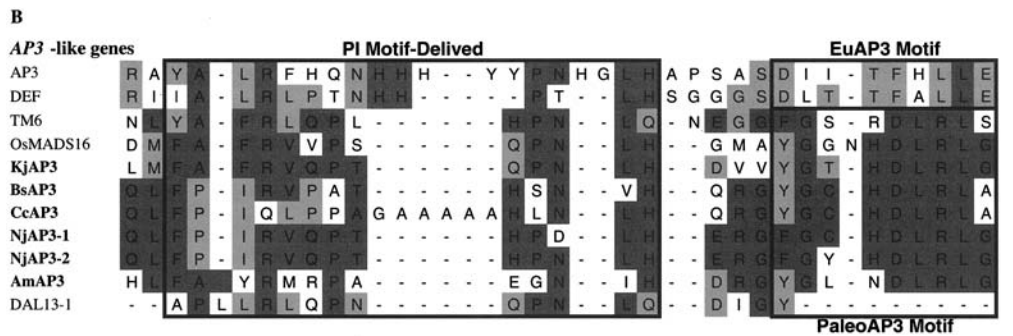
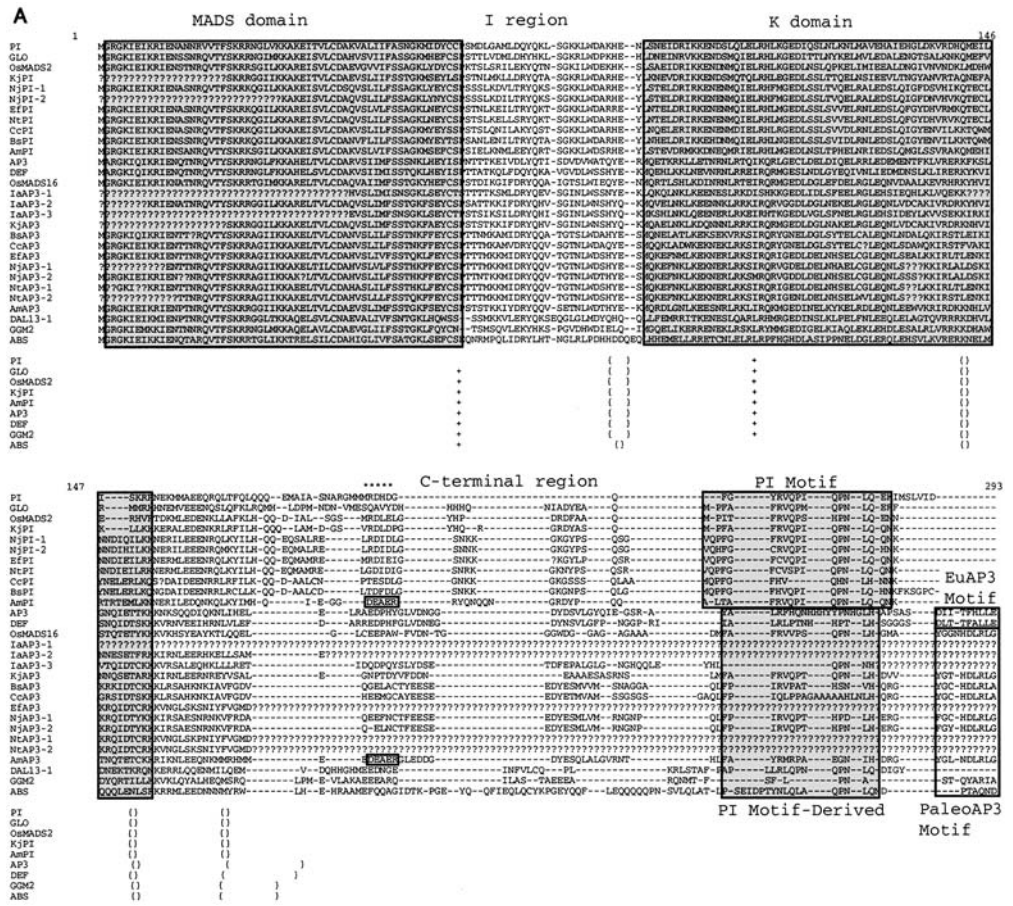
At present, more than 200 sequences of B-class and Bs MADS-box genes from various plant species are available.

We compiled 125 of these sequences into a data matrix. Some genes (e.g., *CjMADS2* of *Cryptomeria japonica*; Fukui et al. 2001) were omitted from the phylogenetic analysis because of deletions in their sequences. *DAL12* (Sundström et al. 1999) and its putative ortholog *GGM15* (Winter et al. 2002) were also eliminated because these genes have an exon 5 of atypical length. These genes are thought to have undergone unusual changes (Winter et al. 2002). As the sequences of several gene sets from closely related plants are highly similar, we used representative sequences from each gene set. Recently, many genes have been isolated from plants in the Ranunculales and it has been shown that B-class genes of these plants have undergone many duplication events (Kramer et al. 2003). The phylogenetic tree of the Ranunculales genes has clades at various phylogenetic levels (Kramer et al. 2003), and we included representative members from each. We omitted *GGM2* (Winter et al. 1999; Becker et al. 2000, 2002; Tanabe et al. 2003) because the clade of gymnosperm B-class genes did not include this gene.

Sequences that are divergent in several regions, such as the I and C-terminal regions, complicate alignment of this gene family and the previously reported alignments of MADS-box genes vary greatly. We accepted the alignment method of Johansen et al. (2002). All of the protein sequences, including both previously reported sequences and those identified in this work, were aligned exon by exon according to the known gene structures, as the individual exons are likely homologous (Fig. 2A). Genes for which genomic sequences were available were aligned first, and subsequently genes for which only the cDNA sequence is known were grouped and aligned to the others. Alignment was initially performed using the program CLUSTAL W version 1.7 with a gap-opening penalty of 15.00 and a gap extension penalty of 6.66, followed by manual adjustment, taking into account gene structure. Prior to phylogenetic analysis, regions in which indels (insertion/deletion) were present in some sequences were excluded from the data set, and the sequence lengths were adjusted. As the alignment was not straightforward at the terminal regions of exons 1 and 2, these sites were also omitted before analysis. The

Fig. 2A,B. Sequence alignment of B-class MADS-box genes. **A** Predicted amino acid sequences of MADS-box genes from basal angiosperms aligned with representatives of eudicots and monocots. *Question marks* indicate that the sequence is ambiguous. The MADS and K domains and characteristic motifs of the B/Bs genes are *boxed*. The identical five amino acids between the *AmAP3* and *AmPI* genes are also *boxed* and the sites of this motif are marked by *asterisks*. *Brackets* indicate the positions of introns located between codons (zero-phase introns). *Plus signs* indicate the position of introns located within codons. They are all phase 2 introns, i.e., they are located between the second and third codon positions. DEF and GLO are derived from *Antirrhinum majus*; PI, AP3, ABS/AGL32 from *Arabidopsis thaliana*; OsMADS2 and OsMADS16 from *Oriza sativa*; DAL13-1 from *Picea abies*; and GGM2 from *Gnetum gnemon*.

B Alignment of the C-terminal regions of the deduced protein sequences. The names of the genes isolated in this study are in *bold*. The motifs identified by Kramer et al. (1998) are *boxed*. *Dark gray* indicates identical amino acids and *light gray* represents conserved amino acids



alignment of 125 MADS-box genes provided a matrix of 154 protein residue sites. Figure 1 shows the scheme of exons 1–6 and the sequence regions used for phylogenetic analysis.

Maximum-likelihood (ML) analysis was carried out using the package MOLPHY version 2.3 (Adachi and Hasegawa 1996). A preliminary neighbor-joining (NJ) tree was constructed with the programs ProtML and Njdist,

using the JTT model. A local rearrangement search was then carried out to construct the ML tree, using the ProtML program. Local bootstrap probabilities for branches were estimated using resampling by the estimated log-likelihood (RELL) method (Kishino et al. 1990; Hasegawa and Kishino 1994). We also compared several competing phylogeny hypotheses using the user's tree option in ProtML, as described by Adachi and Hasegawa (1996). NJ analysis was performed using PHYLIP version 3.6a (Felsenstein 1995). Evolutionary distances were calculated with the PROTDIST program using the JTT model, and a gene tree was constructed with the NJ method (Saitou and Nei 1987) using the NEIGHBOR program. Statistical support for internal branches was estimated by bootstrap probabilities with 100 replicates using the programs SEQBOOT, PROTDIST, NEIGHBOR, and CONSENSE. Parsimony analyses were conducted using PAUP* 4.0b10 (Swofford 2001). A most parsimonious (MP) tree was generated with a heuristic search, using Tree Bisection Reconnection (TBR) and the MulTrees option. Bootstrap support with 1,000 replications was calculated by performing a heuristic search using the TBR option, but not the MulTrees option. Templeton tests were also estimated using PAUP* 4.0b10 with the options of nonparametric test (Templeton 1983).

To select the outgroup of the B-class genes, we first constructed a phylogenetic tree using 38 sequences of MIKC-type MADS-box genes from *Arabidopsis thaliana*. Although the reconstructed ML tree of *Arabidopsis* MADS-box genes showed that the *ABS/AGL32* gene (the Bs gene of *A. thaliana*) is sister to the AP3/PI clade, support was very low for the monophyly of the B/Bs clade (data not shown). It has been noted elsewhere that the inter-clade relationships within the B and Bs clade are ambiguous (Nam et al. 2003; Soltis and Soltis 2003). We then used genes of the *AGL15* and *AGL24* lineages as the outgroup in the following analyses, because these lineages have a close relationship to the B-class and Bs genes in our tree. In several published phylogenetic trees, the relationships between genes of the *AGL15* and *AGL24* lineages and the B-class genes are rather close (Alvarez-Buylla et al. 2000; Theissen et al. 2000; Johansen et al. 2002; Kofuji et al. 2003; Nam et al. 2003; Parenicova et al. 2003; Tanabe et al. 2003). Phylogenetic analysis of the *Arabidopsis* MADS-box genes showed that *AGL63* belongs to the Bs clade (Martinez-Castilla and Alvarez-Buylla 2003; Parenicova et al. 2003; Tanabe et al. 2003), but we did not use this gene because of its high divergence and its lack of motifs specific to the B/Bs genes.

The divergence times between lineages were estimated with the LINTREE program (Takezaki et al. 1995) using the linearized tree method. Yoder and Yang's likelihood method (2000) was also employed, using PAML version 3.12 (Yang 2002) with the model of local clocks, which was implemented using ML, allowing different evolutionary rates for some lineages while assuming rate constancy in others. The value for the gamma shape parameter $\alpha(1.46204)$ of the Poisson-correction (PC) gamma distance obtained using the likelihood method was estimated with the codeml program of PAML version 3.12 (Yang 2002).

The Dayhoff distance, which takes into account backward and parallel mutations, can be computed from the PC gamma distance with $\alpha = 2.25$ (Nei and Kumar 2000).

Results

Phylogenetic analysis of AP3- and PI-like MADS-box genes

Total RNA was isolated from floral buds or floral organs of basal angiosperms shown in Table 1. cDNA of these species was synthesized using MADS-box-specific degenerate primers, and sequenced. We then compared the inferred protein sequences against sequences in protein databases and selected those that were closely related to the sequences of B-class MADS-box proteins. Eight PI-like and twelve AP3-like sequences were isolated (Fig. 2A).

Relationships of PI and AP3 homologues isolated in this study to other PI and AP3 genes were inferred using maximum-likelihood, parsimony, and neighbor-joining methods (Fig. 3). Four monophyletic groups (angiosperm AP3 homologues, angiosperm PI homologues, gymnosperm AP3/PI homologues, and Bs genes) with moderately high bootstrap values were recognized with all three methods. Maximum-likelihood analysis placed the genes of *Amborella* and Nymphaeales at the base of the AP3 and PI lineages (Fig. 3A). This relationship was not rejected by either parsimony or neighbor-joining methods (Fig. 3B and C), but bootstrap support of the interior branches was low.

Several recent phylogenetic analyses have suggested that *Amborella* is sister to all other angiosperms (Mathews and Donoghue 1999; Parkinson et al. 1999; Qiu et al. 1999, 2001; Soltis et al. 1999; Savolainen et al. 2000; Zanis et al. 2002). However, the topology of the AP3 and PI lineages in this study suggests that either the Nymphaeales genes alone are sisters to those of other angiosperms, or that *Amborella* and Nymphaeales genes together form a clade that is sister to the remaining angiosperms. Therefore, we compared three hypotheses for the root of the angiosperms, using a method described previously (Fig. 4; Zanis et al. 2002). A tree-length difference test was applied to these hypotheses. As shown in Fig. 3B, the strict consensus tree of twelve MP trees (3215 steps) supports hypothesis C. We built the preliminary constraint tree for hypotheses A and B using Tree-View version 1.6.5 and performed MP analysis enforcing the topological constraints imposed by these hypotheses. MP analysis produced 74 trees with 3221 steps for hypothesis A and 1593 trees with 3217 steps for hypothesis B (Table 2). A Templeton test indicated that the difference between the three hypotheses is not significant (Table 2; Templeton 1983).

We isolated three divergent AP3 homologues from *Illicium anisatum*. These genes, designated *IaAP3-1*, *IaAP3-2*, and *IaAP3-3*, fall into two clades: one clade consisting of *IaAP3-1* and *IaAP3-3*, and another consisting of *IaAP3-2* and *KjAP3* (*Kadsura japonica*). The monophyly of these clades is strongly supported under all methods of phyloge-

Table 2. Results of Templeton test

Hypothesis	Trees of same length (<i>n</i>)	Tree length	Difference in tree lengths	<i>Z</i>	<i>P</i> ^a
C (MP tree)	12	3215	0	0	1
B	1593	3217	2	-0.277-0	0.782-1
A	74	3221	6	-0.654-0.600	0.513-0.548

^aApproximate probability of getting a more extreme test statistic under the null hypothesis of no difference (two-tailed test)

Fig. 3A-C. Phylogenetic tree of B-class MADS-box genes from angiosperms and gymnosperms. Classification follows Soltis et al. (1999) and Endress (2001). *Eudicots** indicates genes from eudicots except for the Ranunculales. *Magnoliids*** indicates genes from magnoliids except for the monocots. **A** Maximum-likelihood tree using the JTT model with local bootstrap support estimated by resampling of the estimated log-likelihood method with 1,000 replications. **B** Strict consensus tree of 12 equally parsimonious trees. *Numbers above branches* indicate bootstrap values (>25%) from 1,000 replications. **C** Neighbor-joining tree with bootstrap support from 100 replications indicated *above branches*

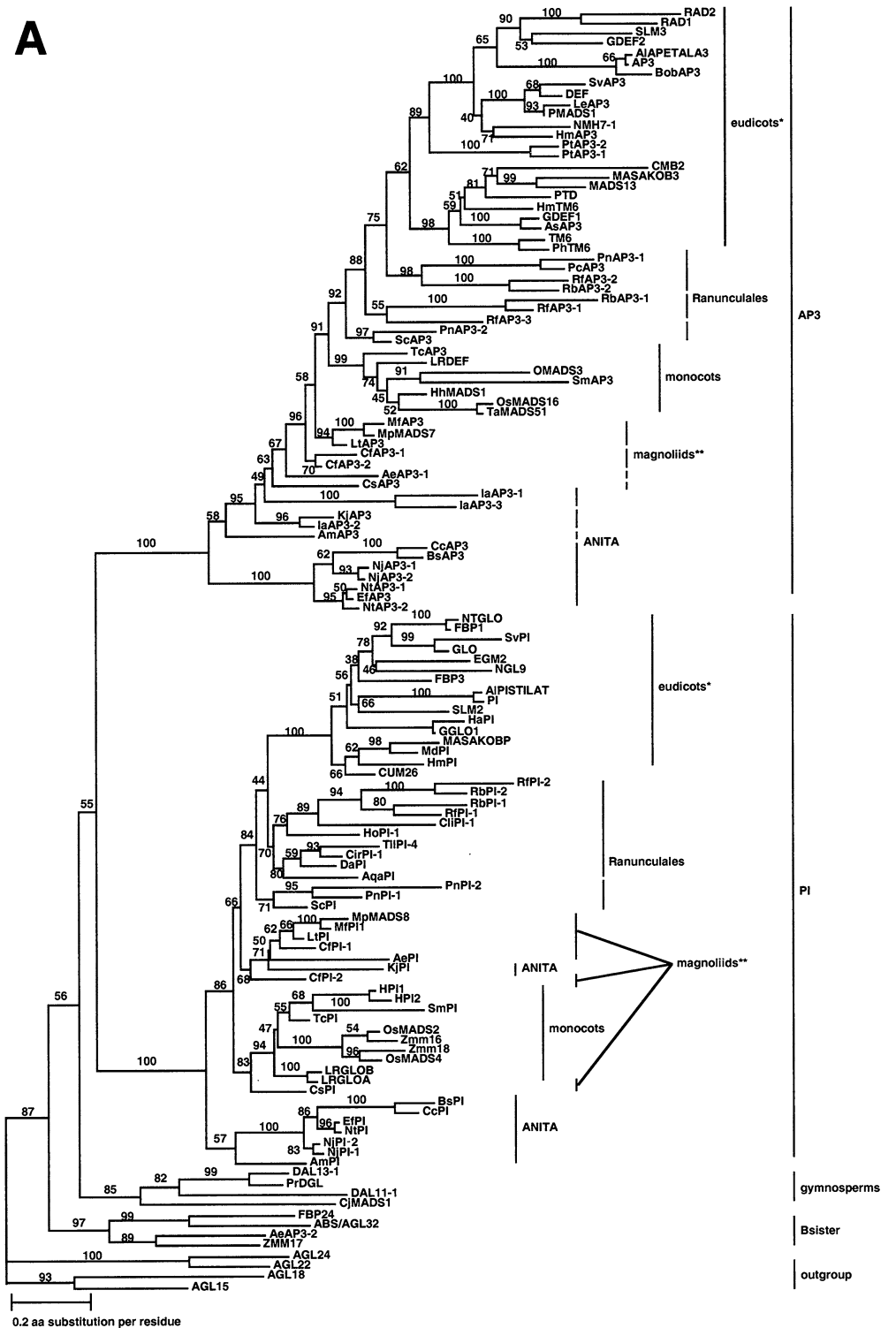
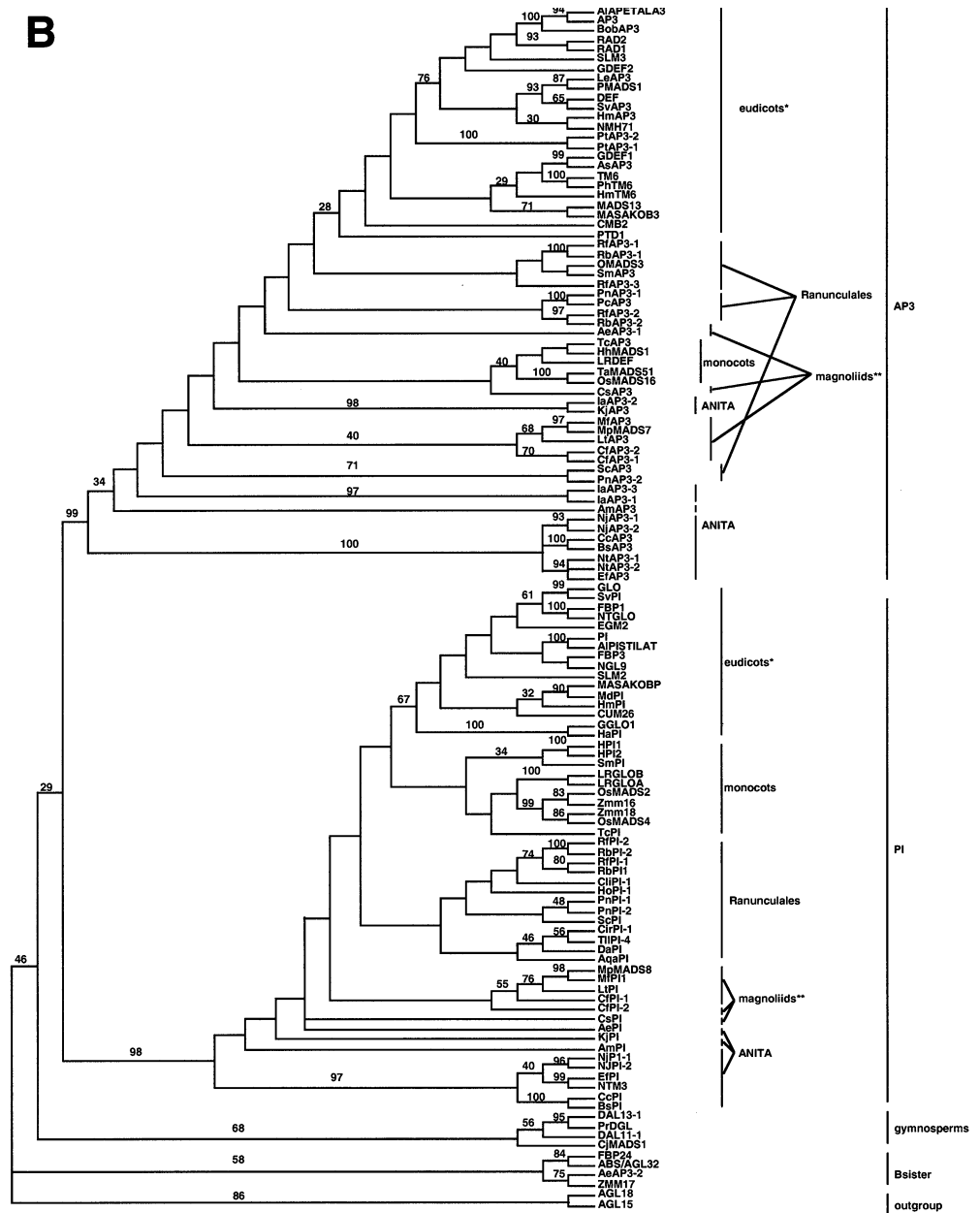


Fig. 3A–C. Continued

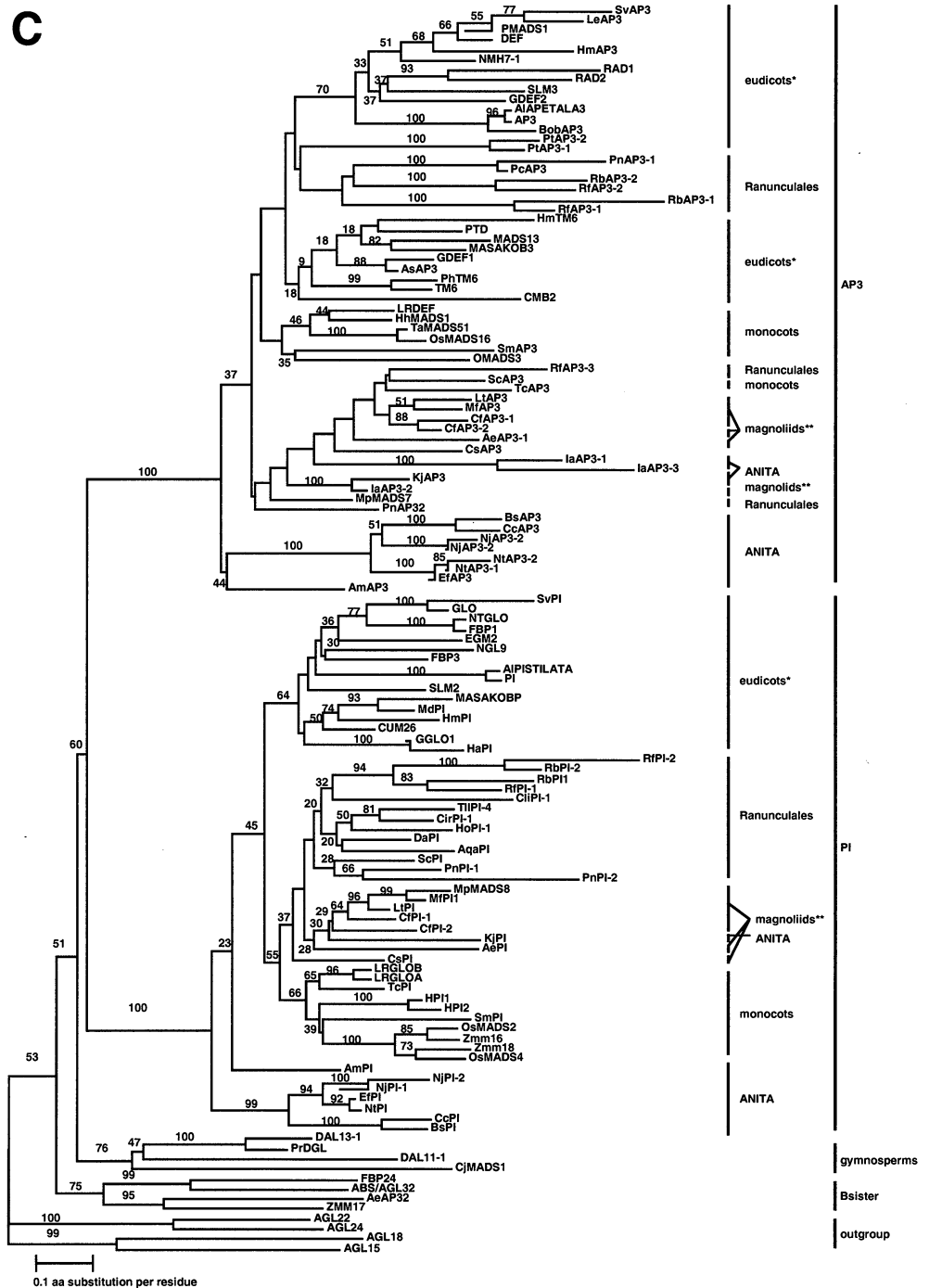


netic inference. Although only one *AP3* homologue has been isolated from *Kadsura japonica*, this finding suggests that the *IaAP3-1/IaAP3-3* and *IaAP3-2/KjAP3* lineages were established before the existence of the common ancestor of *Illicium anisatum* and *K. japonica*. It has been suggested that Austrobaileyaceae, Trimeniaceae, Illiciaceae, and Schisandraceae form a clade at the base of the angiosperms rather than *Amborella* and Nymphaeales (Mathews and Donoghue 1999; Parkinson et al. 1999; Qiu et al. 1999; Soltis et al. 1999). Our data suggest that the genes from *Illicium* (Illiciaceae) and *Kadsura* (Schisandraceae) may be sisters to the genes of the remaining angiosperms (the *AP3* lineage of the ML tree; Fig. 3A). However, the low bootstrap values mean that this result needs further testing.

Features of the isolated *AP3*-like and *PI*-like gene sequences of basal angiosperms

It is known that the genes of the *AP3* and *PI* lineages have several specific sequence features. First, in all previously reported genes of the *PI* lineage there is a deletion of four residues in the K domain. The four-residue deletion also exists in the K domain of *KjPI* (*K. japonica*), but is not found in the sequences of any of the *PI* homologues of *Amborella* or the five species of Nymphaeales (Fig. 2A). To further investigate this deletion in the *PI* homologues, the portion of the genomic DNA that contains the ORF region of the *AmPI* (*A. trichopoda*) and *KjPI* genes was sequenced. Analysis of the exon/intron organization showed that the sequence of the 4-aa site of the K domain

Fig. 3A–C. Continued



corresponds to a 12-bp deletion in exon 5. The length of exon 5 of the *KjPI* gene is 30 bp and is equal to that of the *PI* genes of *Arabidopsis thaliana*, *Antirrhinum majus*, and *Oryza sativa* (Table 3). By contrast, exon 5 sequence of *AmPI* is 42 bp long, as in the genes of the *AP3*, gymnosperm B-class, and *Bs* lineages that are sisters to the *PI* lineage.

The C-terminal region of the *AP3* and *PI* lineages also has specific sequence features. As shown in Fig. 1, the C-terminal region is encoded by exons 6 and 7 (and sometimes 8). In most MADS-box genes of known structure, exon 6 is

42 bp long (Johansen et al. 2002). In contrast, exon 6 in the *AP3* and *PI* homologues is 45 bp long. It has been suggested that the ancestral length of exon 6 of all of the MIKC-type MADS-box genes is 42 bp and that an insertion of 3 bp occurred in the ancestral genes of the *AP3* and *PI* lineages after the divergence of angiosperms and gymnosperms (Johansen et al. 2002; Winter et al. 2002). As shown in Table 2, exon 6 is 45 bp long in *AmAP3* (*A. trichopoda*), *AmPI*, and *KjPI*, as in other genes of the *AP3/PI* clade. This result supports the previous reports (Winter et al. 2002; Johansen et al. 2002).

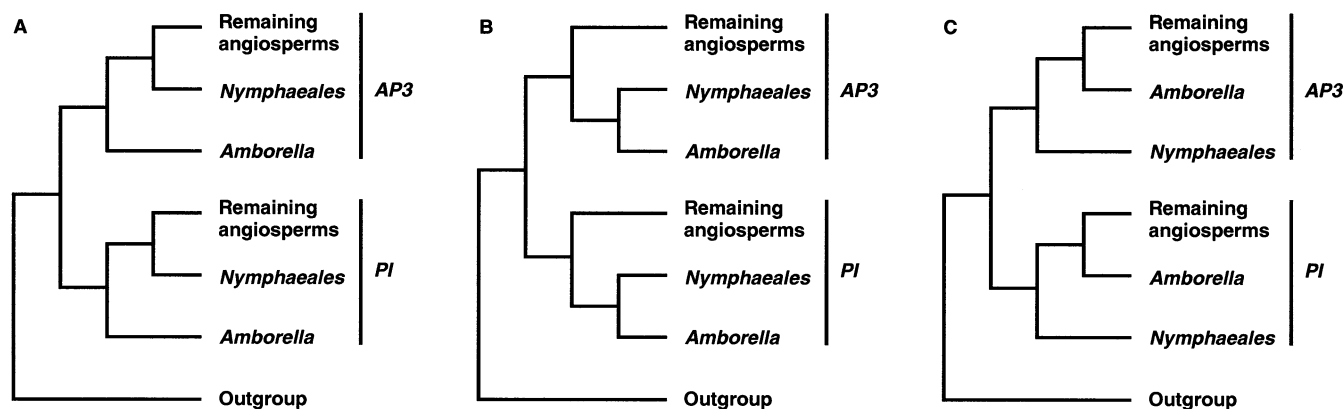


Fig. 4A–C. The three phylogenetic hypotheses for the placement of the root of the angiosperms (Zanis et al. 2002). **A** *Amborella* is sister to all remaining angiosperms. **B** *Amborella* and *Nymphaeales* are sister to all other extant angiosperms. **C** *Nymphaeales* is sister to all other angiosperms

Table 3. Structure of the exons of MADS-box genes

Clade	Gene	Exon length (bp)							
		Exon 1	2	3	4	5	6	7	8
B class									
<i>AP3</i>	<i>AP3</i>	188	67	62	100	42	45	195	
	<i>DEF</i>	188	67	62	100	42	45	180	
	<i>CMB2</i>	188	67	62	100	42	45	141	
	<i>SILKY1</i>	188	67	62	100	42	46	178	
	<i>AmAP3</i>	188	67	62	100	42	45	159	
<i>PI</i>	<i>PI</i>	255 ^a		62	100	30	45	135	
	<i>GLO</i>	188	67	62	100	30	45	156	
	<i>OsMADS2</i>	188	67	62	100	30	45	138	
	<i>KjPI</i>	188	67	62	100	30	45	141	
	<i>AmPI</i>	188	67	62	100	42	45	129	
Gymnosperms	<i>DAL13</i>	? ^b	?	62	100	42	42	?	
	<i>GGM2</i>	182	67	62	100	42	42	138	
Bsister	<i>ABS/AGL32</i>	188	70	165 ^a		42	42	252	
Out group	<i>AGL24</i>	182	82	62	100	42	42	134	19
	<i>AGL15</i>	182	61	68	100	42	42	152	166
	<i>AGL63</i>	188	73	174 ^a		63	150		

^aIntrons of *PI*, *ABS/AGL32*, and *AGL63* are lacking between exons 1 and 2 or 3 and 4

^bQuestion marks indicate unknown exon length

We found a characteristic sequence in the downstream portion of exon 6 of the *AP3* and *PI* homologues from *Amborella*. In the 5' region in exon 7 of the *AmAP3* and *AmPI* genes there is a sequence of five identical residues DEAER (Fig. 2A). However, we cannot be certain that this 5-aa sequence is synapomorphic for the B-class genes because the strong divergence of the C-terminal region complicates alignment of the sequence. By contrast, in the 3' region of the C domain of *AP3* and *PI* homologues, several motifs have been identified (Kramer et al. 1998; Kramer and Irish 2000; Becker et al. 2002). All previously reported *PI* homologues share a *PI* motif, while both a *PI* motif-derived and a *AP3* motif are shared by the *AP3* homologues so far reported. The *AP3* motif was classified into EuAP3 and PaleoAP3 motifs based on their conserved amino acid residues. We determined the sequence of the C-terminal regions of several genes isolated in this study. As

shown in Fig. 2B, the predicted protein sequences of these genes contain the motifs characteristic of *AP3* and *PI* homologues.

Estimates of divergence times

The divergence times among the genes of B-class lineage were estimated. Although several studies have determined molecular estimates of divergence times among the MADS-box genes, the B-class genes were eliminated from these analyses because evolutionary rates of *PI* and *AP3* homologues are heterogeneous (Purugganan 1997; Nam et al. 2003). To select genes for which a molecular clock is applicable, we performed the two-cluster test (Takezaki et al. 1995) with PC distance (Nei and Kumar 2000; Nam et al. 2003). The genes that had evolved significantly differently

from other genes at >3% level were eliminated. This analysis suggested that the *AP3*- and *PI*-like genes of eudicots, Ranunculales, Piperales, monocots, and Cabombaceae of Nymphaeales evolved significantly more rapidly than genes in the remaining lineages at 1 or 3% levels. Therefore, these genes were eliminated, and a linearized tree with a PC distance was constructed for the remaining genes (Fig. 5). To calibrate the time scale of the linearized tree, a divergence time between angiosperms and extant gymnosperms was used as a calibration point (node a in Fig. 5). Previous studies have revealed that extant gymnosperms (conifers, cycads, Gnetales, and Ginkgoales) are monophyletic (Hasebe et al. 1992; Winter et al. 1999; Bowe et al. 2000; Chaw et al. 2000; Donoghue and Doyle 2000), indicating that the divergence time of extant gymnosperms occurred after the angiosperms/extant gymnosperms split. We have made reference to the fossil records of seed plants for the divergence time of angiosperms and extant gymnosperms. Paleontological studies (Stewart and Rothwell 1993) showed that floras consisting of cordaites and seed ferns had reached a great diversification by the Mississippian and Pennsylvanian (Carboniferous) between 280 and 345 Ma. Cordaites, an extinct plant group, are thought to have a close relationships to conifers and may represent a base branch of seed plants (Eames 1952; Stewart and Rothwell 1993; Taylor and Hickey 1996). We therefore assumed that the divergence time between angiosperms and extant gym-

nosperms may be sometime in the Carboniferous at around 340 Ma (Sanderson and Doyle 2001). Martin et al. (1993) and Sanderson (2003) also used a calibration point of 330 Ma (mid-Carboniferous) for the angiosperm/extant gymnosperm divergence. Based on *rbcL* and *Rrn18* gene sequences and using four calibration points of the divergence time between land plants, Savard et al. (1994) obtained an estimate of the divergence time between extant angiosperms (9 species) and extant cycads and conifers (12 species) of 275–290 Ma. Goremykin et al. (1997) used 58 genes from six completely sequenced chloroplast genomes and estimated the split between *Pinus* and extant angiosperms (*Nicotiana*, *Oryza*, and *Zea*) as 348 Ma, assuming that Marchantia diverged at 450 Ma. A recent molecular study by Soltis et al. (2002) that used the assumed age of 125 Ma of the earliest angiosperms as a calibration point yielded an estimate of 343.7 Ma as the divergence time between extant angiosperms (*Austrobailea* and *Chloranthus*) and extant gymnosperms (*Cycas*, *Ginkgo*, *Gnetum*, and *Pinus*) using parsimony analysis and a combined data set of four genes.

Table 4 shows the results, including time estimates obtained using PC, PC gamma, and Dayhoff distances. When the PC distance was used, the divergence time between the *AP3* and *PI* lineages (node b in Fig. 5) was estimated to be 301 Ma. The divergence time was estimated at 293 Ma from the Dayhoff distance and at 287 Ma from the PC gamma distance. Yoder and Yang's method (2000) with a local clock model was also used, including all of the angiosperm sequences. The local clock model allows different evolutionary rates for some lineages. This method estimated the divergence time between the *AP3* and *PI* lineages as 316 Ma.

We also estimated the age of the most recent common ancestor of all living angiosperms (Table 4). Because the *AP3* and *PI* lineages are in a clade of the angiosperm B-class genes, two results for the first divergence time of the extant angiosperms (nodes c and d in Fig. 5) were obtained from a linearized tree. Using the linearized tree method with a PC distance gave an age of extant angiosperms of 208 Ma from the *AP3* lineage and 170 Ma from the *PI* lineage. Analyses with the Dayhoff distance and the PC gamma distance gave smaller estimates than that obtained when using PC distance. Estimates generated using Dayhoff distance were 170 Ma for the *AP3* lineage and 151 Ma for the *PI* lineage, and estimates using PC gamma distance were 161 Ma for the *AP3* lineage and 145 Ma for the *PI* lineage. The likelihood method gave almost the same estimates as the linearized tree method: 196 Ma for the *AP3* lineage and 157 Ma for the *PI* lineage.

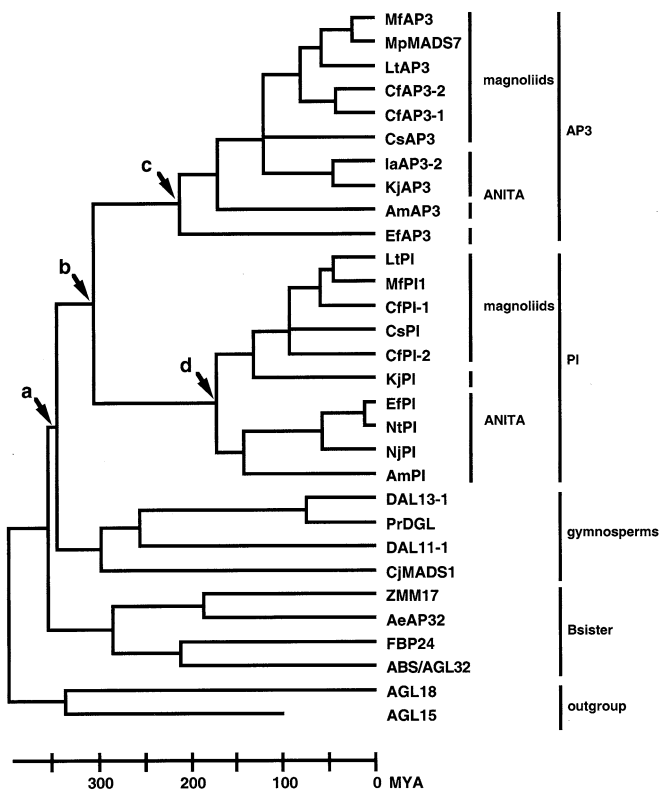


Fig. 5. Linearized tree used for estimation of divergence times. The tree topology and time scale are based on the result with the PC distance (see text). Letters at nodes indicate divergence points of major clades

Discussion

Phylogenetic relationship of basal angiosperms

We have reported here the cloning and characterization of seventeen new genes of the *AP3* and *PI* lineages from basal

Table 4. Estimates of divergence times^a (mya) of B-class MADS-box genes

Node in Fig. 5	Linearized tree method			ML method
	PC distance	Dayhoff distance	PC gamma distance ^b	WAG model
	Alpha ∞	2.25	1.46204	1.03829
(b) Divergence between <i>AP3</i> and <i>PI</i> clades	301 ± 35	293 ± 42	287 ± 47	316 ± 17
(c) First divergence of extant angiosperms in <i>AP3</i> clade	208 ± 26	170 ± 22	161 ± 22	196 ± 13
(d) First divergence of extant angiosperms in <i>PI</i> clade	170 ± 20	151 ± 20	145 ± 21	157 ± 10

^a The gymnosperm/angiosperm split (ca. 340 Ma) was used for calibrating the time scale

^b Times were computed using 29 genes that did not include *AmPI* because of the unclear tree topology

angiosperms. In recent decades, several molecular phylogenetic studies have been performed to deduce the relationship of basal angiosperms. Early phylogenetic analyses using *rbcL* sequences led to the proposal that *Ceratophyllum* may be sister to the remaining angiosperms (Chase et al. 1993). Subsequent studies have identified species of the ANITA grade as the earliest diverging lineages of angiosperms. Analyses of *atpB* (Savolainen et al. 2000); phytochrome genes (Mathews and Donoghue 1999); a combined 18S, *rbcL*, and *atpB* data set (Soltis et al. 1999); a data set of five mitochondrial, plastid, and nuclear genes (Parkinson et al. 1999; Qiu et al. 1999); and noncoding plastid *trnT-trnF* sequences (Borsch et al. 2003) have suggested that *Amborella* is the sister of all other angiosperms, followed by Nymphaeales, and then a clade consisting of *Austrobaileya*, Trimeniaceae, Illiciales, and Schisandraceae. Other analyses link *Amborella* with Nymphaeales or reverse these two taxa, with both lines basal to other angiosperms (Barkman et al. 2000; Graham and Olmstead 2000; Qiu et al. 2000). However, another recent study that used the complete chloroplast genome sequences of thirteen land plants suggested that the monocot lineage diverged earlier than the other groups of angiosperms, including *Amborella* (Goremykin et al. 2003). In the present study, we used *AP3*-like and *PI*-like MADS-box genes, which play an important role in floral development. Molecular phylogenetic analysis of floral homeotic genes with the ML method revealed *Amborella* and Nymphaeales as basalmost lineages in the angiosperms (Fig. 3A). Our data, however, do not allow us to conclusively determine the order of the branching relationship among the putative B-class genes of *Amborella*, Nymphaeales, and the remaining angiosperms (Table 2).

The gene structure of the *PI* homologues of basal angiosperms supports the idea that *Amborella* and Nymphaeales are basalmost lineages in angiosperms. Among the protein sequences deduced from the cDNAs of the basal angiosperm genes, the *PI* homologues of *Amborella* and Nymphaeales do not have the 4-aa deletion that is characteristic of other angiosperm *PI* homologues. The most parsimonious explanation for these results is that the 4-aa deletion was established in the *PI* lineage after the separation of *Amborella* and Nymphaeales from other angiosperms (Fig. 6). The 4-aa deletion of *PI* homologues is in the upstream part of the third putative amphipathic α -helix in the K domain, called K3 (Yang et al. 2003). It has

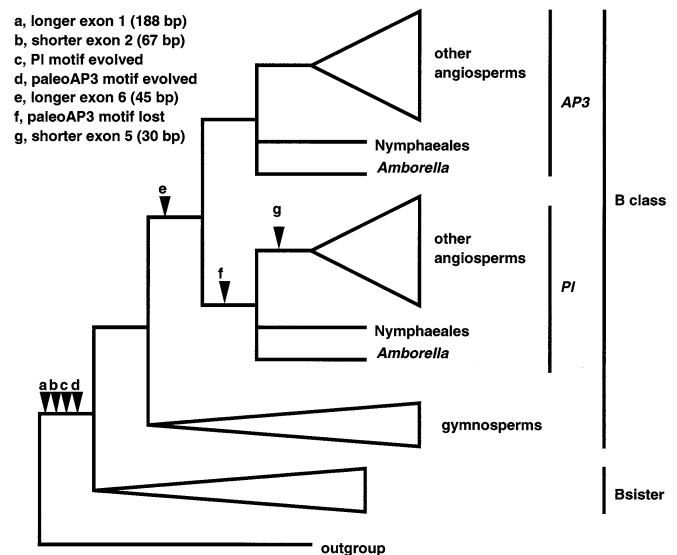


Fig. 6. Phylogeny of B-class genes. A highly simplified phylogenetic tree of B-class and Bs genes is shown, with *AP3* and *PI* symbolizing gene clades. Arrows with letters at nodes define putative apomorphies, as described in the figure

been reported that the K3 region is more important for the interaction between the *Arabidopsis AP3* and *PI* proteins than either the I region or the C-terminal region (Yang et al. 2003). It is likely that the deletion in the ancient *PI* gene modified the function and protein-protein interaction characteristics of this protein.

Previous studies have shown that the molecular evolution of the B-class genes of some species has been rapid (Purugganan 1997; Barrier et al. 2001; Nam et al. 2003). The two-cluster test used in the present study revealed that rapid evolution might have occurred in several lineages of the angiosperm *AP3* and *PI* homologues. Barrier et al. (2001) have suggested that the accelerated evolutionary rates of the MADS-box genes may accompany rapid morphological diversification in adaptive radiations of Hawaiian silversword alliance. Lamb and Irish (2003) also have suggested that changes in protein sequence and function of B-class MADS-box proteins influence floral morphologies among the angiosperms. To understand what changes in the sequence of *AP3* and *PI* genes may underlie morphological diversification, investigation of the influence of rapid molecular evolution and the 4-aa deletion on the function of these

genes will be important. Further analysis of *AP3* and *PI* genes may thus contribute to our understanding not only of the phylogenetic relationship of angiosperms but also of the morphological diversification of flowering plants.

Estimates of divergence times

Using the linearized tree and likelihood methods, the divergence time between the *AP3* and *PI* lineages was estimated as approximately 34–53 million years after the assumed divergence between angiosperms and extant gymnosperms at 340 Ma (Table 4). However, given the error margin and variation of our various estimations (Table 4), this result does not exclude the possibility that the divergence between the *AP3* and *PI* lineages occurred immediately after the angiosperms/extant gymnosperms split (Winter et al. 2002). The divergence between the *AP3* and *PI* lineages appears to have occurred somewhere between immediately after to several tens of millions of years after the split between angiosperms and extant gymnosperms.

We also estimated the age of the most recent common ancestor of the extant angiosperms. At present, the oldest definite angiosperm fossils are pollen grains of lower Cretaceous age, 132–141 million years old (Hughes 1994; Brenner 1996). Molecular-based estimates of the divergence times between the lineages of extant angiosperms range widely. The first molecular study by Ramshow et al. (1972) obtained an estimate of 350–420 Ma for the higher angiosperms/buckwheat divergence based on the cytochrome c protein, calibrated with the bird/mammal split at 280 Ma. Several subsequent analyses obtained basically an estimate of the divergence time between monocots and eudicots. Using the *gapC* gene, calibrated with an animal fossil and the presumed divergence of plants, animals, and fungi at 1000 Ma, Martin et al. (1989) dated the split between monocots (two species) and dicots (*Magnolia* and six eudicots) as 319 Ma. These studies gave far older ages for the angiosperms than the oldest fossil record. Subsequent studies used the divergence time between land plants for

calibration and estimated the divergence time between monocots and eudicots as 200 ± 40 Ma (Wolfe et al. 1989), 230–350 Ma (Brandl et al. 1992), 300 Ma (Martin et al. 1993), 200 Ma (Laroche et al. 1995), 160 Ma (Goremykin et al. 1997), and 239 Ma (Nam et al. 2003). A study by Sanderson (1997) with a novel method, nonparametric rate shift (NPRS), used *rbcL* genes of 36 land plants that include many basal angiosperms. He dated the age of the most recent common ancestor of extant angiosperms as 165 Ma, calibrated with the divergence of *Matchantia* and remaining land plants at 450 Ma. Using almost the same data set of 31 plants, Thorne et al. (1998) applied a model-based Bayesian approach and estimated that the angiosperm root node is 51% as old as the most recent common ancestor of vascular plants (i.e., ~200 Ma; Sanderson and Doyle 2001). Subsequently, Sanderson and Doyle (2001) presented a series of analyses that illustrated the effect of various factors of age estimation using different tree topologies (Mathews and Donoghue 1999; Parkinson et al. 1999; Qiu et al. 1999; Soltis et al. 1999; Savolainen et al. 2000). They obtained estimates of ~140–190 Ma for the base divergence of extant angiosperms.

The present estimate used the *AP3*-like and *PI*-like MADS-box genes of basal angiosperms and the linearized tree and likelihood methods (Takezaki et al. 1995; Yoder and Yang 2000), calibrated with an approximate angiosperm/extant gymnosperm split at 340 Ma. Our analysis suggests that the extant angiosperm divergence began at about 145–208 Ma (upper Triassic to Jurassic). This estimate coincides with the estimates obtained by Sanderson (1997), Thorne et al. (1998), and Sanderson and Doyle (2001). The common feature of our analysis and these previous analyses is that all have used a dense sample of basal angiosperms.

Acknowledgements We are grateful to Prof. Dr. Hirohisa Kishino, Dr. Tomoaki Nishiyama, Miss Yumi Miyamoto, and two anonymous reviewers for their useful and kind comments. This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Appendix. The sequences of cDNAs isolated in this study are shown in Table 5.

Table 5. The MADS-box genes used in this study

Lineage	Gene ^a	Category	Order	Family	Species	Accession No.
B class						
<i>AP3</i>	<i>SvAP3</i>	Euastrid I	Lamiales	Oleaceae	<i>Syringa vulgaris</i>	AAC42584
	<i>DEF</i>	Euastrid I	Lamiales	Scrophulariaceae	<i>Antirrhinum majus</i>	CAA44629
	<i>LeAP3</i>	Euastrid I	Solanales	Solanaceae	<i>Lycopersicon esculentum</i>	AAC42583
	<i>PMADSI</i>	Euastrid I	Solanales	Solanaceae	<i>Petunia × hybrida</i>	Q07472
	<i>TM6</i>	Euastrid I	Solanales	Solanaceae	<i>Lycopersicon esculentum</i>	S23731
	<i>PhTM6</i>	Euastrid I	Solanales	Solanaceae	<i>Petunia × hybrida</i>	AAF73933
	<i>GDEF1</i>	Eusatrifid II	Asterales	Asteraceae	<i>Gerbera hybrid</i>	CAA08802
	<i>GDEF2</i>	Eusatrifid II	Asterales	Asteraceae	<i>Gerbera hybrid</i>	CAA08803
	<i>AsAP3</i>	Eusatrifid II	Asterales	Asteraceae	<i>Anisocarpus scabridus</i>	AAD51221
	<i>NMH7-1</i>	Eurosid I	Fabales	Fabaceae	<i>Medicago sativa</i>	AAB48660
	<i>MADS13</i>	Eurosid I	Rosales	Rosaceae	<i>Malus × domestica</i>	CAC80856
	<i>MASAKO B3</i>	Eurosid I	Rosales	Rosaceae	<i>Rosa rugosa</i>	BAB63261

Table 5. Continued

Lineage	Gene ^a	Category	Order	Family	Species	Accession No.
	<i>PTD</i>	Eurosid I	Malpighiales	Saliceaceae	<i>Populus balsamifera</i>	AAC13695
	<i>AP3</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	P35632
	<i>ALAPETALA3</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	AAF25590
	<i>BoBAP3</i>	Eurosid II	Brassicales	Brassicaceae	<i>Brassica oleracea</i>	AAB17139
	<i>HmAP3</i>	Astrids	Cornales	Hydrangeaceae	<i>Hydrangea macrophylla</i>	AAF73931
	<i>HmTM6</i>	Astrids	Cornales	Hydrangeaceae	<i>Hydrangea macrophylla</i>	AAF73932
	<i>RaD2</i>	Eudicots	Caryophyllales	Polygonaceae	<i>Rumex acetosa</i>	AAA80305
	<i>RaD1</i>	Eudicots	Caryophyllales	Polygonaceae	<i>Rumex acetosa</i>	AAA80304
	<i>SLM3</i>	Eudicots	Caryophyllales	Caryophyllaceae	<i>Silene latifolia</i>	CAA56657
	<i>CMB2</i>	Eudicots	Caryophyllales	Caryophyllaceae	<i>Dianthus caryophyllus</i>	T10715
	<i>PtAP3-2</i>	Eudicots	Proteales	Buxaceae	<i>Pachysandra terminalis</i>	AAC42586
	<i>PtAP3-1</i>	Eudicots	Proteales	Buxaceae	<i>Pachysandra terminalis</i>	AAC42585
	<i>ScAP3</i>	Basal eudicots	Ranunculales	Papaveraceae	<i>Sanguinaria canadensis</i>	AAD31696
	<i>PnAP3-2</i>	Basal eudicots	Ranunculales	Papaveraceae	<i>Papaver nudicaule</i>	AAC42589
	<i>PcAP3</i>	Basal eudicots	Ranunculales	Papaveraceae	<i>Papaver californicum</i>	AAC42587
	<i>PnAP3-1</i>	Basal eudicots	Ranunculales	Papaveraceae	<i>Papaver nudicaule</i>	AAC42588
	<i>RbAP3-2</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus bulbosus</i>	AAD31697
	<i>RbAP3-1</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus bulbosus</i>	AAC42591
	<i>RfAP3-1</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus ficaria</i>	AAC42569
	<i>RfAP3-2</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus ficaria</i>	AAD31698
	<i>RfAP3-3</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus ficaria</i>	AAO26531
	<i>CfAP3-1</i>	Magnoliids	Laurales	Calycanthaceae	<i>Calycanthus floridus</i>	AAF73928
	<i>CfAP3-2</i>	Magnoliids	Laurales	Calycanthaceae	<i>Calycanthus floridus</i>	AAF73929
	<i>MpMADS7</i>	Magnoliids	Magnoliales	Magnoliaceae	<i>Magnolia praecocissima</i>	BAB70742
	<i>MfAP3</i>	Magnoliids	Magnoliales	Magnoliaceae	<i>Michelia figo</i>	AAC42592
	<i>LtAP3</i>	Magnoliids	Magnoliales	Magnoliaceae	<i>Liriodendron tulipifera</i>	AAC42593
	<i>AeAP3-1</i>	Magnoliids	Piperales	Aristolochiaceae	<i>Asarum europaeum</i>	AAF73927
	<i>CsAP3</i>	Magnoliids	Chloranthales	Chloranthaceae	<i>Chloranthus spicatus</i>	AAF73930
	<i>OsMADS16</i>	Magnoliids	Poales	Poaceae	<i>Oryza sativa</i>	AAD19872
	<i>TaMADS51</i>	Magnoliids	Poales	Poaceae	<i>Triticum aestivum</i>	BAA33459
	<i>OMADS3</i>	Magnoliids	Asparagales	Orchidaceae	<i>Oncidium</i> cv.	AAO45824
	<i>LRDEF</i>	Magnoliids	Liliales	Liliaceae	<i>Lilium regale</i>	BAB91550
	<i>TcAP3</i>	Magnoliids	Dioscoreales	Dioscoreaceae	<i>Tacca chantieri</i>	AAF73935
	<i>Hh-MADS1</i>	Magnoliids	Asparagales	Hemerocallidaceae	<i>Hemerocallis</i> cv.	AAG35773
	<i>SmAP3</i>	Magnoliids	Alismatales	Alismataceae	<i>Sagittaria montevidensis</i>	AAF73934
	<i>KjAP3</i>	ANITA	Austrobaileyales	Schisandraceae	<i>Kadsura japonica</i>	AB154848
	<i>IaAp3-1</i>	ANITA	Austrobaileyales	Illiciaceae	<i>Illicium anisatum</i>	AB154843
	<i>IaAp3-2</i>	ANITA	Austrobaileyales	Illiciaceae	<i>Illicium anisatum</i>	AB154844
	<i>IaAp3-3</i>	ANITA	Austrobaileyales	Illiciaceae	<i>Illicium anisatum</i>	AB154846
	<i>BsAP3</i>	ANITA	Nymphaeales	Cabombaceae	<i>Brasenia schreberi</i>	AB158355
	<i>CcAP3</i>	ANITA	Nymphaeales	Cabombaceae	<i>Cabomba caroliniana</i>	AB158353
	<i>EfAP3</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Euryale ferox</i>	AB158349
	<i>NjAP3-1</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Nuphar japonicum</i>	AB158357
	<i>NjAP3-2</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Nuphar japonicum</i>	AB158358
	<i>NtAP3-1</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Nymphaea tetragona</i>	AB158351
	<i>NtAp3-2</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Nymphaea tetragona</i>	AB166793
	<i>AmAP3</i>	ANITA	Amborellales	Amborellaceae	<i>Amborella trichopoda</i>	AB154845
<i>PI</i>	<i>FBP1</i>	Euastrid I	Solanales	Solanaceae	<i>Petunia × hybrida</i>	Q03488
	<i>NTGLO</i>	Euastrid I	Solanales	Solanaceae	<i>Nicotiana tabacum</i>	Q03416
	<i>FBP3</i>	Euastrid I	Solanales	Solanaceae	<i>Petunia × hybrida</i>	S60288
	<i>GLO</i>	Euastrid I	Lamiales	Scrophulariaceae	<i>Antirrhinum majus</i>	Q03378
	<i>SvPI</i>	Euastrid I	Lamiales	Oleaceae	<i>Syringa vulgaris</i>	AAC42576
	<i>GGLO1</i>	Euastrid II	Asterales	Asteraceae	<i>Gerbera</i> cv.	CAA08804
	<i>HaPI</i>	Eusatrid II	Asterales	Asteraceae	<i>Helianthus annuus</i>	AAN47199
	<i>EGM2</i>	Eusatrid II	Myrtales	Myrtaceae	<i>Eucalyptus grandis</i>	AAC78283
	<i>NGL9</i>	Eurosid I	Fabales	Fabaceae	<i>Medicago sativa</i>	AAK77938
	<i>MASAKO BP</i>	Eurosid I	Rosales	Rosaceae	<i>Rosa rugosa</i>	BAB63261
	<i>MdPI</i>	Eurosid I	Rosales	Rosaceae	<i>Malus × domestica</i>	CAC28021
	<i>CUM26</i>	Eurosid I	Cucurbitales	Cucurbitaceae	<i>Cucumis sativus</i>	AAD02250
	<i>PI</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	P48007
	<i>APISTILLATA</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	AAF25591
	<i>HmPI</i>	Astrids	Cornales	Hydrangeaceae	<i>Hydrangea macrophylla</i>	AAF73940
	<i>SLM2</i>	Eudicots	Caryophyllales	Caryophyllaceae	<i>Silene latifolia</i>	CAA56656
	<i>DaPI-1</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Delphinium ajacis</i>	AAC42577
	<i>RbPI-1</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus bulbosus</i>	AAC42574
	<i>RbPI-2</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus bulbosus</i>	AAC42575

Table 5. Continued

Lineage	Gene ^a	Category	Order	Family	Species	Accession No.
	<i>RfPI-1</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus ficaria</i>	AAC42573
	<i>RfPI-2</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Ranunculus ficaria</i>	AAD31700
	<i>TlPI-4</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Trollius laxus</i>	AAO26550
	<i>HoPI-1</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Helleborus orientalis</i>	AAO26526
	<i>ClPI-1</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Clematis integrifolia</i>	AAO26520
	<i>CirPI-1</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Cimicifuga racemosa</i>	AAO26513
	<i>AqaPI</i>	Basal eudicots	Ranunculales	Ranunculaceae	<i>Aquilegia alpina</i>	AAO26500
	<i>ScPI</i>	Basal eudicots	Ranunculales	Papaveraceae	<i>Sanguinaria canadensis</i>	AAD31699
	<i>PnPI-1</i>	Basal eudicots	Ranunculales	Papaveraceae	<i>Papaver nudicaule</i>	AAC42570
	<i>PnPI-2</i>	Basal eudicots	Ranunculales	Papaveraceae	<i>Papaver nudicaule</i>	AAC42571
	<i>MpMADS8</i>	Magnoliids	Magnoliales	Magnoliaceae	<i>Magnolia praecocissima</i>	BAB70743
	<i>CfPI-1</i>	Magnoliids	Laurales	Calycanthaceae	<i>Calycanthus floridus</i>	AAF73937
	<i>CfPI-2</i>	Magnoliids	Laurales	Calycanthaceae	<i>Calycanthus floridus</i>	AAF73938
	<i>MfPI-1</i>	Magnoliids	Magnoliales	Magnoliaceae	<i>Michelia figo</i>	AAC42578
	<i>LiPI-1</i>	Magnoliids	Magnoliales	Magnoliaceae	<i>Liriodendron tulipifera</i>	AAC42579
	<i>AePI</i>	Magnoliids	Piperales	Aristolochiaceae	<i>Asarum europaeum</i>	AAF73936
	<i>CsPI</i>	Magnoliids	Chloranthales	Chloranthaceae	<i>Chloranthus spicatus</i>	AAF73939
	<i>OsMADS2</i>	Magnoliids	Poales	Poaceae	<i>Oryza sativa</i>	T03894
	<i>ZMM16</i>	Magnoliids	Poales	Poaceae	<i>Zea mays</i>	CAC33848
	<i>ZMM18</i>	Magnoliids	Poales	Poaceae	<i>Zea mays</i>	CAC33849
	<i>OsMADS4</i>	Magnoliids	Poales	Poaceae	<i>Oryza sativa</i>	T03902
	<i>LRGLOA</i>	Magnoliids	Liliales	Liliaceae	<i>Lilium regale</i>	BAB91551
	<i>LRGLOB</i>	Magnoliids	Liliales	Liliaceae	<i>Lilium regale</i>	BAB91552
	<i>TcPI</i>	Magnoliids	Dioscoreales	Dioscoreaceae	<i>Tacca chantieri</i>	AAF73942
	<i>HP11</i>	Magnoliids	Asparagales	Hyacinthaceae	<i>Hyacinthus orientalis</i>	AAD22493
	<i>HP12</i>	Magnoliids	Asparagales	Hyacinthaceae	<i>Hyacinthus orientalis</i>	AAD22494
	<i>SmPI</i>	Magnoliids	Alismatales	Alismataceae	<i>Sagittaria montevidensis</i>	AAF73941
	<i>KjPI</i>	ANITA	Austrobaileyales	Schisandraceae	<i>Kadsura japonica</i>	AB154849
	<i>BsPI</i>	ANITA	Nymphaeales	Cabombaceae	<i>Brasenia schreberi</i>	AB158356
	<i>CcPI</i>	ANITA	Nymphaeales	Cabombaceae	<i>Cabomba caroliniana</i>	AB158354
	<i>EjPI</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Euryale ferox</i>	AB158350
	<i>NjPI-1</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Nuphar japonicum</i>	AB158359
	<i>NjPI-2</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Nuphar japonicum</i>	AB158360
	<i>NiPI</i>	ANITA	Nymphaeales	Nymphaeaceae	<i>Nymphaea tetragona</i>	AB158352
	<i>AmPI</i>	ANITA	Amborellales	Amborellaceae	<i>Amborella trichopoda</i>	AB154841
Gymnosperm	<i>CJMADS1</i>	Gymnosperms	Coniferales	Cupressaceae	<i>Cryptomeria japonica</i>	AAL05440
	<i>DAL13-1</i>	Gymnosperms	Coniferales	Pinaceae	<i>Picea abies</i>	AAF18376
	<i>PrDGL</i>	Gymnosperms	Coniferales	Pinaceae	<i>Pinus radiata</i>	AAF28863
	<i>DAL11-1</i>	Gymnosperms	Coniferales	Pinaceae	<i>Picea abies</i>	AAF18372
	<i>GGM2</i>	Gymnosperms	Gnetales	Gnetaceae	<i>Gnetum gnemon</i>	CAD18858
Bsister	<i>FBP24</i>	Euastrid I	Solanales	Solanaceae	<i>Petunia × hybrida</i>	AAK21255
	<i>ABS/AGL32</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	CAC85664
	<i>AeAP3-2</i>	Magnoliids	Piperales	Aristolochiaceae	<i>Asarum europaeum</i>	AAF73927
	<i>ZMM17</i>	Magnoliids	Poales	Poaceae	<i>Zea mays</i>	CAC81053
Outgroup	<i>AGL15</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	Q38847
	<i>AGL18</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	CAB68145
	<i>AGL22</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	AAD22365
	<i>AGL24</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	AAC6313
	<i>AGL63</i>	Eurosid II	Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	AAD21695

^a Gene names were simplified to make the paper understandable to a wide audience. Categories follow Soltis et al. (1999) and Endress (2001)

References

- Adachi J, Hasegawa M (1996) MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. *Comput Sci Monogr* 28:1–150
- Alvarez-Buylla ER, Liljegren SJ, Pelaz S, Gold SE, Burgeff C, Ditta GS, Vergara-Silva F, Yanofsky MF (2000) MADS gene evolution beyond flowers, expression in pollen, endosperm, guard cells, roots, and trichomes. *Plant J* 24:457–466
- Angiosperm Phylogeny Group (APG) II (2003) An update of angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. *Bot J Linn Soc* 141:399–436
- Barkman TJ, Chenery G, McNeal JR, Lyons-Weiler J, Ellisens W, Moore G, Wolfe AD, dePamphilis CW (2000) Independent and combined analyses of sequences from all three genomic compartments converge on the root of flowering plant phylogeny. *Proc Natl Acad Sci USA* 97:13166–13171
- Barrier M, Robichaux RH, Purugganan MD (2001) Accelerated regulatory gene evolution in an adaptive radiation. *Proc Natl Acad Sci USA* 98:10208–10213
- Becker A, Theissen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol* 29:464–489
- Becker A, Winter KU, Meyer B, Saedler H, Theissen G (2000) MADS gene diversity in seed plants 300 million years ago. *Mol Biol Evol* 17:1425–1434

- Becker A, Kaufmann K, Freialdenhoven A, Vincent C, Li MA, Saedler H, Theissen G (2002) A novel MADS-box gene subfamily with a sister-group relationship to class B floral homeotic genes. *Mol Genet Genomics* 266:942–950
- Borsch T, Hilu KW, Quandt D, Wilde V, Neinhuis C, Barthlott W (2003) Noncoding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *J Evol Biol* 16:558–576
- Bowe LM, Coat G, dePamphilis CW (2000) Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales closest relatives are conifers. *Proc Natl Acad Sci USA* 97:4092–4097
- Bowman JL (1997) Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. *J Biosci* 22:515–527
- Bowman J, Smith DR, Meyerowitz EM (1991) Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development* 112:1–20
- Brandl R, Mann W, Sprinzl M (1992) Estimation of the monocot–dicot age through tRNA sequences from the chloroplast. *Proc R Soc Lond Ser B* 249:13–17
- Brenner GJ (1996) Evidence for the earliest stage of angiosperm pollen evolution: a paleoequatorial section from Israel. In: Taylor DW, Hickey LJ (eds) Flowering plant origin, evolution and phylogeny. Chapman and Hall, New York, pp 91–115
- Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y-L, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Systsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedren M, Gaut BS, Jansen RK, Kim K-J, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang Q-Y, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Learn GH, Graham SW Jr, Barrett SCH, Dayanandan S, Albert VA (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann MO Bot Gard* 80:528–580
- Chaw SM, Parkinson CL, Cheng Y, Vincent TM, Palmer JD (2000) Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proc Natl Acad Sci USA* 97:4086–4091
- Crane PR, Friis EM, Pedersen KR (1995) The origin and early diversification of angiosperms. *Nature* 374:27–33
- Coen ES, Meyerowitz EM (1991) The war of the whorls: genetic interactions controlling flower development. *Nature* 353:31–37
- Donoghue MJ, Doyle JA (2000) Demise of the anthophyte hypothesis? *Curr Biol* 10:R106–R109
- Eames AJ (1952) Relationships of Ephedrales. *Phytomorphology* 2:79–100
- Endress PK (2001) The flowers in extant basal angiosperms and inferences on ancestral flowers. *Int J Plant Sci* 162:1111–1140
- Felsenstein J (1995) PHYLIP (Phylogeny Inference Package) version 3.572c. Department of Genetics, University of Washington, Seattle
- Fukui M, Futamura N, Mukai Y, Wang Y, Nagao A, Shinohara K (2001) Ancestral MADS box genes in Sugi, *Cryptomeria japonica*, D. Don (Taxodiaceae), homologous to the B function genes in angiosperms. *Plant Cell Physiol* 42:566–575
- Goremykin V, Hansmann S, Martin W (1997) Evolutionary analysis of 58 proteins encoded in six completely sequenced chloroplast genomes: revised molecular estimates of two seed plant divergence times. *Plant Syst Evol* 206:337–351
- Goremykin VV, Hirsch-Ernst KI, Wolf S, Hellwig FH (2003) Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that amborella is not a basal angiosperm. *Mol Biol Evol* 20:1499–1505
- Goto K, Meyerowitz EM (1994) Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA*. *Genes Dev* 8:1548–1560
- Graham SW, Olmstead RG (2000) Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. *Am J Bot* 87:1712–1730
- Hasebe M, Kofuji R, Ito M, Kato M, Iwatsuki K, Ueda K (1992) Phylogeny of gymnosperms inferred from *rbcL* gene sequences. *Bot Mag Tokyo* 105:673–679
- Hasebe M, Wen CK, Kato M, Banks JA (1998) Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. *Proc Natl Acad Sci USA* 95:6222–6227
- Hasegawa M, Kishino H (1994) Accuracies of the simple methods for estimating the bootstrap probability of a maximum-likelihood tree. *Mol Biol Evol* 11:142–145
- Hughes NF (1994) The enigma of angiosperm origins. Cambridge University Press, Cambridge
- Irish VF (2003) The evolution of floral homeotic gene function. *BioEssays* 25:637–646
- Jack T, Brockman LL, Meyerowitz EM (1992) The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* 68:683–697
- Johansen B, Pedersen LB, Skipper M, Frederiksen S (2002) MADS-box gene evolution – structure and transcription patterns. *Mol Phylog Evol* 23:458–480
- Judd W, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2002) Plant systematics: a phylogenetic approach, 2nd edn. Sinauer, Sunderland, MA
- Kishino H, Miyata T, Hasegawa M (1990) Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J Mol Evol* 30:151–160
- Kofuji R, Sumikawa N, Yamasaki M, Kondo K, Ueda K, Ito M, Hasebe M (2003) Evolution and divergence of MADS-box gene family based on genome wide expression analyses. *Mol Biol Evol* 20:1963–1977
- Kramer EM, Irish VF (2000) Evolution of the petal and stamen development programs: evidence from comparative studies of the lower eudicots and basal angiosperms. *Int J Plant Sci* 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 E, Stilio VSD, Schluter PM (2003) Complex patterns of gene duplication in the *APETALA3* and *PISTILLATA* lineages of the Ranunculaceae. *Int J Plant Sci* 164:1–11
- Lamb RS, Irish VF (2003) Functional divergence within the *APETALA3/PISTILLATA* floral homeotic gene lineages. *Proc Natl Acad Sci USA* 100:6558–6563
- Larouche J, Li P, Bousquet J (1995) Mitochondrial DNA and monocot–dicot divergence time. *Mol Biol Evol* 12:1151–1156
- Ma H, Yanofsky MF, Meyerowitz EM (1990) *AGL1-AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. *Genes Dev* 5:484–495
- Martin W, Gierl A, Saedler H (1989) Molecular evidence for pre-Cretaceous angiosperm origins. *Nature* 339:46–48
- Martin W, Lydiate D, Brinkmann H, Forkmann G, Saedler H, Cerff R (1993) Molecular phylogenetics in angiosperm evolution. *Mol Biol Evol* 10:140–162
- Martinez-Castilla LP, Alvarez-Buylla ER (2003) Adaptive evolution in the *Arabidopsis* MADS-box gene family inferred from its complete resolved phylogeny. *Proc Natl Acad Sci USA* 100:13407–13412
- Mathews S, Donoghue MJ (1999) The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* 286:947–950
- Mathews S, Donoghue MJ (2000) Basal angiosperm phylogeny inferred from duplicated phytochromes A and C. *Int J Plant Sci* 161:S41–S55
- Meyerowitz EM, Bowman J, Brockman LL, Drews GN, Jack T, Sieburth LE, Weigel D (1991) A genetic and molecular model for flower development in *Arabidopsis thaliana*. *Dev Suppl* 1:157–167
- Mouradov A, Hamdorf B, Teasdale RD, Kim J, Winter K-U, Theissen G (1999) A *DEF/GLO*-like MADSbox gene from a gymnosperm: *Pinus radiata* contains an ortholog of angiosperm B class floral homeotic genes. *Dev Genet* 25:245–252
- Munster T, Pahnke J, Di Rosa A, Kim JT, Martin W, Saedler H, Theissen G (1997) Floral homeotic genes were recruited from homologous MADS genes preexisting in the common ancestor of ferns and seed plants. *Proc Natl Acad Sci USA* 94:2415–2420
- Nam J, dePamphilis CW, Ma H, Nei M (2003) Antiquity and evolution of the MADS-box gene family controlling flower development in plants. *Mol Biol Evol* 20:1435–1447
- Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, New York
- Parenicova L, Folter S, Kieffer M, Hrner D, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, Angenent GC, Colombo L (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. *Plant Cell* 15:1538–1551
- Parkinson CL, Adams KL, Palmer JD (1999) Multigene analyses identify the three earliest lineages of extant flowering plants. *Curr Biol* 9:1485–1488

- Purugganan MD (1997) The MADS-box floral homeotic gene lineages predate the origin of seed plants: phylogenetic and molecular clock estimates. *J Mol Evol* 45:392–396
- Qiu YL, Lee J, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M, Zimmer EA, Chen Z, Savolainen V, Chase MW (1999) The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* 402:404–407
- Qiu YL, Lee J, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M, Zimmer EA, Chen Z, Savolainen V, Chase MW (2000) Phylogeny of basal angiosperms: analyses of five genes from three genomes. *Int J Plant Sci* 157:S3–S27
- Qiu YL, Lee J, Whitlock BA, Bernasconi-Quadroni F, Dombrowska O (2001) Was the ANITA rooting of angiosperm phylogeny affected by long-branch attraction? *Mol Biol Evol* 18:1745–1753
- Ramshaw JAM, Richardson DL, Meatyard BT, Brown RH, Richardson M, Thompson EW, Boulter D (1972) The time of origin of flowering plants determined by using amino acid sequence data of cytochrome *c*. *New Phytol* 71:773–779
- Richmann JL, Meyerowitz EM (1997) MADS domain proteins in plant development. *Biol Chem* 378:1079–1101
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol Biol Evol* 14:1218–1231
- Sanderson MJ (2003) Molecular data from 27 proteins do not support a Precambrian origin of land plants. *Am J Bot* 90:954–956
- Sanderson MJ, Doyle JA (2001) Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. *Am J Bot* 88:1499–1516
- Savard L, Li P, Strauss SH, Chase MW, Michaud M, Bousquet J (1994) Chloroplast and nuclear gene sequences indicate late Pennsylvanian time for the last common ancestor of extant seed plants. *Proc Natl Acad Sci USA* 91:5163–5167
- Savolainen V, Chase MW, Hoot SB, Morton CM, Soltis DE, Bayer C, Fay MF, De Bruijn AY, Sullivan S, Qiu Y-L (2000) Phylogenetics of flowering plants based upon a combined analysis of plastid *atpB* and *rbcL* gene sequences. *Syst Biol* 49:306–362
- Shindo S, Ito M, Ueda K, Kato M, Hasebe M (1999) Characterization of MADS genes in gymnosperm *Gnetum pavifolium* and its implication for the evolution of reproductive organs in seed plants. *Evol Dev* 3:180–190
- Soltis DE, Soltis PS (2003) The role of phylogenetics in comparative genetics. *Plant Physiol* 132:1790–1800
- Soltis PS, Soltis DE, Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402:402–404
- Soltis PM, Soltis DE, Savolainen V, Crane PR, Barraclough TG (2002) Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. *Proc Natl Acad Sci USA* 99:4430–4435
- Sommer H, Beltran JP, Huijser P, Pape H, Lonngig WE, Saedler H, Schwarz-Sommer Z (1990) Deficiens, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. *EMBO J* 9:605–613
- Stewart WN, Rothwell GW (1993) Paleobotany and the evolution of plants. Cambridge University Press, New York
- Sundström J, Carlsbecker A, Svensson M, Urban J, Theissen G, Engström P (1999) MADSbox genes active in developing pollen cones of Norway spruce (*Picea abies*) are homologous to the B-class floral homeotic genes in angiosperms. *Dev Genet* 25:253–266
- Swofford DL (2001) PAUP*: phylogenetic analysis using parsimony (*and other methods), v 4. Sinauer, Sunderland, Mass
- Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. *Mol Biol Evol* 12:823–833
- Tanabe Y, Uchida M, Hasebe M, Ito M (2003) Characterization of the *Selaginella remotifolia* MADS-box gene. *J Plant Res* 116:71–75
- Taylor DW, Hickey LJ (1996) Flowering plant origin, evolution and phylogeny. Chapman and Hall, New York
- Templeton AR (1983) Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37:221–244
- Theissen G (2001) Development of floral organ identity: stories from the MADS house. *Curr Opin Plant Biol* 4:75–85
- Theissen G, Saedler H (2001) Floral quartets. *Nature* 409:469–471
- Theissen G, Kim JT, Saedler H (1996) Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. *J Mol Evol* 43:484–516
- Theissen G, Becker A, Rosa AD, Kanno A, Kim JT, Munster T, Winter K-U, Saedler H (2000) A short history of MADS-box genes in plants. *Plant Mol Biol* 42:115–149
- Thorne RF (1992) Classification and geography of the flowering plants. *Bot Rev* 58:225–348
- Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate molecular evolution. *Mol Biol Evol* 15:1647–1657
- Trevisan L (1988) Angiospermous pollen (monosulcate-trichotomosulcate phase) from the very early Lower Cretaceous of Southern Tuscany (Italy): some aspects. 7th Int Palynol Congr (Brisbane) Abstr 165
- Tröbner W, Ramirez L, Motte P, Hue I, Huijser P, Lonngig WE, Saedler H, Sommer H, Sommer ZS (1992) GLOBOSA: a homeotic gene which interacts with DEFICIENS in the control of Antirrhinum floral organogenesis. *EMBO J* 11:4693–4704
- Winter KU, Becker A, Munster T, Kim JT, Saedler H, Theissen G (1999) MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. *Proc Natl Acad Sci USA* 96:7342–7347
- Winter KU, Weiser C, Kaufmann K, Bohne A, Kirchner C, Kanno A, Saedler H, Theissen G (2002) Evolution of class B floral homeotic proteins: obligate heterodimerization originated from heterodimerization. *Plant J* 19:587–596
- Wolfe KH, Gouy M, Yang Y-W, Sharp PM, Li W-H (1989) Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc Natl Acad Sci USA* 86:6201–6205
- Yang Y, Fannig L, Jack T (2003) The K domain mediates heterodimerization of the *Arabidopsis* floral organ identity proteins, *APETALA3* and *PISTILLATA*. *Plant J* 33:47–59
- Yang Z (2002) Phylogenetic analysis by maximum likelihood (PAML), v 3.13. University College London, London
- Yoder AD, Yang Z (2000) Estimation of primate speciation dates using local molecular clocks. *Mol Biol Evol* 17:1081–1090
- Zanis MJ, Soltis DE, Soltis PS, Mathews S, Donoghue MJ (2002) The root of the angiosperms revisited. *Proc Natl Acad Sci USA* 99:6848–6853