Evolution of Alcohol Dehydrogenase Genes in Peonies (*Paeonia*): Phylogenetic Relationships of Putative Nonhybrid Species

Tao Sang, Michael J. Donoghue,† and Daming Zhang‡*

*Department of Botany and Plant Pathology, Michigan State University; †Department of Organismic and Evolutionary Biology, Harvard University; and ‡Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing

Alcohol dehydrogenase genes were amplified by PCR, cloned, and sequenced from 11 putative nonhybrid species of the angiosperm genus *Paeonia.* Sequences of five exons and six intron regions of the *Adh* gene were used to reconstruct the phylogeny of these species. Two paralogous genes, *Adh1A* and *Adh2,* were found; an additional gene, *Adh1B,* is also present in section *Moutan.* Phylogenetic analyses of exon sequences of the *Adh* genes of *Paeonia* and a variety of other angiosperms imply that duplication of *Adh1* and *Adh2* occurred prior to the divergence of *Paeonia* species and was followed by a duplication resulting in *Adh1A* and *Adh1B.* Concerted evolution appears to be absent between these paralogous loci. Phylogenetic analysis of only the *Paeonia Adh* exon sequences, positioning the root of the tree between the paralogous genes *Adh1* and *Adh2,* suggests that the first evolutionary split within the genus occurred between the shrubby section *Moutan* and the other two herbaceous sections *Oneapia* and *Paeonia.* Restriction of *Adh1B* genes to section *Moutan* may have resulted from deletion of *Adh1B* from the common ancestor of sections *Oneapia* and *Paeonia.* A relative-rate test was designed to compare rates of molecular change among lineages based on the divergence of paralogous genes, and the results indicate a slower rate of evolution within the shrubby section *Moutan* than in section *Oneapia.* This may be responsible for the relatively long branch length of section *Oneapia* and the short branch length between section *Moutan* and the other two sections found on the *Adh,* ITS (nrDNA), and *mat*K (cpDNA) phylogenies of the genus. *Adh1* and *Adh2* intron sequences cannot be aligned, and we therefore carried out separate analyses of *Adh1A* and *Adh2* genes using exon and intron sequences together. The Templeton test suggested that there is not significant incongruence among *Adh1A,* ITS, and *mat*K data sets, but that these three data sets conflict significantly with *Adh2* sequence data. A combined analysis of *Adh1A,* ITS, and *mat*K sequences produced a tree that is better resolved than that of any individual gene, and congruent with morphology and the results of artificial hybridization. It is therefore considered to be the current best estimate of the species phylogeny. Paraphyly of section *Paeonia* in the *Adh2* gene tree may be caused by longer coalescence times and random sorting of ancestral alleles.

Introduction

Low-copy-number nuclear genes in plants (Okamuro and Goldberg 1989) are a potentially rich source of information for phylogenetic studies. However, the phylogenetic utility of such genes remains underexplored due largely to the difficulties of distinguishing orthology from paralogy and detecting concerted evolution among members of a gene family (Sanderson and Doyle 1992). Chloroplast DNA (cpDNA) and nuclear ribosomal DNA (nrDNA) are the only widely used sequence data for phylogeny reconstruction in plants. Limited sources of independent gene phylogenies, however, may hamper our ability to obtain accurate species phylogenies when hybridization, lineage sorting, lateral gene transfer, or high homoplasy are involved. Conflicting cpDNA and nrDNA phylogenies have been reported in some plant groups (Soltis and Kuzoff 1995; Mason-Gamer and Kellogg 1996; Soltis, Johnson, and Looney 1996; Sang, Crawford, and Stuessy 1997) and in such cases it will be critical to obtain additional independent gene phylogenies and compare and combine them for stronger hypotheses of the species phylogenies (Hillis 1995).

Key words: alcohol dehydrogenase, *Adh, Paeonia,* phylogeny, gene duplication, relative-rate test.

Address for correspondence and reprints: Tao Sang, Department of Botany and Plant Pathology, Michigan State University, East Lansing, Michigan 48824. E-mail: sang@pilot.msu.edu.

Mol. Biol. Evol. 14(10):994–1007. 1997

In the present study, we cloned and sequenced alcohol dehydrogenase (*Adh*) genes in the angiosperm genus *Paeonia* (Paeoniaceae) to investigate the phylogenetic utility of a low-copy nuclear gene family at the interspecific level in plants and to better understand the evolution of peonies. *Adh* genes are among the best characterized nuclear protein-coding genes in plants. In the majority of flowering plants that have been studied, two to three *Adh* loci have been identified, each containing 10 exons and 9 introns (Gottlieb 1982; Dennis et al. 1985; Llewellyn et al. 1987; Trick et al. 1988; Wolyn and Jelenkovis 1990; Morton, Gaut, and Clegg 1996). In *Arabidopsis* and *Arabis* (Brassicaceae), however, a single *Adh* locus is present, which consists of seven exons and six introns (Chang and Meyerowitz 1986; Miyashita Innan, and Terauchi 1996). In *Paeonia californica,* two *Adh* loci were detected by enzyme electrophoresis (Zona et al. 1991).

Phylogenetic uses of *Adh* genes have been concerned with both high and low taxonomic levels. Sequences of amino acids derived from *Adh* gene sequences of vertebrates and plants were analyzed together using yeasts as outgroups (Yokoyama and Harry 1993). Sequences of *Adh* genes, including both exons and introns, have been used to resolve phylogenetic relationships within *Drosophila* (Jeffs, Holmes, and Ashburner 1994; Russo, Takezaki, and Nei 1995; Nurminsky et al. 1996). In plants, phylogenies of *Adh* genes and divergence rates were inferred based on exon sequences of representative grasses and palms (Gaut et al. 1996; Morton, Gaut, and Clegg 1996).

Table 1
Taxonomy and Sample Localities *Paeonia* Species Included in this Study, the Number of Clones Screened, and the Numbers of *Adh1A, Adh1B, Adh2,* and Recombined Clones Obtained

Species	Sample Localities	Abbreviation	No. Screened	*Adh1A*	*Adh1B*	*Adh2*	Recombined
Section *Moutan*							
Subsection *Delavayanae*							
P. delavayi	Lijiang, Yunnan, China	DEL	8		2	4	2
P. lutea	Mt. Xi, Yunnan, China	LUT1	4	3		1	
	Bomi, Tibet	LUT2	6		2	4	
Subsection *Vaginatae*							
P. rockii	Wenxian, Gansu, China	ROC1	6		1	4	1
	Shenglongija, Hubei, China	ROC2	4	1		3	
P. suffruticosa subsp. *spontanea*	Mt. Ji, Shaanxi, China	SPO	6	1	3	2	
P. szechuanica	Marekang, Sichuan, China	SZE	7	3		4	
Section *Oneapia*							
P. brownii	Modoc, Calif.	BRW	8			8	
P. californica	Los Angeles, Calif.	CAL1	4	3		1	
	St. Louis Obispo, Calif.	CAL2	4	1		3	
Section *Paeonia*							
P. anomala	Yiling, Xinjiang, China	ANO	12	9[a]		4[b]	
P. lactiflora	Chicheng, Hebei, China	LAC	19	8[a]		6[a], 5[b]	
P. tenuifolia	Sofia, Bulgaria	TEN	11	7		4[b]	
P. veitchii	Mt. Taibei, Shaanxi, China	VEI1	6			6	
	Mongda, Qinghai, China	VEI2	7	2		4	1

[a] Clones obtained from two PCR reactions.
[b] Clones of PCR products of the *Adh2*-specific primers (fig. 1).

Paeonia consists of approximately 35 species placed in three sections, *Moutan, Oneapia,* and *Paeonia* (Stern 1946; Pan 1979; Tzanoudakis 1983). Section *Moutan* contains five diploid shrubby species occurring in central and western China. Section *Oneapia* contains two diploid perennial herbaceous species endemic to Pacific North America. Section *Paeonia* comprises approximately 28 diploid and tetraploid herbaceous species found in eastern Asia, central Asia, the western Himalayas, and the Mediterranean region. The genus has been placed in its own family, Paeoniaceae, and often in its own order, Paeoniales (Takhtajan 1969, 1987; Thorne 1992); its broader relationships within angiosperms have been controversial (Keefe and Moseley 1978; Melville 1983; Cronquist 1988), although it appears to be related to Crassulaceae based on *rbc*L and 18S ribosomal DNA sequences (Chase et al. 1993; Rice, Donoghue, and Olmstead 1997; Soltis et al. 1997). Recent studies using sequences of the internal transcribed spacer (ITS) region of nrDNA and the *mat*K gene of cpDNA revealed complex reticulated evolution within section *Paeonia* (Sang, Crawford, and Stuessy 1995, 1997).

Here we focus on reconstructing *Adh* gene trees for all of the putative nonhybrid species (Sang, Crawford, and Stuessy 1997). We have identified orthologous and paralogous *Adh* genes based on sequence divergence and phylogenetic analyses. Sequence divergence has also helped us determine whether concerted evolution of paralogous genes has occurred. We have taken advantage of *Adh* gene duplications for rooting purposes and to test relative rates of molecular divergence among peony lineages. Exon and intron sequences of the orthologous genes are analyzed together, and the resulting gene trees are compared with previous nrDNA and cpDNA phylogenies. A combined analysis of *Adh,* ITS, and *mat*K sequences was conducted to estimate the species phylogeny.

Materials and Methods

Amplification, Cloning, and Sequencing

Total DNAs were isolated from leaf tissue using the CTAB method (Doyle and Doyle 1987) and purified in CsCl/ethidium bromide gradients. Leaves of most of the *Paeonia* species included in this study were collected from natural populations in Bulgaria, China, and the United States (table 1).

Two PCR primers, AdhF1 and AdhR1, were designed for maximal coverage of *Adh* genes using regions conserved across the eudicot families Brassicaceae, Fabaceae, Solanaceae, and Rosaceae (Chang and Meyerowitz 1986; Llewellyn et al. 1987; Ellison, Yu, and White 1990; Wolyn and Jelenkovis 1990). These primers amplify a large portion of the peony *Adh* genes (fig. 1). The PCR reactions were carried out under standard conditions with *Taq* DNA polymerase (Gibco BRL Life Technologies) on a GeneAmp PCR system 9600 (Perkin Elmer). The PCR reactions include the following cycles: (1) 94°C, 2 min; (2–5) 94°C, 45 s—55°C, 1.5 min—72°C, 1 min; (6–36) 94°C, 30 s—55°C, 40 s—72°C, 1 min; (37) 72°C, 5 min. The PCR products were examined with 1% agarose gel in TBE buffer.

PCR reactions using primers AdhF1 and AdhR1 produced a single band of approximately 1.4 kb for the peony DNAs that we were able to amplify (including some hybrid species). Direct sequencing of purified PCR products with either of these primers produced only about 200 bp of readable sequence. Based on sequences obtained from several species, two new primers, AdhF2

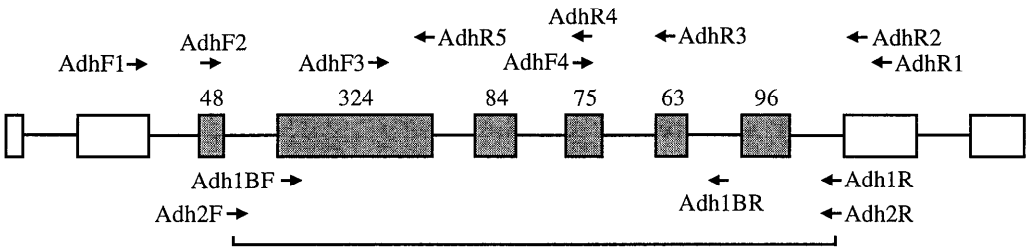

FIG. 1.—Diagram of the *Adh* gene in peonies. Shaded boxes represent the exons amplified and sequenced, with sizes indicated above by the number of base pairs; open boxes represent the remaining exons of *Adh* genes as documented in other eudicots. The line below the gene represents the region sequenced and used in phylogenetic analyses. Arrows indicate the locations and directions of the PCR and sequencing primers: AdhF1, 5′-TACTTCTGGGAAGCYAAGG-3′; AdhF2, 5′-CCTCGCATATTTGGTCACGAAG-3′; AdhF3, 5′-GTTGGAACTTCAACTTTCAG-3′; AdhF4, 5′-TCAAGGATTATTGGTGTTGA-3′; AdhR1, 5′-CCCTTRAGMGTCCTCTCATTC-3′; AdhR2, 5′-GGGCACACCAACAAGTACTG-3′; AdhR3, 5′-CTGTCACACCAAATTTCTTGG-3′; AdhR4, 5′-TCAACACCAATAATCCTTGA-3′; AdhR5, 5′-AGAACACAAACTTTATCAAG-3′; Adh1R, 5′-CATCCCTGCAACAAGATA-3′; Adh1BF, 5′-TGAAGGTGTCACGGATCTT-3′; Adh1BR, 5′-GAGATTAGAACAAACCACTGT-3′; Adh2F, 5′-TGGGTTTCCTTGATTGATTG-3′; Adh2R, 5′-CATCCCTGCAACATGCAT-3′. Adh1R is an *Adh1*-gene-specific primer. Adh1BF and Adh1BR are *Adh1B*-gene-specific primers. Adh2F and Adh2R are *Adh2*-gene-specific primers.

and AdhR2, were designed in the conserved regions and led to approximately 500 bp of readable sequence (fig. 1). When AdhF2 and AdhR2 were used as PCR primers, a single band was obtained which was much stronger than that obtained for the same species using primers AdhF1 and AdhR1. AdhF2 and AdhR2 also amplified the species that we could not amplify with AdhF1 and AdhR1. The results suggest that AdhF2 and AdhR2 match *Adh* sequences of peonies better than the first set of primers, and we used PCR products amplified with these two primers for cloning and subsequent sequencing.

PCR products were ligated with plasmids and transformed into *E. coli* competent cells, which were selected on the plates containing ampicillin and X-Gal (Original TA Cloning Kit, Invitrogen). Ten to 15 white *E. coli* clones were picked and cultured for isolating plasmids. Purified plasmid DNAs were digested with *Eco*RI and amplified with primers AdhF2 and AdhR2 to check whether they contain the correct inserts. The plasmids with correct inserts were sequenced by at least one primer, usually AdhR2 (reading three introns), to screen variation among clones. Clones with sequences determined to be different were sequenced for both strands in their entirety and included in the phylogenetic analyses. The locations and sequences of the PCR and sequencing primers used in this study are given in figure 1 (also see *Results*). Sequencing was conducted using ABI370A and ABI373A automated DNA sequencers with the *Taq* Cycle Sequencing DyDeoxy Terminator reagents (Applied Biosystems). Sequences were edited with the program SeqEd and aligned manually. Sequences obtained in this study have been assigned GenBank accession numbers AF009041–AF009068.

Phylogenetic Analyses

Maximum parsimony, as implemented in PAUP 3.1.1 (Swofford 1993), was used to infer relationships based on nucleotide substitutions in aligned *Adh* sequences. Heuristic searches were performed using TBR branch swapping MULPARS on 100 starting trees derived using the random option in PAUP. Branch-and-bound was employed when there were less than 18 sequences in an analysis. ACCTRAN option was used for character optimization. Bootstrap analyses (Felsenstein 1985) were carried out with 500 replicates, using simple taxon addition. Nucleotide substitutions were weighted equally, and gaps were treated as missing information.

Analysis I

Sequences of five *Adh* exons of peony species and other angiosperms were aligned (without gaps) and analyzed using monocot sequences for rooting purposes. *Adh* sequences of the following taxa were obtained from GenBank: *Arabidopsis thaliana* (M12196), cotton (*Gossypium hirsutum*; U49061), apple (*Malus domestica*; Z48234), strawberry (*Fragaria ananassa*; X15588), tobacco (*Nicotiana tabacum*; X81853), petunia (*Petunia hybrida*; X54106), teppary bean (*Phaseolus acutifolius*; Z23171), pea (*Pisum sativum*; X06281), white clover (*Trifolium repens*; X14826), maize (*Zea mays*; *Adh1,* X04049; *Adh2,* X02915), and barley (*Hordeum vulgare*; *Adh1,* X12732; *Adh2,* X12733; *Adh3,* X12734).

Analysis II

In previous phylogenetic analyses of *Paeonia* using ITS and cpDNA sequences (Sang, Crawford, and Stuessy 1995, 1997), gene trees were rooted along the longest branch, thereby separating section *Oneapia* from the other two sections. This approach was necessitated by the great distance of any potential outgroups. Analysis I (above), which included distantly related outgroups, did not help resolve the root within *Paeonia* (see below). Nevertheless, it did suggest that the duplication giving rise to the *Adh1* and *Adh2* genes occurred before the diversification of *Paeonia.* Analysis II was designed to include only *Paeonia Adh* genes, with trees rooted between the two paralogous genes (see *Discussion* for references on this approach). Since introns could not be aligned between the two paralogous genes, only exon sequences were analyzed.

Analysis III

In order to obtain better resolution of interspecific relationships, both exons and introns of the orthologous genes, *Adh1* and *Adh2,* were analyzed separately. Based

on the results of Analysis II and additional information on rates of evolution (see below) the trees were rooted between section *Moutan* and the other two sections. Several pseudogenes, identified based on deletions in the exons, were also included in these analyses.

Templeton Test

Incongruence among *Adh1A, Adh2,* ITS, and *mat*K data sets was tested using the Templeton test (Templeton 1983; Larson 1994). This test was chosen because it can identify potentially conflicting individual nodes among trees while entire data sets are compared (Mason-Gamer and Kellogg 1996). Only the 11 species included in this study were retained in ITS and *mat*K data sets (Sang, Crawford, and Stuessy 1997), and *Adh1A* and *Adh2* data sets were reduced by randomly choosing a single clone representing a species in each data set.

Combined Analysis

Since the Templeton test implies a significant incongruence between the *Adh2* data set and each of the other data sets but not among *Adh1A,* ITS, and *mat*K, a combined analysis of these three data sets was conducted to better estimate the species phylogeny (de Queiroz, Donoghue, and Kim 1995; Huelsenbeck, Bull, and Cunningham 1996). However, we also conducted a combined analysis all four data sets, *Adh1A, Adh2,* ITS, and *mat*K.

Comparisons of Sequence Divergence

Sequence divergences for all nucleotides, for synonymous sites, and for nonsynonymous sites were estimated using Jukes-Cantor corrections (Jukes and Cantor 1969) as calculated by MEGA 1.02 (Kumar, Tamura, and Nei 1993). Pseudogenes were not included in comparisons of sequence divergence.

When concerted evolution occurs between paralogous genes, it is expected that sequence divergence between paralogues will be significantly lower within than among taxa (Dover 1982; Arnheim 1983; Hughes 1995). Concerned evolution between the paralogous genes of *Paeonia* was examined (1) by comparing average sequence divergence within species and among species within each section, and (2) by comparing average sequence divergence within each section and among the three sections.

We have made use of gene duplication to test relative rates of molecular evolution (fig. 2). If a duplication resulting in two paralogous genes (1 and 2) occurred prior to the divergence of two taxa (A and B) (fig. 2*A*), sequence divergence between genes A1 and A2 ($d_{A1,A2}$) should be equal to the divergence between B1 and B2 ($d_{B1,B2}$)—that is, $d_{A1,A2} - d_{B1,B2} = 0$—if molecular evolution occurred at the same rate in taxa A and B and concerted evolution between paralogous genes 1 and 2 was absent. The variance of ($d_{A1,A2} - d_{B1,B2}$) is equal to the variance of ($d_{A1,A2}$) plus the variance of ($d_{B1,B2}$) minus twice the covariance of ($d_{A1,A2}$, $d_{B1,B2}$). The covariance of ($d_{A1,A2}$, $d_{B1,B2}$) equals the variance of x, where x is the estimated length of the interior branch on the unrooted tree for the four genes (fig. 2*B*),

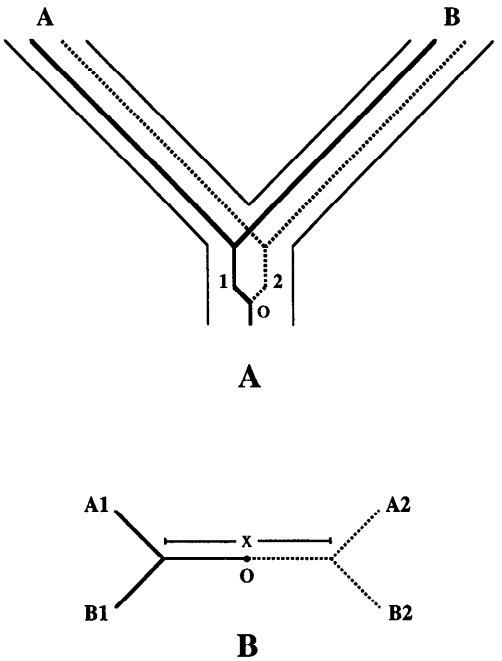

FIG. 2.—*A,* Duplication of an ancestral gene at point O, giving rise to genes 1 and 2 prior to the divergence of taxa A and B. *B,* An unrooted tree of the four genes present in A and B, with the interior branch length x.

which equals half of $d_{A1,A2} + d_{B1,B2} - d_{A1,B1} - d_{A2,B2}$ (Wu and Li 1985; M. Sanderson, personal communication). The variance of x equals $P(1 - P)/[L(1 - P/0.75)^2]$, where $P = 0.75(1 - e^{-x/0.75})$ and L is the number of sites compared (Li and Tanimura 1987).

Results

Adh Genes in *Paeonia*

The number of clones screened and the number of clones found for each paralogous *Adh* gene in each species are shown in table 1. Two types of sequences, which we have called *Adh1* and *Adh2,* were discovered in peony species; the introns of these two forms are highly diverged and could not be aligned. After the first round of cloning and screening, only *Adh1* clones were obtained from *P. anomala* or *P. tenuifolia,* and only *Adh2* clones were found in *P. brownii. Adh1*- and *Adh2*-specific primers were then designed based on the obtained sequences in the intron regions of the other species to amplify these three species (fig. 1). *Adh2* genes were cloned successfully from *P. anomala* and *P. tenuifolia* using the *Adh2*-specific primers (Adh2F and Adh2R). Amplification of the *Adh1* gene of *P. brownii* using the *Adh1*-specific primers (Adh1R) with AdhF2 still failed. This was probably due to poor quality of the template DNA; the DNA of *P. brownii* was isolated from dry specimens, while fresh leaves were used for the rest of the species.

Two types of *Adh1* genes were cloned from the species of the shrubby section *Moutan.* The type that is also found in the rest of genus is designated as *Adh1A,* and the type that is limited to section *Moutan* is called *Adh1B.* To better establish the distribution of the *Adh1B*

gene in peonies, a pair of *Adh1B*-specific primers, Adh1BF and Adh1BR (fig. 1), were designed and used in PCR and DNAs of all species included in the study. The *Adh1B* gene was amplified for all the studied individuals belonging to section *Moutan,* but from none of the species belonging to the other two sections. The *Adh1B* locus, therefore, appears to be present only in section *Moutan.* Both *Adh1A* and *Adh1B* genes were cloned from the same individual of *P. suffruticosa* subsp. *spontanea* (table 1). The absence of *Adh1A* clones for DEL and LUT2 and of *Adh1B* clones for LUT1, ROC2, and SZE is most likely due to screening an insufficient number of clones. An exhaustive search for these genes was not carried out, because it does not appear to be critical for understanding paralogy and orthology of *Adh* genes or interspecific relationships of the peony species in this study.

Four clones, two from *P. delavayi,* and one each from *P. rockii* and *P. veitchii,* have combined sequences of *Adh1* and *Adh2* genes. To determine whether these resulted from amplification of genomic interlocus recombinations, a combination of *Adh1*- and *Adh2*-specific primers, Adh2F and Adh1R, was used to amplify the genomic DNAs of these three species. None of the PCR reactions yielded any bands, implying that the recombined sequences we obtained were an artifact of particular PCR reactions (Bradley and Hillis 1997). We, therefore, urge caution in interpreting sequences of PCR products of low-copy-number nuclear genes.

Adh2 pseudogenes were cloned from *P. delavayi, P. lactiflora,* and *P. suffruticosa* subsp. *spontanea.* The pseudogenes found in *P. delavayi* and *P. lactiflora* have a 1-bp deletion in the 324-bp exon (at different sites). A 2-bp deletion in this exon is found in the pseudogene of *P. suffruticosa* subsp. *spontanea.* These deletions resulted in the stop codons in the exon. Both the pseudogene and the normal *Adh2* gene were cloned from an individual of *P. delavayi* and an individual of *P. suffruticosa* subsp. *spontanea.* Only the pseudogene was found in *P. lactiflora* after screening 11 *Adh2* clones, which implies that the normal *Adh2* gene is either absent from this individual plant or could not be amplified for some unknown reasons.

Phylogenetic Analyses

Analysis I, including exon sequences of all peony *Adh* clones and other angiosperm *Adh* genes, generated 181 most-parsimonious trees; the strict consensus of these is shown in figure 3. The resolved and well-supported clades within eudicots include *Paeonia,* Rosaceae (including apple and strawberry), Fabaceae (including pea, teppary bean, and white clover), and Solanaceae (including tobacco and petunia). A clade containing cotton and *Arabidopsis* is more weakly supported (72% bootstrap). Relationships among these major clades, however, are not supported by bootstrap values higher than 50%, despite the relatively long branch lengths evident in figure 3.

All peony *Adh* sequences form a single clade, within which there are two major groups corresponding to *Adh1* and *Adh2.* This result implies that a duplication

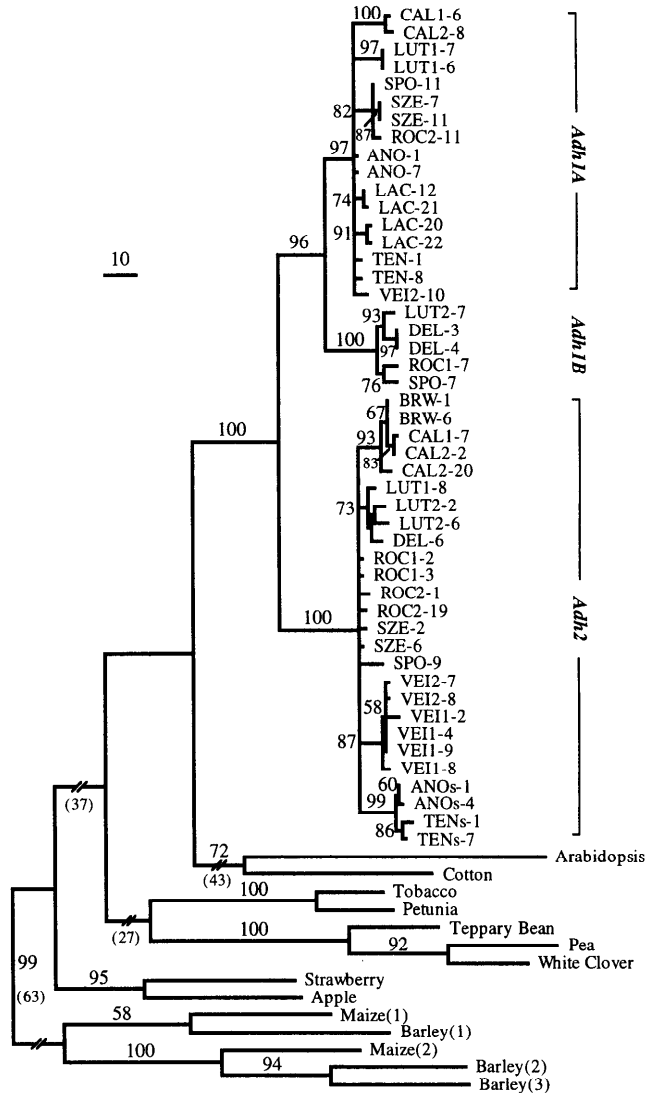

FIG. 3.—The strict consensus tree of 181 most-parsimonious trees of length 1,297 (consistency index = 0.49; retention index = 0.75) based on sequences of five exons of *Adh* genes of peonies and other angiosperms. Branch lengths are drawn proportional to the number of nucleotide substitutions (scale represents 10 substitutions); branch lengths are reduced in the several cases indicated by broken branches and numbers in parentheses. All other numbers associated with branches are bootstrap percentages above 50%. For species and sample abbreviations, see table 1; a number following an "–" represents a clone number. ANOs and TENs represent clones from PCR products amplified by *Adh2*-gene-specific primers (fig. 1).

event occurred prior to the diversification of the peonies. The *Adh1A* and *Adh1B* clades split at the base of the *Adh1* clade. Most other relationships within peonies, except for the monophyly of section *Oneapia,* are poorly resolved in this analysis.

Analysis II, involving peony exon sequences only, yielded six most-parsimonious trees; the strict consensus of these is shown in figure 4 (rooted between *Adh1* and *Adh2*). In the *Adh1A* clade, species of the herbaceous sections *Oneapia* and *Paeonia* form a monophyletic group supported by a 72% bootstrap value, suggesting that the root of the genus may be near the shrubby section *Moutan.* In the *Adh2* clade, however, the relation-

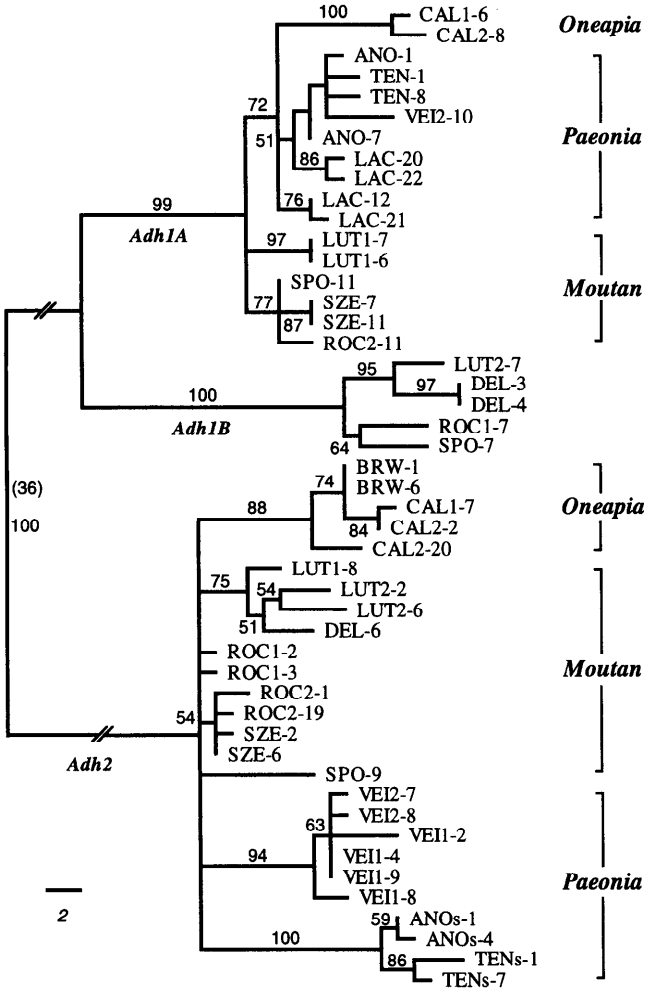

FIG. 4.—The strict consensus tree of six most-parsimonious trees of length 192 (consistency index = 0.86; retention index = 0.97) based on sequences of five exons of the cloned *Adh1* and *Adh2* genes of peonies. Branch lengths are proportional to the number of nucleotide substitutions (scale represents two substitutions). The branch between *Adh1* and *Adh2* genes is reduced; its length is given in parentheses. All other numbers are bootstrap percentages above 50%.

ship among sections is not clearly resolved, although section *Oneapia* is supported as monophyletic. Despite low bootstrap support for a link between any two of the sections, the highest bootstrap value of 36% supports the relationship of sections *Oneapia* and *Paeonia*; this compares to 17% for *Oneapia* and *Moutan* and 0.1% for *Moutan* and *Paeonia.*

Analysis III (*Adh1A* and *Adh2* analyzed separately, with introns added) provided better resolution of relationships within *Paeonia* (fig. 5*A* and *B*). Monophyly of section *Oneapia,* section *Moutan,* and the subsections *Delavayanae* and *Vaginatae* of section *Moutan* is strongly supported. The monophyly of section *Paeonia* is also strongly supported on the *Adh1A* tree. These results are concordant with the relationships previously obtained based on ITS and *mat*K sequences (fig. 6*A* and *B*; Sang, Crawford, and Stuessy 1997). Discordance among these four gene trees is found within section *Paeonia.* In particular, section *Paeonia* is paraphyletic in the *Adh2* tree; sequences from two populations of *P.*

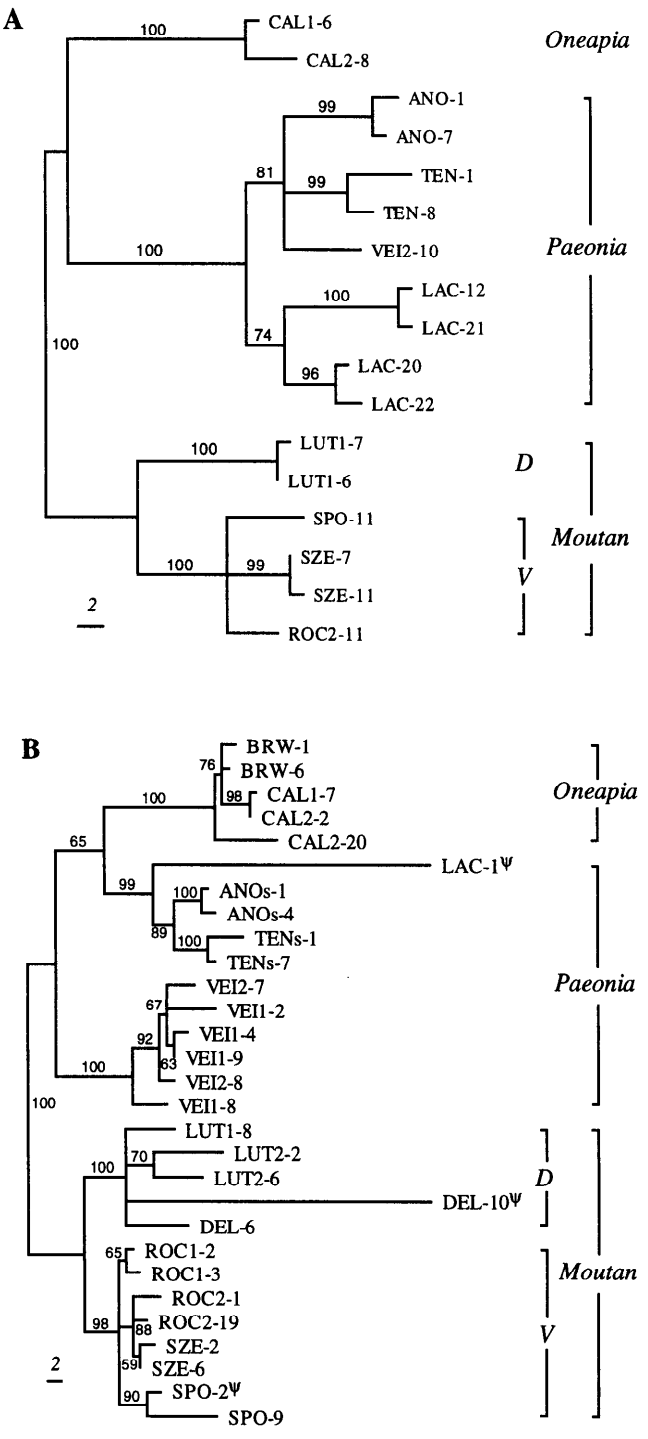

FIG. 5.—*Adh* gene trees of *Paeonia* species. *A, Adh1A* gene tree; strict consensus of three most-parsimonious trees of length 129 (consistency index [CI] = 0.95; retention index [RI] = 0.97). *B, Adh2* gene tree; strict consensus of 24 most-parsimonious trees of length 279 (CI = 0.93; RI = 0.95). Pseudogenes are marked by ψ. Branch lengths are proportional to the number of nucleotide substitutions (scale represents two substitutions). Numbers associated with branches are bootstrap percentages above 50%. *D* and *V* represent subsections *Delavayanae* and *Vaginatae* of section *Moutan,* respectively.

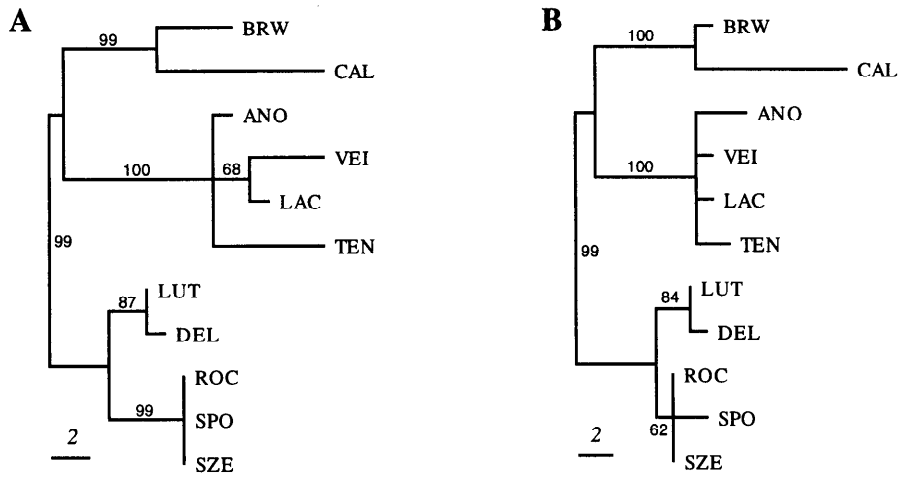

FIG. 6.—Peony gene trees of nrDNA and cpDNA. *A*, ITS (nrDNA) gene tree; strict consensus of two most-parsimonious trees of length 50 (consistency index [CI] = 0.94; retention index [RI] = 0.94). *B*, *mat*K gene (cpDNA) tree; strict consensus of four most-parsimonious trees of length 40 (CI = 0.95; RI = 0.96). For the sake of comparison, ITS and *mat*K sequences of the 11 species represented in the *Adh* gene trees (Sang, Crawford, and Stuessy 1995, 1997) were reanalyzed and rooted between section *Moutan* and the other two sections. Branch lengths are proportional to the number of nucleotide substitutions (scale represents two substitutions). Numbers associated with branches are bootstrap percentages above 50%.

veitchii form their own clade, which is a sister group to the clade containing section *Oneapia* and the remaining species of section *Paeonia.* Discordance is also found between the *Adh1A* and ITS trees: *P. lactiflora* and *P. veitchii* are sister groups on the ITS tree (supported by a 68% bootstrap value), whereas *P. anomala, P. tenuifolia,* and *P. veitchii* form a monophyletic group in the *Adh1A* trees (supported by 81% bootstrap).

The Templeton test was conducted to determine whether the discordance among gene trees is significant (table 2). When the *Adh2* topology (section *Paeonia* paraphyletic) is used as the constraint topology, parsimony analysis of the *Adh1A,* ITS, and *mat*K data sets are each seen to be significantly worse. In contrast, when the relationships of section *Paeonia* found in the ITS tree are used as the constraint topology, parsimony analysis of the *Adh1A* data set is not significantly worse. A combined analysis of *Adh1A,* ITS, and *mat*K data sets resulted in the single most-parsimonious tree shown in figure 7*A*. The single most-parsimonious tree generated from the combined analysis of all four data sets is shown in figure 7*B*. These two trees are better resolved than any individual gene trees but differ by the positions of *P. veitchii* and *P. lactiflora.*

Table 2
Templeton Test for Incongruence Between the *Adh2* Data Set and the *Adh1A,* ITS, and *mat*K Data Sets, and Between the *Adh1A* and ITS Data Sets

Data Set	L	Constraint	Lc	*N*	Ts	*P*
Adh1A ...	112	*Adh2*[a]	129	15	0	<0.01
ITS.....	51	*Adh2*	58	7	0	<0.01
*mat*K	41	*Adh2*	46	7	1	<0.05
Adh1A ...	112	ITS[b]	115	3	0	>0.1

NOTE.—L: lengths of trees resulting from analyses without constraint; Lc: lengths of trees resulting from analyses with constraint; *N:* number of characters undergoing step changes after constraint analysis; Ts: test statistic; *P:* probability. For the ITS and the *mat*K data sets, results of comparing consensus trees are given, which are essentially the same as the results of comparing individual parsimonious trees.

[a] Paraphyletic relationships of section *Paeonia* on the *Adh2* tree used as the constraint topology: VEI((CAL, BRW), (ANO, TEN, LAC)).

[b] Relationships within section *Paeonia* on the ITS tree used as the constraint topology: ANO(VEI, LAC)TEN.

It is noteworthy that clones of each species are united on the *Adh1A* trees (fig. 5*A*). On the *Adh2* trees, however, the clones of each species are seen to be directly linked for only half of the species (fig. 5*B*). Two clones from one population of *P. rockii* (ROC2-1 and ROC2-19) form a clade with *P. szechuanica,* while two clones from another population of *P. rockii* (ROC1-2 and ROC1-3) form a trichotomy with this clade and *P. suffruticosa* subsp. *spontanea.* Clones from the same population of *P. lutea* (LUT2-2 and LUT2-6) are also grouped together but do not form a monophyletic group with the LUT1-8 clone from a different population. In *P. californica,* clones from two populations (CAL1-7 and CAL2-2) form a well-supported group, while a second clone (CAL2-20) from one of the two populations appears as the sister group of a clade containing *P. californica* and *P. brownii.*

Sequence Divergence

We compared average sequence divergence between *Adh1A* and *Adh2* genes within and among species within each section, as well as within and among sections (table 3). Likewise, we compared *Adh1A* and *Adh1B* genes within and among species within section *Moutan.* Within each section, average overall sequence divergence of *Adh1A* and *Adh2* genes is slightly higher within species than among them, suggesting that concerted evolution between the two paralogous loci has not occurred within species after the divergence of each section (Dover 1982; Hughes 1995). Almost identical average sequence divergence between *Adh1A* and *Adh2* genes is found within sections and among sections, suggesting that concerted evolution has not occurred within

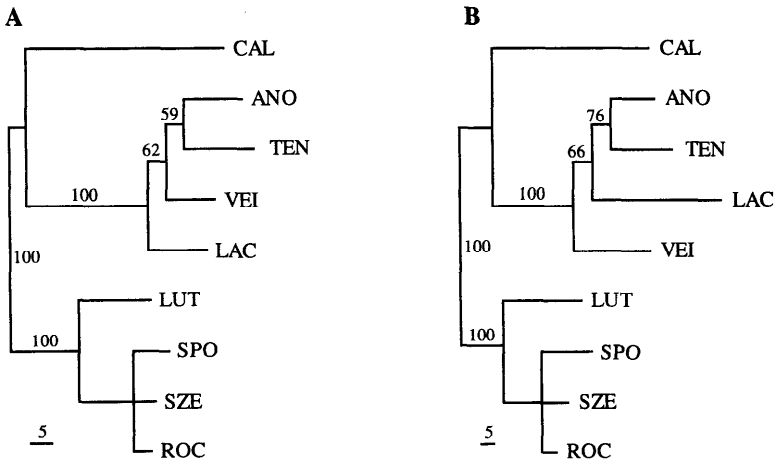

FIG. 7.—Trees obtained from combined analyses. *A*, The single most-parsimonious tree obtained from combined analysis of *Adh1A*, ITS, and *mat*K sequences; tree length = 197, consistency index [CI] = 0.96, retention index [RI] = 0.96. The following *Adh1A* clones were chosen randomly to represent species in the analysis: CAL2-6, ANO-1, TEN-8, VEI-10, LAC-12, LUT1-7, SPO-11, SZE-7, ROC2-11. *B*, The single most-parsimonious tree obtained from combined analysis of *Adh1A, Adh2,* ITS, and *mat*K sequences; tree length = 366, CI = 0.95, RI = 0.92. The follwoing *Adh2* clones were chosen randomly to represent species in the analysis: CAL2-7, ANOs-1, TENs-1, VEI2-7, LAC-1, LUT2-6, SPO-9, SZE-2, ROC1-2. Branch lengths are proportional to the number of nucleotide substitutions (scale represents five substitutions). Numbers associated with branches are bootstrap percentages above 50%.

any section following the divergence of the three sections. Average sequence divergence between *Adh1A* and *Adh1B* is very similar within species and among species in section *Moutan,* again implying that concerted evolution has not occurred between these two paralogous loci.

The average divergence between *Adh1A* and *Adh2* genes is highest in section *Oneapia* and lowest in sec-

Table 3
Average Percent Divergence (with the range) of Exon Sequences Between Two Pairs of Paralogous Genes (*Adh1A* and *Adh2, Adh1A* and *Adh1B*) and Exon and Intron Sequences of Each of the *Adh1A* and *Adh2* Genes at all ($\bar{d}$), Synonymous ($\bar{d}_s$), and Nonsynonymous ($\bar{d}_n$) Sites

	$\bar{d}$	$\bar{d}_s$	$\bar{d}_n$
Adh1A–Adh2			
Within section *Moutan*	8.21 (7.19–8.92)	34.01 (29.60–39.07)	1.88 (1.54–2.58)
Within species	8.21 (7.53–8.92)	34.06 (31.66–36.95)	1.86 (1.54–2.16)
Among species	8.12 (7.19–8.92)	33.74 (29.60–39.07)	1.82 (1.54–2.37)
Within section *Oneapia*[a]	9.36 (9.27–9.45)	36.50 (35.66–37.79)	2.74 (2.37–3.01)
Within species	9.39 (9.27–9.45)	36.24 (35.66–36.92)	2.83 (2.59–3.01)
Among species	9.27 (9.27–9.27)	37.30 (36.80–37.79)	2.48 (2.37–2.58)
Within section *Paeonia*	8.68 (7.88–9.80)	33.55 (30.64–39.13)	2.55 (2.16–3.42)
Within species	8.99 (8.40–9.63)	34.88 (33.50–36.92)	2.64 (2.16–3.21)
Among species	8.80 (8.05–9.80)	33.79 (31.70–39.13)	2.64 (2.16–3.42)
Within three sections	8.58 (7.19–9.80)	33.87 (29.60–39.13)	2.35 (1.54–3.21)
Among three sections	8.56 (7.19–9.98)	33.54 (27.60–41.23)	2.39 (1.74–3.31)
Adh1A–Adh1B			
Within section *Moutan*	5.30 (4.82–5.66)	17.97 (16.42–20.43)	1.87 (1.64–2.27)
Within species	5.15 (4.82–5.32)	17.45 (16.44–18.16)	1.81 (1.64–1.95)
Among species	5.21 (4.99–5.49)	17.46 (16.42–19.50)	1.88 (1.64–2.27)
Adh1A[b]			
Exons	1.36 (0.31–2.37)	3.64 ± 0.89[d]	0.69 ± 0.23[d]
Introns	4.68 (1.09–7.30)	—	—
Adh2[c]			
Exons	2.14 (0.47–3.67)	7.53 ± 1.34[d]	0.59 ± 0.17[d]
Introns	4.54 (0.52–6.47)	—	—

[a] Since there are only two species in section *Oneapia,* and only *Adh2* genes were cloned from *P. brownii,* within-species divergence is represented by divergence within only *P. californica,* and divergence between *Adh2* of *P. brownii* and *Adh1A* of *P. californica* is calculated as among-species divergence.

[b] One clone is randomly chosen to represent one of nine species from which *Adh1A* genes were cloned.

[c] One clone is randomly chosen to represent one of 10 species from which *Adh2* genes were cloned.

[d] Standard errors of $\bar{d}_s$ and $\bar{d}_n$ were calculated by MEGA, and $\bar{d}_s$ is significantly higher in *Adh2* than in *Adh1* ($p < 0.05$), while $\bar{d}_n$ is not significantly different between *Adh1* and *Adh2* (*t*-test; Kumar, Tamura, and Nei 1993).

Table 4
Relative-Rate Tests Between Section *Oneapia* and Section *Moutan* Based on Overall Percent Divergence of Paralogous Genes, *Adh1A* and *Adh2*

A1,A2	B1,B2	$d_{A1,A2}$±SE	$d_{B1,B2}$±SE	$d_{A1,B1}$	$d_{A2,B2}$	$d_{A1,A2} - d_{B1,B2}$±SE
CAL1c6, CAL1c7[a]	LUT1c7, LUT2c6[a]	9.45 ± 0.0161	8.92 ± 0.0151	0.0221	0.0286	0.53 ± 0.95
CAL1c6, CAL1c7	LUT1c7, LUT2c2[b]	9.45 ± 0.0161	8.40 ± 0.0141	0.0221	0.0253	1.05 ± 0.87
CAL1c6, CAL1c7	SPOc11, SPOc9	9.45 ± 0.0161	8.57 ± 0.0146	0.0189	0.0302	0.88 ± 0.95
CAL1c6, CAL1c7	SZEc7, SZEc2[a]	9.45 ± 0.0161	8.05 ± 0.0137	0.0221	0.0221	1.40 ± 0.90*
CAL1c6, CAL1c7	SZEc7, SZEc6[b]	9.45 ± 0.0161	7.88 ± 0.0132	0.0221	0.0205	1.57 ± 0.87**
CAL1c6, CAL1c7	ROC2c11, ROC2c19[a]	9.45 ± 0.0161	8.05 ± 0.0137	0.0205	0.0221	1.40 ± 0.88*
CAL1c6, CAL2c7	ROC2c11, ROC1c2[b]	9.45 ± 0.0161	7.53 ± 0.0128	0.0205	0.0205	1.92 ± 0.86**
CAL1c6, CAL3c2[b]	SZEc7, SZEc2	9.27 ± 0.0159	8.05 ± 0.0137	0.0221	0.0205	1.22 ± 0.88*
CAL1c6, CAL3c2	SZEc7, SZEc6	9.27 ± 0.0159	7.88 ± 0.0132	0.0221	0.0189	1.39 ± 0.86*
CAL1c6, CAL3c2	ROC2c11, ROC2c19	9.27 ± 0.0159	8.05 ± 0.0137	0.0205	0.0205	1.22 ± 0.87*
CAL1c6, CAL3c2	ROC2c11, ROC1c2	9.27 ± 0.0159	7.53 ± 0.0128	0.0205	0.0189	1.74 ± 0.85**

[a] The two clones with the highest overall divergence value within the species.
[b] The two clones with the lowest overall divergence value within the species.
* $P < 0.1$, ** $P < 0.05$ (t-test, $n = 642$).

tion *Moutan,* and the ranges in divergence between species of these two groups do not overlap (table 3). The average divergence value for section *Paeonia* is intermediate, but the range of divergence overlaps with those of the other two sections. Our relative-rate test comparing divergences between section *Oneapia* and section *Moutan* included species of the two sections from which we had cloned both *Adh1A* and *Adh2* genes; the highest and lowest divergence values within each species are included in the test (table 4). The results indicate that the highest divergence value for *P. californica* is not significantly higher than the divergence values of the paralogous loci in *P. lutea* and *P. suffruticosa* subsp. *spontanea,* but both the higher and lower divergence values in *P. californica* are higher ($P < 0.1$ or 0.05) than those of *P. rockii* and *P. szechuanica.* Divergence of the two genes was not compared by the relative-rate test between section *Paeonia* and the other two sections owing to the overlap in the ranges of divergence.

Discussion

Our discovery that there are two major paralogous genes, *Adh1A* and *Adh2,* present in peonies and an additional gene, *Adh1B,* in the shrubby species is in agreement with the findings that two to three *Adh* loci usually exist in angiosperms (Gottlieb 1982; Morton, Gaut, and Clegg 1996). Identification of two major loci, *Adh1A* and *Adh2,* in *P. californica* is also concordant with the results of enzyme electrophoresis for this species (Zona et al. 1991).

The phylogenetic analysis of five *Adh* exon sequences of peonies along with a variety of other angiosperms (fig. 3) suggests that the duplication giving rise to *Adh1* and *Adh2* occurred prior to the diversification of *Paeonia.* However, the fact that peony *Adh1* and *Adh2* genes form a single clade in figure 3 implies that this duplication does not predate the diversification of eudicots or of all angiosperms. This confirms that there have been a number of separate duplication events within angiosperms (e.g., in the peony lineage, with grasses, etc.; see Morton, Gaut, and Clegg 1996) and cautions against superficial comparisons of *Adh* genes across angiosperms (e.g., genes labeled *Adh1* in different groups may not be homologous).

Based on our phylogenetic analyses, the duplication giving rise to *Adh1A* and *Adh1B* occurred later from an ancestral *Adh1* gene member. Finding *Adh1B* only in section *Moutan* leads to two alternative explanations: (1) a duplication of *Adh1A* and *Adh1B* genes occurred only in the ancestral lineage of section *Moutan*; or (2) a duplication of *Adh1A* and *Adh1B* genes occurred in the ancestor of the entire peony clade followed by subsequent loss of the *Adh1B* locus from the other two sections. Our finding that *Adh1B* genes form the sister group of all *Adh1A* genes (figs. 3 and 4) is best explained by the second hypothesis. Sequence divergence data also appear to support this explanation. Under the first hypothesis we would expect sequence divergence between *Adh1A* and *Adh1B* within section *Moutan* to be lower than the divergence of *Adh1A* genes between section *Moutan* and either of the other two sections unless *Adh1B* genes evolved very rapidly in section *Moutan,* perhaps under selection. The observed sequence divergence between *Adh1A* and *Adh1B* within section *Moutan* (averaging 5.30%), however, is much higher than *Adh1A* sequence divergence between this section and the other two sections (averaging 1.54%). Synonymous substitutions are almost 10 times higher than nonsynonymous substitutions between *Adh1A* and *Adh1B* genes in section *Moutan* (table 2), which suggests that strong positive selection has not been a major factor during the evolution of *Adh1B* genes (Hughes and Nei 1988, 1989). Taken together, these arguments favor the hypothesis of an early duplication followed by a loss (fig. 8).

Using sequences of low-copy nuclear gene families as phylogenetic markers provides an opportunity for rooting trees when outgroups are unavailable or very highly diverged. An example of this application is the rooting of the entire tree of life (Gogarten et al. 1989; Iwabe et al. 1989; Brown and Doolittle 1995; Lawson, Charlebois, and Dillon 1996). Based on Analysis I, it is clear that peony *Adh1* and *Adh2* genes are more closely related to each other than they are to any other angiosperm *Adh* genes included in the analysis. When the root

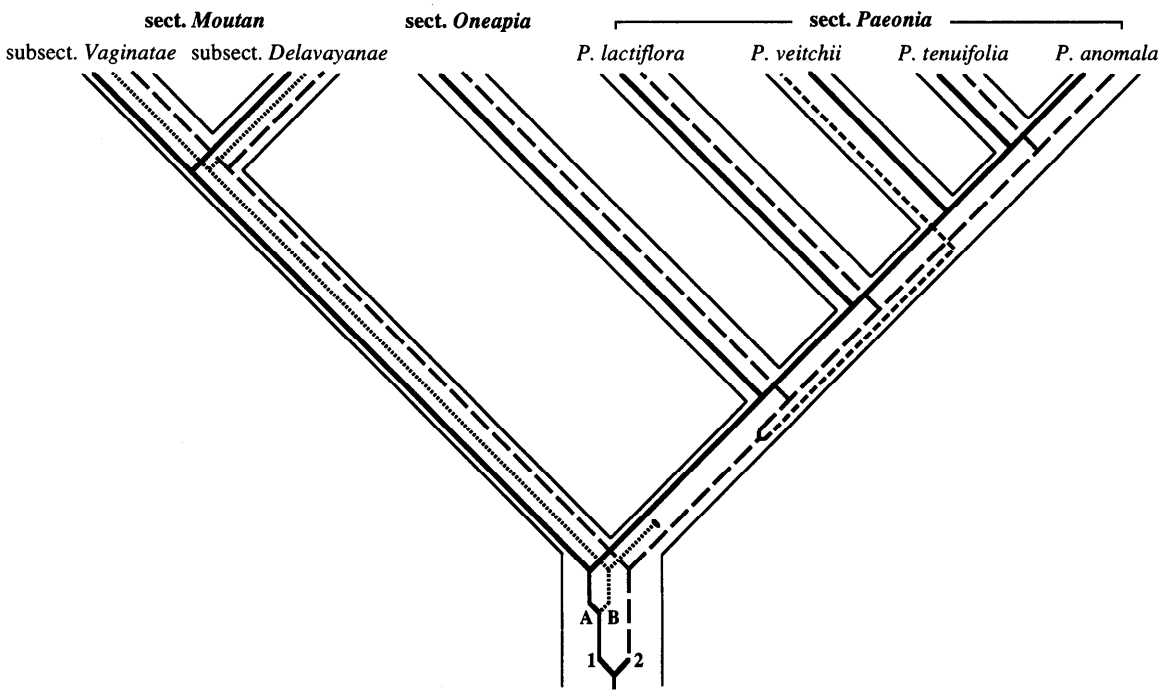

FIG. 8.—Peony species and *Adh* gene trees. The species tree is outlined by thin solid lines and represents all four studied diploid species of section *Paeonia* and two subsections of section *Moutan*; species within sections *Oneapia* and *Mouton* are not distinguished because their positions are concordant on all gene trees (figs. 5 and 6). Phylogenies of the *Adh1A, Adh1B,* and *Adh2* genes are represented by thick solid lines, dotted lines, and dashed lines, respectively. Duplications leading to *Adh1* (1) and *Adh2* (2) and to *Adh1A* (A) and *Adh1B* (B) are shown at the base of the species tree.

is positioned between *Adh1* and *Adh2* sequences in Analysis II, a monophyletic group containing sections *Oneapia* and *Paeonia* is obtained in the *Adh1A* clade, implying that the root of the genus lies between section *Moutan* and the lineage containing the other two sections. Although the same result was not supported above the 50% bootstrap level within the *Adh2* portion of the tree, support for the monophyly of *Oneapia* and *Paeonia* is higher than other possible relationships among the three sections. The observation that sectional relationships are less well resolved for *Adh2* could be a function of longer coalescence times for *Adh2* genes than for *Adh1A* genes (see below).

The view that the first evolutionary split occurred between section *Moutan* and the rest of the genus accords well with our hypothesis of a duplication followed by a deletion of *Adh1B* from sections *Oneapia* and *Paeonia.* When the peony tree is rooted between section *Moutan* and the other two sections, only one deletion of *Adh1B* is required; that is, in the common ancestor of sections *Oneapia* and *Paeonia.* Other rootings would require independent deletions of *Adh1B* from section *Oneapia* and section *Paeonia.*

Better understanding of the early evolution of *Paeonia* provides insights into morphological evolution and biogeography. The previous hypothesis of rooting—between the New World section *Oneapia* and the Old World sections—implied that the first evolutionary split within the genus may have been coupled with an intercontinental disjunct distribution, and that the herbaceous habit may well have been the ancestral condition from which the shrubby habit of section *Moutan* was derived. The new view of rooting—between the woody section and the two herbaceous sections—allows that the ancestral habit of peonies may have been shrubby (depending on the habit of close outgroups) and implies that the Eurasia–North America disjunction occurred during early diversification of the herbaceous lineage.

Sequence divergence among the sections, however, appears to conflict with our rooting hypothesis if a molecular clock is assumed. Sequence divergence of *Adh1A* genes is highest between section *Oneapia* and the other two sections (average 2.03%) and lowest between section *Moutan* and the other two sections (average 1.55%). For *Adh2* genes, the lowest divergence value is also between section *Moutan* and the other sections (average 2.27%), while the highest average sequence divergence is found between section *Paeonia* and the others (2.31%). These results are concordant with those based on ITS and *mat*K coding sequences and the *psb*A-*trn*H intergenic spacer region of cpDNA. That is, studies of these sequences also show the lowest sequence divergence between section *Moutan* and other two sections and the highest divergence between section *Oneapia* and the others (Sang, Crawford, and Stuessy 1997).

If the first evolutionary split within peonies was between section *Moutan* and the other two sections, it would appear that rates of evolution either decreased in section *Moutan* or increased in the other sections, particularly section *Oneapia.* Relative rates of molecular divergence between sections *Moutan* and *Oneapia* can be gauged by examining the divergence of *Adh1A* and *Adh2* in each section. If divergence rates have been the same in the two sections, and if concerted evolution has

not occurred, then the divergence of the two paralogous genes should be the same in each section (Wu and Li 1985). We used a relative-rate test based on the divergence of paralogous genes, *Adh1A* and *Adh2,* to assess whether paralogous genes diverged significantly more rapidly in section *Oneapia* than in section *Moutan.* The results indicate that rates have been slower in at least a part of section *Moutan* than in section *Oneapia* given that concerted evolution between *Adh1A* and *Adh2* loci is not detected. Although it is not possible to carry out a similar relative-rate test for ITS or *mat*K, lower sequence divergence in section *Moutan* appears to be correlated among nrDNA, cpDNA, and *Adh* genes. Such a correlation has been found in monocots, where synonymous substitution rates of both *Adh* genes and the chloroplast *rbc*L gene appear to be lower in palms than in grasses (Gaut et al. 1996). The longer generation time of palms was postulated by Gaut et al. (1996) to be responsible for their slower rate of molecular evolution. Generation time could also explain differences in rate within peonies, as the shrubby species of section *Moutan* are likely to have longer generation times than the herbaceous species of the other two sections.

Based on the foregoing conclusions, separate phylogenetic analyses of *Adh1A* and *Adh2* genes including both exons and introns were carried out (Analysis III), and the individual trees were rooted between section *Moutan* and the other two sections. Relationships within *Paeonia* on each resulting gene tree are better resolved and, based on the Templeton test, generally congruent with each other and with the ITS and *mat*K trees, except that section *Paeonia* is paraphyletic in *Adh2* trees. The unusual position of *P. veitchii* on the *Adh2* phylogeny may have been caused by either hybridization or random sorting of ancestral alleles. The hybridization hypothesis is less likely here given that no polymorphism has been found at the nrDNA, *Adh1A,* or *Adh2* loci in *P. veitchii,* and the phylogenetic positions of this diploid species do not conflict among the ITS, cpDNA, and *Adh1A* trees. Although additional independent nuclear loci need to be examined to rule out the hybridization hypothesis, the lineage-sorting hypothesis is currently favored. It is possible, however, that the ancestral allele cloned from *P. veitchii* is still maintained in all or some other species of section *Paeonia* but was not discovered due to the rarity of the allele and/or the limited intraspecific sampling.

The tree resulting from the combined analysis of *Adh1A,* ITS, and *mat*K sequences (fig. 7*A*) correlates better with morphology and the results of artificial hybridization experiments than the tree generated by the combined analysis of all four data sets (fig. 7*B*). *Paeonia lactiflora* has relatively broad leaflets and is placed in subsection *Foliolatae,* while the remaining three species have narrow leaflets and are placed in subsection *Paeonia* (Stern 1946). Artificial hybrids can easily be obtained from crosses among *P. anomala, P. veitchii,* and *P. tenuifolia* but are very difficult to obtain from hybridization between any of these three species and *P. lactiflora* (Saunders and Stebbins 1938). The sister group relationship of *P. anomala* and *P. tenuifolia* is supported by the production of only one terminal flower per stem (Stern 1946). *Paeonia lactiflora,* in contrast, has more than one flower on each stem; this is likely to be the ancestral state because it also occurs in section *Oneapia* and subsection *Delavayanae* of section *Moutan. Paeonia veitchii* seems to be intermediate on this character; some populations have one terminal flower and aborted flower buds on side branches (Pan 1979). Owing to the congruence of the tree based on the three data sets (fig. 7*A*) with other data we consider this to be the current best estimate of the species phylogeny, and on this basis we have reconstructed the *Adh* gene phylogenies shown in figure 8. Furthermore, in the case of considering the tree based on the four data sets to be the species phylogeny, a lineage-sorting hypothesis still has to be invoked to account for the paraphyly of section *Paeonia* on the *Adh2* phylogeny, and an additional hypothesis is needed to explain the incongruent relationships of *P. lactiflora* and *P. veitchii* between the *Adh1* phylogeny and this hypothetical species phylogeny.

In comparison with the *Adh1* tree, wherein alleles form monophyletic groups within each species, *Adh2* alleles are not directly linked within five species or even within section *Paeonia* (fig. 5*A* and *B*). *Adh2* genes appear to have longer coalescence times than *Adh1A* genes, which might also explain the higher divergence at synonymous sites in *Adh2* (table 3). It is possible that longer coalescence times in *Adh2* are a function of there being a larger number of copies than there are of *Adh1A* (Kreitman 1991; Moore 1995). This is suggested by the observation that we found a larger number of distinct clones of the *Adh2* gene in the majority of peony species. The number of distinct clones, however, may not precisely reflect the number of gene copies in the genome, because we cannot distinguish whether two *Adh2* clones represent recently duplicated loci or heterozygous alleles at a single locus. Furthermore, small sequence differences among clones may be caused by *Taq* polymerase errors during PCR. Nevertheless, cloning both a normal *Adh2* gene and a pseudogene from the same genome of *P. delavayi* and *P. suffruticosa* subsp. *spontanea* provides evidence for the previous existence of more than one *Adh2* locus in these two species. Of the three *Adh2* pseudogenes cloned in this study, two (LAC-1 and DEL-10) are highly divergent from the normal *Adh2* gene (fig. 5*B*), indicating that their origins may be quite ancient. In contrast, clone SPO-2 is even less diverged than the normal *Adh2* clone of the same plant (SPO-9), which implies a recent origin of the pseudogene or a PCR error. Variation in copy numbers among different genes, and perhaps among species (as well as the occurrence of pseudogenes), is characteristic of the evolution of multigene families (Morton, Gaut, and Clegg 1996). In view of this dynamism, it is interesting to note that the copy number of the *Adh* gene family appears to be relatively stable in flowering plants, where two to three loci are usually found (Gottlieb 1982; Morton, Gaut, and Clegg 1996).

Phylogenetic inference at lower taxonomic levels (e.g., the interspecific level) in plants may be complicated by lineage sorting, hybridization, and/or limited resolu-

tion. A number of independent gene trees, therefore, are needed to represent the species tree (Pamilo and Nei 1988; Wu 1991; Hudson 1992; Maddison 1995). With respect to lineage sorting, it has been argued that the phylogeny of maternally inherited mitochondrial DNA has better chance of being congruent with the species tree than a nuclear gene phylogeny owing to a smaller effective population size and therefore shorter coalescence times (Moore 1995). The same argument may apply to chloroplast DNA phylogenies of the majority of flowering plants whose cpDNA is maternally inherited (Corriveau and Cloeman 1988; Mogensen 1996). Unlike mitochondrial DNA, however, sequences of cpDNA have rarely been used for phylogenetic inference at the species level owing to limited variation. For example, the *mat*K coding region, which is the most rapidly evolving coding sequence known in cpDNA (Olmstead and Palmer 1994), seems to have evolved more slowly than *Adh* exons have in peonies (average divergence of 0.85% for the *mat*K sequences of the 11 species; compare table 3). The noncoding regions of cpDNA also appear to be evolutionarily conserved and provided very limited resolution of interspecific relationships in *Paeonia* (Sang, Crawford, and Stuessy 1997). ITS sequences are now widely used phylogenetic markers at lower taxonomic levels in plants but also show limited variation in many cases (Baldwin et al. 1995).

These observations highlight the need to extend the sampling of low-copy-number nuclear genes in phylogenetic studies, particularly at lower taxonomic levels. It is noteworthy that the *Adh1A* gene tree for *Paeonia* is better resolved and better supported than the ITS and *mat*K phylogenies (figs. 5 and 6). This may be due largely to the variable introns of the *Adh* genes, which are more variable than ITS sequences among the same species (average divergence of 3.11% for ITS; compare table 3). Likewise, Gottlieb and Ford (1996) demonstrated the phylogenetic utility of *PgiC* gene sequences in *Clarkia* (Onagraceae). Although it may be difficult in practice to identify paralogous and orthologous sequences, the extra effort required to make use of nuclear gene families may often be worth the effort. Knowledge of duplications in the evolutionary history of a gene can be useful in rooting phylogenetic trees and in assessing relative rates of molecular evolution.

Acknowledgments

We thank M. Sanderson for help in designing the relative-rate test, A. Davelos, A. Jarosz, and J. Smith for discussing the statistical tests, W. Qu for help with several calculations, and D. Ferguson, D. Hibbett, E. Pine, R. Mason-Gamer, and S. Mathews for advice on cloning and sequencing methods. D. Hibbett and two anonymous reviewers provided helpful comments on the manuscript. T.S. was supported by a Mercer Postdoctoral Fellowship from the Arnold Arboretum of Harvard University.

LITERATURE CITED

ARNHEIM, N. 1983. Concerted evolution of multigene families. Pp. 38–61 *in* M. NEI and R. K. KOEHN, eds. Evolution of gene and protein. Sinauer, Sunderland, Mass.

BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL, and M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. **82**:247–277.

BRADLEY, R. D., and D. M. HILLIS. 1997. Recombinant DNA sequences generated by PCR amplification. Mol. Biol. Evol. **14**:592–593.

BROWN, J. R., and W. F. DOOLITTLE. 1995. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. Proc. Natl. Acad. Sci. USA **92**: 2441–2445.

CHANG, C., and E. M. MEYEROWITZ. 1986. Molecular cloning and DNA sequence of the *Arabidopsis thaliana* alcohol dehydrogenase. Proc. Natl. Acad. Sci. USA **83**:1408–1412.

CHASE, M. W., D. E. SOLTIS, R. G. OLMSTEAD et al. (42 coauthors). 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. Mo. Bot. Gard. **80**: 528–580.

CORRIVEAU, J. L., and A. W. CLOEMAN. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. Am. J. Bot. **75**:1443–1458.

CRONQUIST, A. 1988. The evolution and classification of flowering plants. Allen Press, Lawrence, Kans.

DENNIS, E. S., M. M. SACHS, W. L. GERLACH, E. J. FINNEGAN, and W. J. PEACOCK. 1985. Molecular analysis of alcohol dehydrogenase 2 (*Adh2*) gene of maize. Nucleic Acids Res. **13**:727–743.

DE QUEIROZ, A., M. J. DONOGHUE, and J. KIM. 1995. Separate versus combined analysis of phylogenetic evidence. Annu. Rev. Ecol. Syst. **26**:657–681.

DOVER, G. 1982. Molecular drive: a cohesive mode of species evolution. Nature **299**:111–117.

DOYLE, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Syst. Bot. **17**:144–163.

DOYLE, J. J., and J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**:11–15.

ELLISON, N. W., P. L. YU, and D. W. WHITE. 1990. Nucleotide sequence of a white clover alcohol dehydrogenase cDNA. Nucleic Acids Res. **18**:4913.

FELSENSTEIN, J. 1985. Confidence limits on phylogenetics: an approach using the bootstrap. Evolution **39**:783–791.

GAUT, B. S., B. R. MORTON, B. C. MCGAIG, and M. T. CLEGG. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene *rbc*L. Proc. Natl. Acad. Sci. USA **93**:10274–10279.

GOGARTEN, J. P., H. KILBAK, P. DITTRICH, L. TAIZ, E. J. BOWMAN, B. J. BOWMAN, M. F. MANOLSON, R. J. POOLE, T. DATE, and T. OSHIMA. 1989. Evolution of vacuolar H+-ATPase: implications for the origin of eukaryotes. Proc. Natl. Acad. Sci. USA **86**:6661–6665.

GOTTLIEB, L. D. 1982. Conservation and duplication of isozymes in plants. Science **216**:373–380.

GOTTLIEB, L. D., and V. S. FORD. 1996. Phylogenetic relationships among the sections of *Clarkia* (Onagraceae) inferred from the nucleotide sequences of *PgiC*. Syst. Bot. **21**:45–62.

HILLIS, D. M. 1995. Approaches for assessing phylogenetic accuracy. Syst. Biol. **44**:3–16.

HUDSON, R. R. 1992. Gene trees, species trees, and the segregation of ancestral alleles. Genetics **131**:509–512.

HUELSENBECK, J. P., J. J. BULL, and C. W. CUNNINGHAM. 1996. Combining data in phylogenetic analysis. TREE **11**:152–158.

HUGHES, A. L. 1995. The evolution of the type I interferon gene family in mammals. J. Mol. Evol. **41**:539–548.

HUGHES, A. L., and M. NEI. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature **335**:167–170.

———. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc. Natl. Acad. Sci. USA **86**:958–962.

IWABE, N., K. KUMAM, M. HASEGAWA, S. OSAWA, and T. MIYATA. 1989. Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. USA **86**:9355–9359.

JEFFS, P. S., E. C. HOLMES, and M. ASHBURNER. 1994. The molecular evolution of the alcohol dehydrogenase and alcohol dehydrogenase-related genes in the *Drosophila melanogaster* species subgroup. Mol. Biol. Evol. **11**:287–304.

JUKES, T. H., and C. R. CANTOR. 1969. Evolution of protein molecules. Pp. 21–132 *in* H. N. MUNRO, ed. Mammalian protein metabolism. Academic Press, New York.

KEEFE, J. M., and M. F. MOSELEY JR. 1978. Wood anatomy and phylogeny of *Paeonia* section *Moutan*. J. Arnold Arbor. **59**:274–297.

KREITMAN, M. 1991. Detecting selection at the level of DNA. Pp. 204–221 *in* R. K. SELANDER, A. G. CLARK, and T. S. WHITTAM, eds. Evolution at the molecular level. Sinauer, Sunderland, Mass.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis. Version 1.02. The Pennsylvania State University, University Park.

LARSON, A. 1994. The comparison of morphological and molecular data in phylogenetic systematics. Pp. 371–390 *in* B. SCHIERWATER, B. STREIT, G. P. WAGNER, and R. DESALLE, eds. Molecular ecology and evolution: approaches and applications. Birkhauser Verlag, Basel.

LAWSON, F. S., R. L. CHARLEBOIS, and J.-A. R. DILLON. 1996. Phylogenetic analysis of carbamoylophosphate synthetase genes: complex evolutionary history includes an internal duplication within a gene which can root the tree of life. Mol. Biol. Evol. **13**:970–977.

LI, W.-H., and M. TANIMURA. 1987. The molecular clock runs more slowly in man than in apes and monkeys. Nature **326**: 93–96.

LLEWELLYN, D. J., E. J. FINNEGAN, J. G. ELLIS, E. S. DENNIS, and W. J. PEACOCK. 1987. Structure and expression of an alcohol dehydrogenase 1 gene from *Pisum sativum* (cv. "Greenfeast"). J. Mol. Biol. **195**:115–123.

MADDISON, W. P. 1995. Phylogenetic histories within and among species. Pp. 273–287 *in* P. C. HOCH and A. G. STEPHENSON, eds. Experimental and molecular approaches to plant biosystematics. Missouri Botanical Garden, St. Louis.

MASON-GAMER, R. J., and E. A. KELLOGG. 1996. Testing for phylogenetic conflict among molecular data sets in the tribe *triticeae* (Gramineae). Syst. Biol. **45**:524–545.

MELVILLE, R. 1983. The affinity of *Paeonia* and a second genus of Paeoniaceae. Kew Bull. **38**:87–105.

MIYASHITA, N. T., H. INNAN, and R. TERAUCHI. 1996. Intra- and interspecific variation of the alcohol dehydrogenase locus region in wild plants *Arabis gemmifera* and *Arabidopsis thaliana*. Mol. Biol. Evol. **13**:433–436.

MOGENSEN, H. L. 1996. The hows and whys of cytoplasmic inheritance in seed plants. Am. J. Bot. **83**:383–404.

MOORE, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution **49**:718–726.

MORTON, B. R., B. GAUT, and M. CLEGG. 1996. Evolution of alcohol dehydrogenase genes in the Palm and Grass families. Proc. Natl. Acad. Sci. USA **93**:11735–11739.

NURMINSKY, D. I., E. N. MORIYAMA, E. R. LOZOVAKAYA, and D. L. HARTL. 1996. Molecular phylogeny and genome evolution in the *Drosophila virilis* species group: duplication of the alcohol dehydrogenase gene. Mol. Biol. Evol. **13**: 132–149.

OKAMURO, J. K., and R. B. GOLDBERG. 1989. Regulation of plant gene expression: general principles. Biochem. Plants **15**:1–82.

OLMSTEAD, R. G., and J. D. PALMER. 1994. Chloroplast DNA systematics: A review of methods and data analysis. Am. J. Bot. **81**:1205–1224.

PAMILO, P., and M. NEI. 1988. Relationship between gene trees and species trees. Mol. Biol. Evol. **5**:568–583.

PAN, K.-Y. 1979. *Paeonia*. Pp. 37–59 *in* Flora reipublicae siniccae. Vol. 27. Science Press, Beijing.

RICE, K. A., M. J. DONOGHUE, and R. G. OLMSTEAD. 1997. Analyzing large data sets: *rbc*L 500 revisited. Syst. Biol. (in press).

RUSSO, C. A., N. TAKEZAKI, and M. NEI. 1995. Molecular phylogeny and divergence times of *Drosophilid* species. Mol. Biol. Evol. **12**:391–404.

SANDERSON, M. J., and J. J. DOYLE. 1992. Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. Syst. Biol. **41**:4–17.

SANG, T., D. J. CRAWFORD, and T. F. STUESSY. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using sequences of internal transcribed spacer of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proc. Natl. Acad. Sci. USA **92**:6813–6817.

———. 1997. Chloroplast phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). Am. J. Bot. (in press).

SAUDERS, A. P., and G. L. STEBBINS JR. 1938. Cytogenetic studies in *Paeonia* I. The compatibility of the species and appearance of the hybrids. Genetics **23**:65–111.

SOLTIS, D. E., L. A. JOHNSON, and C. LOONEY. 1996. Discordance between ITS and chloroplast topologies in the *Boykinia* group. Syst. Bot. **21**:169–186.

SOLTIS, D. E., and R. K. KUZOFF. 1995. Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). Evolution **49**:727–742.

SOLTIS, D. E., P. S. SOLTIS, D. L. NICKRENT et al. (15 co-authors). 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. Ann. Mo. Bot. Gard. (in press).

STERN, F. C. 1946. A study of the genus *Paeonia*. The Royal Horticultural Society, London.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony. Version 3.1.1. Illinois Natural HIstory Survey, Champaign.

TAKHTAJAN, A. 1969. Flowering plants: origin and dispersal. Oliver and Boyd, Edinburgh.

———. 1987. Systema Magnoliophytorum. Officina Editoria "Nauka," Leninopoli.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution **37**:221–244.

THORNE, R. F. 1992. Classification and geography of the flowering plants. Bot. Rev. **58**:225–348.

TRICK, M., E. S. DENNIS, K. J. R. EDWARDS, and W. J. PEACOCK. 1988. Molecular analysis of the alcohol dehydrogenase gene family of barley. Plant Mol. Biol. **11**:147–160.

TZANOUDAKIS, D. 1983. Karyotypes of four wild *Paeonia* species from Greece. Nord. J. Bot. **3**:307–318.

WOLYN, D. J., and G. JELENKOVIS. 1990. Nucleotide sequence of an alcohol dehydrogenase gene in octoploid strawberry (*Fragaria* × *Ananassa* Duch.). Plant Mol. Biol. **14**:855–857.

WU, C.-I. 1991. Inferences of species phylogeny in relation to segregation of ancient polymorphisms. Genetics **127**:429–435.

WU, C.-I., and W.-H. LI. 1985. Evidence for higher rates of nucleotide substitution in rodents than man. Proc. Natl. Acad. Sci. USA **82**:1741–1745.

XUE, J., M. JOGENSEN, U. PIHLGREN, and L. RASK. The myrosinase gene family in *Arabidopsis thaliana*: gene organization, expression and evolution. Plant Mol. Biol. **27**: 911–922.

YOKOYAMA, S., and D. E. HARRY. 1993. Molecular phylogeny and evolutionary rates of alcohol dehydrogenase in vertebrates and plants. Mol. Biol. Evol. **10**:1215–1226.

ZONA, S., L. RIESEBERG, R. L. CARTER, and K. DOAN. 1991. Isozyme diversity is low in *Paeonia californica* (Paeoniaceae). Plant Species Biol. **6**:89–93.

BARBARA A. SCHAAL, reviewing editor

Accepted June 23, 1997