

The Potential of *Betv1* Homologues, a Nuclear Multigene Family, as Phylogenetic Markers in Flowering Plants

Jun Wen,^{*,1} M. Vanek-Krebitz,[†] K. Hoffmann-Sommergruber,[†] O. Scheiner,[†] and H. Breiteneder^{†,2}

^{*} Department of Biology, Colorado State University, Fort Collins, Colorado 80523; and [†]Department of General and Experimental Pathology, University of Vienna, A-1090 Vienna, Austria

Received January 28, 1997; revised April 9, 1997

***Betv1* homologues are a ubiquitous group of genes in flowering plants encoding a class of highly conserved defense-related proteins and containing open reading frames from 465 to 480 bp. *Betv1*-like genes consist of two exons interrupted by an intron of 76–359 bp, with the intron position highly conserved. The pairwise *p* distance ranged from 0 to 0.583 among flowering plants. Within plant families, the ranges of the *p* distance were 0–0.403, 0–0.253, and 0.011–0.369, for Apiaceae, Betulaceae, and Fabaceae, respectively. The most striking feature of the *betv1* gene phylogeny was that the multiple sequences from each plant family formed a monophyletic group and sequences from each species were generally more similar to each other than those from other species. The almost exclusive paralogous relationships of genes from the same species suggested that the genes of the multigene family underwent strong concerted evolution. Phylogenies of Betulaceae and Fabaceae inferred from *betv1* gene trees were generally congruent with those based on morphology and other molecules. *Betv1* homologues constitute potential phylogenetic markers at the intrafamilial level or among closely related families in flowering plants.** © 1997 Academic Press

INTRODUCTION

Plant molecular systematics has developed rapidly in the past 10 years. At present most of the molecular

characters are from the chloroplast genes (e.g., *rbcL*, *matK* and *ndhF*) and nuclear ribosomal genes or regions (e.g., 18S, 26S, and the ITS regions). The clear distinction between gene trees and species trees (e.g., Nei, 1987; Li and Graur, 1991; Doyle, 1992) has led to the search for additional phylogenetic markers, especially those from the nuclear genome. Phylogenetic utility of several single- or low-copy nuclear genes has been explored (e.g., Mathews *et al.*, 1995; Russo *et al.*, 1995; Waters, 1995; Gottlieb and Ford, 1996; Mason-Gamer and Kellogg, 1996; Mathews and Sharrock, 1996).

Betv1 homologues are a ubiquitous group of genes and were identified from a wide range of flowering plants including both dicots and monocots. They encode a class of highly conserved defense-related proteins and contain open reading frames from 465 to 480 bp. Genomic sequences usually consist of two exons interrupted by an intervening sequence (*mdgc10* from *Malus* without intron, Hoffmann-Sommergruber *et al.*, 1997). The sizes of the intron vary from 76 bp for *PfYpr10* from *Pisum fulvum* (GenBank Accession Number U65425, S. Tewari, S. M. Brown, P. Kenyon, and B. Fristensky, unpublished) to 359 bp for *STH-2* from *Solanum tuberosum* (Matton *et al.*, 1993). The intron size variation of the *betv1* homologues is presented in Table 1. The position of the intron is conserved.

Bet v 1, the 17-kDa major allergenic protein of birch pollen, was originally identified by screening a pollen cDNA library with serum IgE from birch pollen allergic patients (Breiteneder *et al.*, 1989). Comparison of the nucleotide sequence of the cloned allergen with the sequences in the EMBL databank known at that time revealed a sequence similarity of 70% to a pea disease resistance response gene (Fristensky *et al.*, 1988). Since then, a growing new class of small intracellular defense-

¹ Address for correspondence: Department of Biology, Colorado State University, Fort Collins, CO 80523. Fax: (970) 491-0649. E-mail: jwen@lamar.ColoState.edu.

² Address for correspondence: Department of General and Experimental Pathology, University of Vienna, AKH-EBO-3Q, Waehringer Guertel 18-20, Vienna, 1090 Austria. Fax: +431-40400-5130. E-mail: Heimo.Breiteneder@akh-wien.ac.at.

TABLE 1
Intron Size Variation of Genomic Clones
of *Betv1* Homologues

Sequence	Species	Intron size (bp)	Reference
PfYpr10	<i>Pisum fulvum</i>	76	Tewari (unpublished)
DRRG49-c	<i>Pisum sativum</i>	88	Chiang and Hadwiger (1990)
gPcPR1-1	<i>Petroselinum crispum</i>	164	Somssich <i>et al.</i> (1988)
gPcPR2	<i>Petroselinum crispum</i>	195	van de Loecht <i>et al.</i> (1990)
STH-2	<i>Solanum tuberosum</i>	359	Matton <i>et al.</i> (1993)
STH-21	<i>Solanum tuberosum</i>	347	Matton <i>et al.</i> (1993)
gAoPR1	<i>Asparagus officinalis</i>	347	Warner <i>et al.</i> (1993)
cagc10	<i>Corylus avellana</i>	81	Hoffmann-Sommergruber <i>et al.</i> (1997)
cagc11	<i>Corylus avellana</i>	94	Hoffmann-Sommergruber <i>et al.</i> (1997)
mdgc32, mdgc36 and mdgc38	<i>Malus domestica</i>	111	Hoffmann-Sommergruber <i>et al.</i> (1997)
mdgc10	<i>Malus domestica</i>	0	Hoffmann-Sommergruber <i>et al.</i> (1997)

related proteins, different from all previously characterized pathogenesis-related proteins, has been described in several plant species. All members of this class of proteins are very similar in amino acid sequence, protein size, and total charge, clearly indicating a common ancestry and probably an identical or related, still unknown, biological function. They establish a ubiquitous new class of evolutionarily conserved pathogenesis- or defense-related proteins. The deduced amino acid sequences range from 154 to 160 residues, and the deduced molecular masses range from 16.6 to 17.5 kDa.

Members of this plant gene family have been found to be constitutively expressed in a variety of species under complex patterns of developmental regulation. The presence of mRNAs coding for Bet v 1 homologues was shown during early developmental phases of root hairs and root epidermal cells in pea (Mylona *et al.*, 1994), roots of soybean seedlings (Crowell *et al.*, 1992), cotyledons of germinating chickpea seeds (Cervantes *et al.*, 1995), and hypocotyledons and cotyledons of yellow lupine (Sikorski *et al.*, 1995). During late developmental phases, the accumulation of mRNAs coding for Bet v 1 homologous proteins was shown for pea in late seed development (Iturriaga *et al.*, 1994), for the yellow lupine in senescent nodules (Sikorski *et al.*, 1995), and for soybean in senescent leaves (Crowell *et al.*, 1992). Moreover, transcripts of genes belonging to the *betv1*

family were detected in mature pollen of trees of the order Fagales (Breiteneder *et al.*, 1989, 1992, 1993; Larsen *et al.*, 1992), celery (Breiteneder *et al.*, 1995), apple (Vanek-Krebitz *et al.*, 1995), and transgenic tobacco (Warner *et al.*, 1993).

On the other hand, transcripts were also found to be inducible by the interaction of the plants with bacteria, fungi, and fungal elicitors as was shown for pea (Fristensky *et al.*, 1988; Chiang and Hadwiger, 1990), alfalfa (Esnault *et al.*, 1993), yellow lupine (Sikorski *et al.*, 1995), and soybean (Crowell *et al.*, 1992), or by wounding as shown for asparagus (Warner *et al.*, 1992) and potato (Matton and Brisson, 1989). Treatment of cell-suspension cultures with fungal elicitors and bacteria also resulted in the induction of these genes in bean (Walter *et al.*, 1990), parsley (Somssich *et al.*, 1986), and birch (Swoboda *et al.*, 1995a,b).

The first indication of a potential function of these proteins came from studies done in the monocot *Asparagus officinalis*. A role was suggested for these proteins as enzymes of the core phenylpropanoid pathway that are involved in the synthesis of cell-wall structural molecules, pigments, UV-light protectants, chemical attractants, and antimicrobial phytoalexins (Warner *et al.*, 1994). Along the same lines, it was proposed by Yazaki *et al.* (1995) that the LEDI-1 protein from *Lithospermum erythrorhizon* might play a significant role in the induction of secondary metabolism, such as shikonin production.

The definition of *betv1* homologue is at present based on sequence similarity. As the most extensive research was done on the Fagales and *betv1*, the term *betv1* homologues (rather than selecting any other member of this group) was chosen. We predict that the function of these proteins, which is unknown to date, will tie them together. First evidence for a rather generalized function was reported recently based on the crystal structure of a Bet v 1 isoform (Gajhede *et al.*, 1996). It was shown that the Bet v 1 molecule contains a large forked cavity that runs through its structure with three openings on the surface. It was suggested that Bet v 1 might be a nucleotide-binding protein, because of a conserved amino acid residue motif in close vicinity to that cavity. This so-called P-loop motif is found in many nucleotide-binding proteins.

In this study, we examined the phylogenetic utility of the *betv1* homologous genes. Betulaceae and Fabaceae were relatively well represented and were therefore used to discuss the phylogenetic potential of *betv1* homologues.

MATERIALS AND METHODS

Sequences and Alignment

Sixty-seven gene sequences were obtained from GenBank and include nine genomic sequences and 58

cDNA sequences (Table 2). These sequences were selected based on their high level of overall similarity to *betv1* genes from *Betula verrucosa*. The proteins they code are very similar in amino acid sequence, protein size, and total charge, suggesting a common ancestry and probably an identical or related, still unknown, biological function. Three genomic sequences including STH-2, STH-21, and gAoPR1 (Table 1) were not included in the analyses because they contain the cDNA sequences pSTH-2, pSTH-21, and AoPR1, respectively. The 67 amino acid sequences were aligned using PILEUP (Devereux *et al.*, 1984) of the GCG program with a gap weight penalty of 5.000 and a gap length weight of 0.300 (see Appendix for the alignment). The aligned amino acid sequences were converted to nucleotide sequences, which were exported to PAUP (version 3.1, Swofford, 1993), PHYLIP (version 3.5, Felsenstein, 1993), and MEGA (version 1.01, Kumar *et al.*, 1993) for phylogenetic analysis. Several sequences are identical at the coding regions (e.g., Bet v 1i and Bet v 1f, and Bet v 1m and Bet v 1n), with a one to two basepair difference in the noncoding regions. The identical sequences of the coding regions were grouped together, reducing the number of the sequences in the phylogenetic analysis to 65. A brief review of the sequences used is provided below.

Betulaceae

Betula verrucosa. The first tree pollen allergen to be included in this class of defense-related proteins on the basis of sequence homology was Bet v 1. Bet v 1 isoforms, a group of highly similar polypeptides of 17 kDa, are constitutively expressed at high levels in mature pollen and together represent the major birch pollen allergen Bet v 1. A cDNA clone encoding the most abundant isoallergen, Bet v 1a, was isolated from a pollen expression library by immunoscreening with a birch pollen allergic patient's serum (Breiteneder *et al.*, 1989). Several additional isoforms were described and their presence in the natural mixture of isoallergens in birch pollen was confirmed (Swoboda *et al.*, 1995a). Isolation of cDNA clones from a cDNA library prepared from cultured birch cells grown in the presence of *Pseudomonas syringae* demonstrated that the *betv1* gene family of pollen allergens includes a subset of genes, different from those constitutively expressed in pollen, that are induced in somatic cells upon their interaction with microorganisms (Swoboda *et al.*, 1995b).

All sequences of *betv1* isoforms were obtained from cDNA synthesized on mRNA isolated from birch pollen. The source of the pollen was a commercial provider for pollen (Allergon, Sweden). The pollen was checked by electron microscopy for purity. In the case of birch

pollen, the material was collected from a plantation of *B. verrucosa* trees in Sweden. The same applies for the alder, hazel, and hornbeam sequences. Genomic sequences were obtained from DNA isolated from leaves of a single Austrian *B. verrucosa* tree (as was done for hazel and apple).

Alnus glutinosa. Aln g 1 was isolated and characterized on the basis of its allergenic potential from pollen mRNA of alder (Breiteneder *et al.*, 1992).

Carpinus betulus. Studies of the major allergen from hornbeam pollen revealed the presence of isogenic variation. The Car b 1 isoforms were characterized by antibody-based assays (Larsen *et al.*, 1992).

Corylus avellana. Cor a 1, the major allergen of hazel pollen, consists of several isoforms, four of which have been studied in detail (Breiteneder *et al.*, 1993).

Apiaceae

Apium graveolens. Api g 1, a major allergen from celery, was identified by its reactivity with IgE from individuals displaying Type I allergic symptoms after the ingestion of celery (Breiteneder *et al.*, 1995).

Daucus carota. The sequence of a pathogenesis-related protein encoding mRNA from carrot somatic embryos was deposited in the EMBL database (Accession Number U47087, X. Liu, G. J. Hwang, and J. L. Zimmerman, unpublished).

Petroselinum crispum. Administration of a cell-wall preparation from the fungus *Phytophthora megasperma* f. sp. *glycinea* in cell-suspension cultures of parsley resulted in rapid and dramatic increase of mRNAs coding for PcPR1 and PcPR2. Regulation of the PR protein synthesis occurred at the transcriptional level (Somssich *et al.*, 1986). Genomic DNA blot experiments indicated that PR1 is a small gene family consisting of approximately three to six genes per haploid genome (Somssich *et al.*, 1988). A single-copy gene encoding pathogenesis-related protein 2 (PR2) in parsley was identified by its transcriptional activation in cultured parsley cells treated with fungal elicitor. The deduced PR2 protein belonged to a class of evolutionarily conserved polypeptides which are closely related to disease resistance in plants. (van de Loecht *et al.*, 1990).

Boraginaceae

Lithospermum erythrorhizon. cDNAs expressed predominantly in dark-grown *Lithospermum* cell cultures were enriched by a cDNA subtraction method. The expression of LEDI-1 was strongly induced in *Lithospermum* cells under aseptic conditions without fungal elicitation. The gene product may play a significant role

TABLE 2
Betv1 Homologous Sequences Used in This Study

Taxon	Sequence designation	Sequence type	Genbank accession no.		Reference
			DNA	Protein	
Betulaceae					
Alnus glutinosa	Aln g 1	cDNA	S50892	P38948	Breiteneder et al. (1992)
Betula verrucosa	Bet v 1a	cDNA	X15877	P15494	Breiteneder et al. (1989)
	Bet v 1b	cDNA	X77200	P45431	Swoboda et al. (1995a)
	Bet v 1c	cDNA	X77265	P43176	Swoboda et al. (1995a)
	Bet v 1d	cDNA	X77266	P43177	Swoboda et al. (1995a)
	Bet v 1e	cDNA	X77267	P43178	Swoboda et al. (1995a)
	Bet v 1f	cDNA	X77268	P43179	Swoboda et al. (1995a)
	Bet v 1g	cDNA	X77269	P43180	Swoboda et al. (1995a)
	Bet v 1h	cDNA	X77270	P43177	Swoboda et al. (1995a)
	Bet v 1i	cDNA	X77274	P43179	Swoboda et al. (1995a)
	Bet v 1j	cDNA	X77271	P43283	Swoboda et al. (1995a)
	Bet v 1k	cDNA	X77272	P43184	Swoboda et al. (1995a)
	Bet v 1l	cDNA	X77273	P43185	Swoboda et al. (1995a)
	Bet v 1m	cDNA	X81972	P43186	Swoboda et al. (1995a)
	Bet v 1n	cDNA	X82028	P43186	Swoboda et al. (1995a)
	Bet v 1-Sc1	cDNA	X77599		Swoboda et al. (1995b)
	Bet v 1-Sc2	cDNA	X77600		Swoboda et al. (1995b)
	Bet v 1-Sc3	cDNA	X77601		Swoboda et al. (1995b)
Carpinus betulus	Car b 1a	cDNA	X66932	P38949	Larsen et al. (1992)
	Car b 1b	cDNA	X66918	P38948	Larsen et al. (1992)
	Car b 1/2	cDNA	X66933	P28959	Larsen et al. (1992)
Corylus avellana	Cor a 1a	cDNA	X70999		Breiteneder et al. (1993)
	Cor a 1b	cDNA	X71000		Breiteneder et al. (1993)
	Cor a 1c	cDNA	X70997		Breiteneder et al. (1993)
	Cor a 1d	cDNA	X70998		Breiteneder et al. (1993)
	cagc10	genomic	Z72439		Hoffmann-Sommergruber et al. (1997)
	cagc11	genomic	Z72440		Hoffmann-Sommergruber et al. (1997)
Apiaceae					
Apium graveolens	Api g 1	cDNA	Z48967	P49372	Breiteneder et al. (1995)
Daucus carota	Gea 20	cDNA	U47087		Liu (unpublished)
Petroselinum crispum	PcPR1-2	cDNA	X12574	P19417	Somssich et al. (1988)
	PcPR1-3	cDNA	X12573	P19418	Somssich et al. (1988)
	PcPR2	cDNA	X58698	P27538	van de Loecht et al. (1990)
	gPcPR2	genomic	X55736		van de Loecht et al. (1990)
Boraginaceae					
Lithospermum erythrorhizon	LEDI-1	cDNA	D38493		Yazaki et al. (1995)
Fabaceae					
Cicer arietinum	CaPR-10a	cDNA	X79706		Cervantes et al. (1995)
Glycine max	SAM22	cDNA	X60043	P26987	Crowell et al. (1992)
	H4	cDNA	X60044		Crowell et al. (1992)
Lupinus luteus	LIR18A	cDNA	X79974		Sikorski et al. (1995)
	LIR18B	cDNA	X79975		Sikorski et al. (1995)
Medicago sativa	SRG1	cDNA	U42752		Truesdell (unpublished)
	MsPR10	cDNA	U13709		Esnault et al. (1993)
Medicago truncatula	MtYpr10	cDNA	Y08641		Gamas (unpublished)
Phaseolus vulgaris	PvPR1	cDNA	X61365	P25985	Walter et al. (1990)
	PvPR2	cDNA	X61364	P25986	Walter et al. (1990)
	Ypr10	cDNA	X96999		Walter et al. (1996)
Pisum elatius	PePR10	cDNA	U57064		Tewari (unpublished)
Pisum fulvum	PfYpr10	genomic	U65425		Tewari (unpublished)
Pisum sativum	pI49	cDNA	X13383		Fristensky et al. (1988)
	pI176	cDNA	X18249		Fristensky et al. (1988)
	DRRG49-c	genomic	J03680	P27047	Chiang and Hadwiger (1990)
	ABR17	cDNA	Z15128		Iturriaga et al. (1994)
	ABR18	cDNA	Z15127		Iturriaga et al. (1994)
	RH2	cDNA	S74512		Mylona et al. (1994)
Vigna unguiculata	VuPR4.2	cDNA	X98608		Lan (unpublished)

TABLE 2—Continued

Taxon	Sequence designation	Sequence type	Genbank accession no.		Reference
			DNA	Protein	
Rosaceae					
<i>Malus domestica</i>	Mal d 1	cDNA	Z48969	P43211	Vanek-Krebitz <i>et al.</i> (1995)
	Mal d 1.2	cDNA	X83672	P43211	Schoening (unpublished)
	Ap15	cDNA	L42952		Atkinson <i>et al.</i> (1996)
	mdgc10	genomic	Z72425		Hoffmann-Sommergruber <i>et al.</i> (1997)
	mdgc32	genomic	Z72426		Hoffmann-Sommergruber <i>et al.</i> (1997)
	mdgc36	genomic	Z72427		Hoffmann-Sommergruber <i>et al.</i> (1997)
	mdgc38	genomic	Z72428		Hoffmann-Sommergruber <i>et al.</i> (1997)
<i>Prunus avium</i>	Pru a 1	cDNA	U66076		Scheurer <i>et al.</i> (unpublished)
Solanaceae					
<i>Solanum tuberosum</i>	pSTH-2	cDNA	M25155	P17642	Matton and Brisson (1989)
	pSTH-21	cDNA	M25156	P17641	Matton and Brisson (1989)
Liliaceae					
<i>Asparagus officinalis</i>	AoPR1	cDNA	X62103		Warner <i>et al.</i> (1992)
	AocPR2A	cDNA	X69652		Warner (unpublished)
	AocPR3A	cDNA	X69653		Warner (unpublished)

in the induction of secondary metabolism, such as shikonin production (Yazaki *et al.*, 1995).

Fabaceae

Cicer arietinum. CaPR-10a was isolated from a cDNA library made from cotyledons of germinating chickpea seeds. Southern analysis of chickpea genomic DNA revealed that the corresponding gene was single-copy. CaPR-10a mRNA was first detected in Northern blots at the time of desiccation in seed development. High transcript levels were maintained in the dry seed and early in germination. Transcript levels increased severalfold upon exposure of imbibed seeds to ABA (Cervantes *et al.*, 1995).

Glycine max. Stress-induced, developmentally regulated soybean genes were investigated. The mRNAs corresponding to these cDNAs, called SAM22 and H4, accumulated predominantly in the roots of soybean seedlings and were present at high levels in the roots and leaves of mature plants. SAM22 accumulation was especially dramatic in senescent leaves. In addition, SAM22 accumulation could be induced in young leaves by wounding or by transpiration-mediated uptake of salicylic acid, methyl viologen, fungal elicitor (chitosan), hydrogen peroxide, or sodium phosphate (pH 6.9). Southern blot analysis indicated that approximately 10 copies of sequences related to SAM22 existed in the soybean genome (Crowell *et al.*, 1992).

Lupinus luteus. Two intracellular pathogenesis-related proteins, L1R18A and L1R18B, from yellow lupine were identified through the study of the development of nitrogen-fixing nodules and the characterization of transcriptionally regulated host plant genes. A significant increase of their expression was observed in

senescent nodules. The accumulation of transcripts of both genes was observed at high levels in the roots of uninfected plants and hypocotyls, cotyledons, and leaves of mature plants infected with *B. lupini* (Sikorski *et al.*, 1995).

Medicago sativa. Accumulation of a transcript encoding the 17-kDa PR MsP10 protein was observed 6 to 30 h after alfalfa leaves were infiltrated with *Pseudomonas syringae* pv. *pisi* (Esnault *et al.*, 1993). Sequence of another *M. sativa* stress response gene 1 (SRG1) mRNA was deposited in the EMBL database (Accession Number U42752; G. M. Truesdell and M. B. Dickmann, unpublished).

Phaseolus vulgaris. Rapidly induced small transcripts were isolated from a cDNA library derived from bean cell-suspension cultures treated with an elicitor from *Colletotrichum lindemuthianum*. Southern blot analysis indicated that PvPR protein gene organization is highly complex with an estimated copy number of more than 12 genes (Walter *et al.*, 1990). Another pathogenesis-related protein Ypr10 from *P. vulgaris* was shown to display complex patterns of developmental, dark-induced, and exogenous stimulus-dependent expression (Walter *et al.*, 1996).

Pisum sativum. The expression of the Group I disease resistance response genes (pI49, pI1276, and pI206) in endocarp tissue of *P. sativum* cv. Alaska was strongly associated with resistant interactions of the pea with the fungus *Fusarium solani*, the bacterium *Pseudomonas syringae* pv. *pisi*, and the fungal elicitor chitosan (Fristensky *et al.*, 1985; Riggleman *et al.*, 1985; Daniels *et al.*, 1987; Fristensky *et al.*, 1988).

The accumulation of disease resistance response

gene (DRRG) mRNAs in pea endocarp tissue occurred 2–4 h after treatment with chitosan or inoculation with the incompatible pathogen *F. solani* f. sp. *phaseoli*. A pea genomic clone DRRG49-c encompassing one DRRG structural gene was obtained using the homologous pI49 cDNA as a probe. The DRRG structural gene is one of a multigene family, and an estimated five copies exist within the pea genome (Chiang and Hadwiger, 1990).

cDNAs encoding two low molecular weight abscisic acid-responsive pea proteins, ABR17 and ABR18, which were synthesized during late seed development *in vivo*, were isolated and characterized. They were induced by culture of embryos *in vitro* and their production markedly enhanced by exogenous ABA. The ABR17 cDNA corresponds to a single-copy gene, but ABR18 is one member of a family of closely related sequences in the pea genome (Iturriaga *et al.*, 1994).

Two-dimensional gel electrophoresis of pea root and root hair proteins revealed the existence of at least 10 proteins present at elevated levels. One of these, RH2, was expressed not only in root hairs but also in root epidermal cells lacking these tubular outgrowths. The expression of *rh2* was developmentally regulated. The authors postulated that *rh2* contributed to a constitutive defense barrier in the root epidermis (Mylona *et al.*, 1994).

Vigna unguiculata. The sequence of the gene coding for PR4.2 from the cowpea was deposited in the EMBL database (Accession Number X98608; V. T. T. Lan, W. J. Broughton, and A. Krause, unpublished). PR4.2 was identified as a pathogenesis-related protein during the *Rhizobium*–legume interaction.

Rosaceae

Malus domestica. Mal d 1, the major allergen from apple, was identified and characterized by the use of allergic individuals' sera and its immunological relationship to Bet v 1 demonstrated by cross-inhibition experiments using recombinant Mal d 1 (Vanek-Krebitz *et al.*, 1995). The screening of a ripe apple cDNA expression library with possibly cross-reacting polyclonal antibodies to spinach sucrose-phosphate synthase resulted in the isolation of cDNA clone pAP15 (Atkinson *et al.*, 1996). Genomic Southern analysis suggested that the gene belonged to a family of at least 15 members. Northern analysis indicated that pAP15 was ripening-related, with an accumulation of sequence-related mRNA coincident with ethylene production during apple fruit ripening (Atkinson *et al.*, 1996). A sequence coding for another isoform of Mal d 1 was deposited in the EMBL database (Accession Number X83672; B. Schoening, W. H. Ziegler, S. Viehls, and W. Baltes, unpublished).

Prunus avium. The sequence of a cDNA coding for the *betv1* homologue from sweet cherry was submitted to the EMBL database (Accession Number U66076; S. Scheurer, K. Metzner, D. Hausteiner, and S. Viehls, unpublished). The gene product was designated as cherry allergen Pru a 1.

Solanaceae

Solanum tuberosum. Potato pSTH-related transcripts were found to be rapidly induced by slicing tubers or stems as well as by the addition of arachidonic or eicosapentaenoic acid to tuber disks or by treatment of leaves with *Phytophthora infestans* mycelial homogenate (Matton and Brisson, 1989).

Liliaceae

Asparagus officinalis. A wound-induced cDNA designated AoPR1 from a suspension of mesophyll cells isolated from cladodes of light-grown *A. officinalis* seedlings was reported. The transcript abundance is up-regulated following cell separation and in chopped mesocotyl tissue from dark-grown seedlings. This was also the first example of a monocot cDNA belonging to this class of intracellular pathogenesis-related proteins (Warner *et al.*, 1992).

Cis-acting regulatory sequences of the AoPR1 gene were isolated. A translational fusion was constructed with the β -glucuronidase (GUS) reporter gene and tested in transgenic tobacco. Histochemical localization of GUS activity in transgenic tobacco demonstrated rapidly induced and persistent activity localized to wound and pathogen (*Botrytis cinerea*) invasion sites. GUS activity was also observed in mature pollen (Warner *et al.*, 1993).

The expression pattern of the AoPR1-GUS gene correlated well, temporally and spatially, with the developmental expression observed for GUS fusions with promoters derived from genes coding for enzymes in the "core phenylpropanoid pathway." Enzymes involved in that pathway catalyze the synthesis of cell-wall structural molecules, pigments, UV-light protectants, chemical attractants, and antimicrobial phytoalexins. The AoPR1 promoter directed similar cell-specific transcription patterns in both asparagus and transgenic tobacco. The AoPR1 promoter contained sequence motifs similar to those implicated as important in the regulation of phenylpropanoid pathway genes (Warner *et al.*, 1994).

Sequence Information on Additional Betv1 Homologues

Additional partial sequences were available at the time of this study but were not included in the analysis. An N-terminal amino acid sequence of Cas s 1, the

major allergen from *Castanea sativa* (Fagaceae), was reported by Kos *et al.* (1993). Schoening *et al.* (1995) identified the N-terminus of a Bet v 1 homologue from *Prunus cerasus* (Rosaceae). A *betv1* homologue from *Taraxacum officinale* (Asteraceae) was discovered by J. D. Bewley and collaborators (personal communication) and M. Vanek-Krebitz from our group identified most of the coding sequence of a *betv1* homologue from *Brassica oleracea* (Brassicaceae). An N-terminal sequence of yet another Bet v 1 homologue from *Solanum melongena* (Solanaceae) was presented at the 52nd Annual Meeting of the American Academy of Allergy, Asthma, and Immunology (Hsieh *et al.*, 1996).

Phylogenetic Analyses

Parsimony analysis of the 67 nucleotide sequences was performed with PAUP 3.1 (Swofford, 1993) using a heuristic search. Indels were treated either as missing data or as new character states. The amount of support for monophyletic groups revealed in the maximally parsimonious tree(s) (MPTs) was examined with 100 bootstrap replicates (Felsenstein, 1985) using PAUP with the random addition and heuristic search options. A maximum likelihood (Felsenstein 1981) tree was generated using the DNAML program of PHYLIP (version 3.5, Felsenstein, 1993). The maximum likelihood analysis was performed with the input order of sequences randomized and the transition/transversion ratio set at 0.8620 based on the actually observed frequencies in the MPTs of the parsimony analysis. A neighbor-joining (Saitou and Nei, 1987) tree was constructed using the uncorrected distances (*p* distances) with the program MEGA (Kumar *et al.*, 1993). Of the 67 sequences sampled, 3 belong to monocots and 64 represent sequences from dicots. The three monocot sequences (AoPR1, AocPR2A, and AocPR3A) were selected as outgroups in the phylogenetic analyses.

Several analyses were performed treating indels in different ways. The first phylogenetic analysis treated gaps as missing data. The second analysis was conducted treating each gap position as a new character state as if it were a fifth base. The third analysis treated the gaps as missing data and then each indel was scored as a separate character (i.e., presence/absence data). Indels may provide phylogenetic information. However, the second approach weights multistate gaps by their length and was therefore not preferred in this study.

RESULTS

Sequence Characteristics

Betv1 homologues have an average nucleotide frequency of 0.2992 for A, 0.2148 for C, 0.2499 for G, and

0.2361 for T. The sequences of the open reading frames were conservative and thus the alignment was straightforward. Alignment of the 67 sequences from seven plant families representing both monocots and dicots required 15 indels, each of 3–12 bp (see Appendix). Among all the 67 sequences of both dicots and monocots compared, 438 of the 495 aligned positions were variable, 412 of which were phylogenetically informative. Of the 27 sequences from Betulaceae, 221 positions were variable with 146 phylogenetically informative. The 20 sequences of Fabaceae have 310 variable sites, 244 of which were phylogenetically informative.

Phylogenetic Analyses

Phylogenetic analyses employed all the available nucleotide sequences of *betv1* homologues, representing sequences of 22 species from seven plant families. Maximum parsimony analysis generated eight maximally parsimonious trees with a length of 2583 steps, a consistency index (CI) of 0.370, a retention index (RI) of 0.760, and a rescaled consistency index (RC) of 0.281. The strict consensus tree is presented in Fig. 1. The eight MPTs differed only in the relative positions of OTUs within the following monophyletic groups: (1) Bet v 1e and Bet v 1j; (2) Bet v 1b, Bet v 1c, and Bet v 1k; and (3) mdgc32, mdgc36, and mdgc38. Treating gaps as new states generated 16 MPTs with a total length of 2631 steps, a CI of 0.388, a RI of 0.769, and a RC of 0.298. The 16 trees were similar to the ones generated by treating gaps as missing data. They had an additional difference in the position of pI49 (*P. sativum*), which was either sister to PePR10 (*P. elatius*) or basal to the group of PePR10 (pI176, PfYpr10). Treating each indel as a separate character produced eight MPTs with the same phylogenetic conclusions as treating gaps as missing data (2602 steps, CI = 0.374, RI = 0.762, and RC = 0.285).

The neighbor-joining tree (not shown) was congruent with the MPTs (Fig. 1) and they differed only in the position of *cagc11* (*Corylus*, Betulaceae). The parsimony analysis suggested that it formed a monophyletic group with the seven sequences from Rosaceae (Mal d 1, Mal d 1.2, Ap15, mdgc10, mdgc32, mdgc36, mdgc38, and Pru a 1). The NJ tree suggested its basal position within the Betulaceae clade. The maximum likelihood tree had a similar topology to the MPTs, with *cagc11* basal to the Rosaceae clade.

Parsimony analysis of sequences from Betulaceae alone generated eight MPTs using Ap15 (*Malus*) as the outgroup. The eight MPTs have a total length of 529 steps, a CI of 0.688, a RI of 0.800, and a RC of 0.550. Their consensus has the same topology as that in Fig. 1. The parsimony analysis of the sequences from Faba-

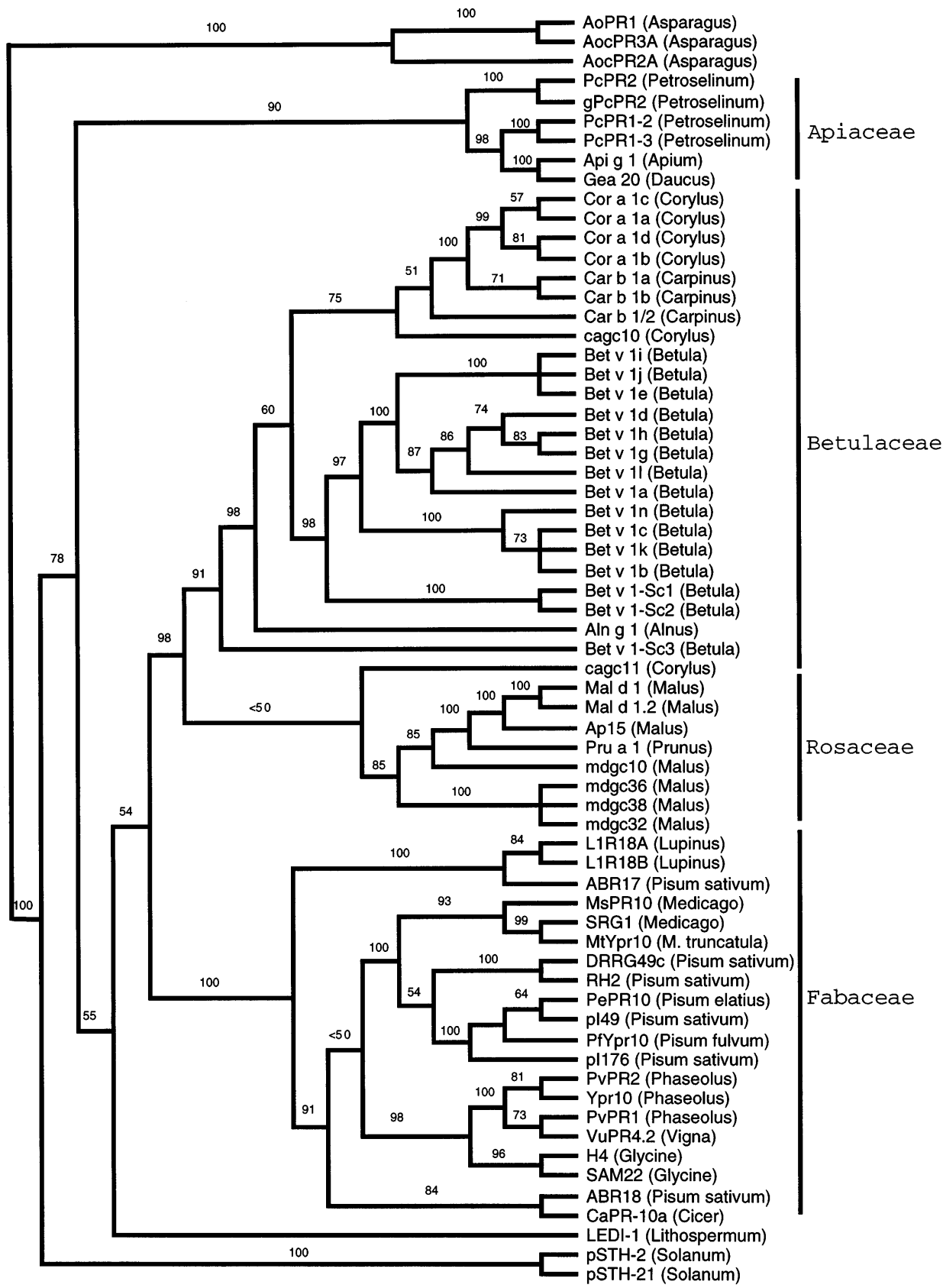


FIG. 1. The strict consensus tree of the eight most parsimonious trees of the *betv1* homologues (2583 steps, CI = 0.370, RI = 0.760, and RC = 0.281). Numbers above lines are the bootstrap values in 100 replicates.

ceae alone produced a single MPT using LEDI-1 (*Lithospermum*) as the outgroup. This Fabaceae tree has an identical topology as shown in Fig. 1, with a total length of 905 steps, a CI of 0.628, a RI of 0.650, and a RC of 0.408. Analyses using multiple outgroups (AP15 of *Malus*, Pru a 1 of *Prunus*, LEDI-1 of *Lithospermum*, and pSTH-2 of *Solanum*) produced the same phylogenetic results for Fabaceae as in Fig. 1.

Sequence Divergence

The pairwise p distance among sequences of the *betv1* homologues ranged from 0 to 0.583. The largest p distance was between AocPR3A (*Asparagus*) and Api g 1 (*Apium*). Within plant families, the ranges of the p distance were 0–0.403, 0–0.253, and 0.011–0.369, for Apiaceae, Betulaceae, and Fabaceae, respectively. Because of the large number of sequences, up to two representatives of each genus were presented in Table 3. The proportions of nonsynonymous differences (P_N) among *betv1* homologues with gap sites and missing data excluded ranged from 0 among some recent orthologues to 0.5430 between AocPR2A (*Asparagus*) and LEDI-1 (*Lithospermum*). The P_N values within Apiaceae, Betulaceae, and Fabaceae were 0–0.3360, 0–0.1630, and 0–0.2895, respectively. Nucleotide sequence differentiation was detected within species. For example, the P_N values among 15 sequences of *Betula verrucosa* ranged from 0 to 0.0729.

DISCUSSION

Concerted Evolution of the *Betv1*-like Genes

One of the most striking features of the *betv1* gene phylogeny (Fig. 1) was that the *betv1* homologues from each plant family constituted a monophyletic group with only the exception of *cagc11* from Betulaceae, which was grouped with homologues from Rosaceae in the maximum parsimony and maximum likelihood trees. We examined gene phylogenies for several nuclear multigene families including the actin genes (McElroy *et al.*, 1990), the cyclin genes (Kouchi *et al.*, 1995), the glutamine synthetase (GS) genes (Doyle, 1991), the phytochrome genes (Mathews *et al.*, 1995; Mathews and Sharrock, 1996), and the small heat-shock genes (Waters, 1995). Contrary to *betv1* homologues, all of the above gene families were shown to form two or more distinct classes in each plant family.

Let us assume that an ancestral species possessed its own *betv1* multigene family. Members of this gene family would have diverged to some degree through time. Derivatives of this species would be expected to demonstrate similar *betv1* gene lineages. Ancient lineages of *betv1* sequences within *betv1* gene families which predate the divergence of the species would be expected to be present. However, concerted evolution,

TABLE 3
Sequence Divergences of the *Betv1* Homologues

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 AocPR1 (<i>Asparagus</i>)	—	0.546	0.578	0.547	0.513	0.516	0.507	0.503	0.501	0.518	0.556	0.557	0.560	0.543	0.566	0.548	0.570	0.550	0.553	0.558
2 PcPR1-3 (<i>Pterocarpus</i>)	0.249	—	0.314	0.308	0.478	0.430	0.422	0.437	0.428	0.430	0.437	0.502	0.498	0.469	0.475	0.499	0.469	0.475	0.484	0.448
3 Api g 1 (<i>Apium</i>)	0.262	0.146	—	0.181	0.454	0.420	0.435	0.426	0.422	0.420	0.466	0.479	0.484	0.489	0.476	0.457	0.478	0.487	0.455	0.470
4 Gea 20 (<i>Daucus</i>)	0.248	0.143	0.084	—	0.447	0.424	0.424	0.429	0.416	0.424	0.438	0.510	0.486	0.481	0.470	0.476	0.489	0.487	0.483	0.485
5 pSTH-21 (<i>Solanum</i>)	0.234	0.215	0.203	0.200	—	0.412	0.419	0.410	0.408	0.429	0.437	0.463	0.455	0.410	0.426	0.440	0.440	0.458	0.431	0.481
6 Cor a 1c (<i>Corylus</i>)	0.243	0.200	0.194	0.196	0.193	—	0.085	0.159	0.155	0.141	0.327	0.423	0.426	0.382	0.405	0.405	0.400	0.411	0.380	0.413
7 Car b 1/2 (<i>Carpinus</i>)	0.239	0.196	0.201	0.196	0.192	0.171	—	0.153	0.151	0.139	0.333	0.420	0.411	0.399	0.409	0.407	0.411	0.411	0.386	0.429
8 Bet v 1a (<i>Betula</i>)	0.237	0.203	0.197	0.198	0.192	0.177	0.070	—	0.070	0.128	0.323	0.410	0.398	0.391	0.398	0.395	0.406	0.406	0.388	0.424
9 Bet v 1b (<i>Betula</i>)	0.236	0.199	0.195	0.192	0.191	0.175	0.073	0.034	—	0.112	0.304	0.401	0.402	0.378	0.377	0.386	0.383	0.396	0.378	0.431
10 Aln g 1 (<i>Alnus</i>)	0.244	0.200	0.194	0.196	0.201	0.168	0.067	0.062	0.054	—	0.338	0.408	0.406	0.393	0.396	0.397	0.400	0.398	0.376	0.422
11 Mal d 1 (<i>Malus</i>)	0.260	0.202	0.214	0.201	0.203	0.157	0.155	0.155	0.146	0.162	—	0.434	0.435	0.397	0.416	0.408	0.394	0.407	0.382	0.453
12 L1R18A (<i>Lupinus</i>)	0.256	0.232	0.220	0.234	0.211	0.199	0.198	0.193	0.189	0.192	0.203	—	0.170	0.314	0.324	0.331	0.355	0.348	0.347	0.443
13 L1R18B (<i>Lupinus</i>)	0.257	0.230	0.222	0.223	0.207	0.200	0.193	0.187	0.189	0.191	0.203	0.80	—	0.321	0.340	0.329	0.357	0.351	0.363	0.450
14 SRG1 (<i>Medicago</i>)	0.251	0.218	0.226	0.222	0.187	0.180	0.188	0.184	0.178	0.185	0.186	0.148	0.151	—	0.139	0.259	0.271	0.276	0.262	0.431
15 p1176 (<i>Pisum</i>)	0.265	0.221	0.220	0.217	0.197	0.193	0.195	0.190	0.180	0.189	0.197	0.153	0.160	0.66	—	0.268	0.278	0.286	0.256	0.452
16 CaPR-10a (<i>Cicer</i>)	0.255	0.232	0.211	0.220	0.202	0.192	0.193	0.187	0.183	0.188	0.192	0.156	0.155	0.123	0.128	0.113	0.241	0.254	0.229	0.462
17 PvPR2 (<i>Phaseolus</i>)	0.260	0.218	0.221	0.226	0.198	0.186	0.191	0.189	0.178	0.186	0.182	0.165	0.166	0.127	0.130	0.113	—	0.062	0.175	0.468
18 Ypr10 (<i>Phaseolus</i>)	0.251	0.221	0.225	0.206	0.191	0.191	0.191	0.189	0.184	0.185	0.188	0.162	0.163	0.129	0.134	0.119	0.29	—	0.188	0.475
19 SAM22 (<i>Glycine</i>)	0.257	0.225	0.210	0.223	0.198	0.180	0.183	0.184	0.179	0.178	0.180	0.164	0.171	0.124	0.122	0.109	82	88	—	0.473
20 LEDI-1 (<i>Lithosper.</i>)	0.253	0.207	0.216	0.223	0.215	0.191	0.198	0.196	0.199	0.195	0.208	0.202	0.205	0.198	0.209	0.212	215	218	217	—

Note. Uncorrected distances (p distances) are shown above the diagonal and observed nucleotide differences are given below the diagonal.

i.e., the tendency of a family of repeated genes to evolve in unison (Zimmer *et al.*, 1980), may occur and result in the pattern that the repeated genes within a species are more similar to each other than those of other species.

The multiple sequences of *betv1* homologues from the same species were generally more similar to each other than those from other species, suggesting that *betv1* homologous genes must have undergone concerted evolution. However, differentiation of genes into two or more groups within a species was also encountered, suggesting concerted evolution was incomplete. Seventeen sequences were available for *B. verrucosa*, sixteen of which (except Bet v 1-Sc3) formed a monophyletic group. Of the six sequences of *Corylus avellana*, the four cDNA sequences formed a monophyletic group and the two genomic sequences (cagc10 and cagc11) were scattered on the tree (Fig. 1). Phylogenetic analysis suggested three groups of *betv1* homologues for *Malus domestica*: Group I of the three cDNA sequences (Mal d 1, Mal d 1.2, and Ap15), Group II of mdgc10, and Group III of mdgc32, mdgc36, and mdgc38. The Group II and Group III sequences were genomic sequences. A more complicated situation was detected in *Pisum*. Eight *betv1* homologous sequences were available for three species of *Pisum*: *P. elatius*, *P. fulvum*, and *P. sativum*. Six of the eight *betv1* homologues formed a monophyletic group; however, ABR17 and ABR18 of *P. sativum* became highly divergent, perhaps with different functions. Therefore, among the *betv1*-like genes, a high level of similarity was usually observed of sequences within the species with some exceptions. However, the usually exclusive monophyly of sequences from each plant family suggested that concerted evolution operated during ancient times.

An alternative to the concerted evolution hypothesis is presented below. If gene copy number were to increase from one to several in the ancestor of a taxon, and the copies then underwent divergence, a pattern of gene evolution similar to the one shown in Fig. 1 could theoretically result, with all of the copies within that taxon forming a somewhat divergent yet monophyletic group (here referred to as the divergence hypothesis). In the majority of the species examined so far, the gene copy number is usually more than one except in *Cicer* (see also Materials and Methods). If the divergence hypothesis were operating, we would expect to see several major paralogues in the gene phylogeny unless the gene divergence was very recent (no speciation events among the terminal taxa in Fig. 1). The more or less strict monophyly of *betv1* homologues from each plant family (only with the exception of cagc11 in the maximum likelihood and the maximum parsimony analyses) supports the concerted evolution hypothesis.

Phylogeny of Fabaceae Inferred from the *Betv1* Gene Tree

The phylogeny of Fabaceae has been examined extensively based on both morphological and molecular data (Polhill and Raven, 1981; Doyle *et al.*, 1992; Crisp and Doyle, 1995). Nucleotide sequences of *betv1* homologues were available for seven genera of Fabaceae: *Cicer* (tribe Cicereae), *Glycine* (tribe Phaseoleae), *Lupinus* (tribe Genisteae), *Medicago* (tribe Trifolieae), *Phaseolus* (tribe Phaseoleae), *Pisum* (tribe Vicieae), and *Vigna* (tribe Phaseoleae).

The *betv1* homologues' nucleotide data suggested a phylogeny for the seven legume genera represented. First, a close relationship among *Glycine*, *Phaseolus*, and *Vigna* was suggested. This was consistent with results from other nuclear genes (Doyle, 1991) and the *rbcL* data (Doyle, 1995). These three genera are members of the tribe Phaseoleae (Polhill and Raven, 1981). Second, *Medicago* (Trifolieae) and *Pisum* (Vicieae) were suggested to be closely related. Third, *Glycine*, *Phaseolus*, *Medicago*, and *Pisum* formed a monophyletic group. Fourth, *Cicer* was suggested to be basal to the ((*Glycine*, *Phaseolus*) (*Medicago*, *Pisum*)) group. Fifth, *Lupinus* was cladistically basal among the seven genera sampled. This phylogeny was in complete agreement with the general phylogeny reconstructed from multiple nuclear genes (Doyle, 1991) except that the Doyle phylogeny did not include *Cicer*.

Despite the presence of concerted evolution, the *betv1*-like genes seemed to be useful for inferring phylogenies at the intrafamilial level in Fabaceae. In this regard, the *betv1* genes are similar to the *rbcS* genes encoding the small subunit of ribulose-1,5-bisphosphatase carboxylase (Meagher *et al.*, 1989). The complicated branching pattern of *Pisum betv1* homologues (Fig. 1), however, suggested that we need to be cautious in interpreting our results when the sampling of the genes is inadequate.

Betv1 Phylogeny of Betulaceae

Betulaceae, the birch family, consists of six genera: *Alnus*, *Betula*, *Carpinus*, *Corylus*, *Ostrya*, and *Ostryopsis*. *Corylus* was sometimes treated as belonging to its own family Corylaceae. Sequences of *betv1* homologues were available for four of the six genera: *Alnus*, *Betula*, *Carpinus*, and *Corylus*. The data suggested a close relationship between *Carpinus* and *Corylus*. *Betula* was found to be sister to the *Carpinus*–*Corylus* group. The phylogenetic position of *Alnus* was not well resolved. The parsimony and maximum likelihood trees suggested a basal position of *Alnus*. The neighbor-joining tree showed that it was basal to the *Carpinus*–*Corylus* group with the Kimura two-parameter distance. The NJ tree with the uncorrected *p* distance suggested that *Alnus* formed a sister group with *Betula*.

Data from morphology, ITS sequences, and *rbcL* data (Crane, 1989; Bousquet *et al.*, 1992) suggested that *Alnus* formed a sister group with *Betula*, and *Carpinus* and *Corylus* are sister to each other. The *betv1* data therefore were largely congruent with previous studies.

Potential Phylogenetic Utility of *Betv1* Homologues

Criteria for a potential phylogenetic marker should consider: (1) size of the gene or DNA region, (2) demonstration of organismal relationships, (3) ease of sequencing, (4) level of sequence divergence, and (5) ubiquitous occurrence.

Size. The open reading frames of *betv1* homologues ranged between 465 to 480 bp. Compared with sequences commonly used for phylogenetic purposes, *betv1* homologues are similar in size to the ITS regions (ITS-1 and ITS-2, each <300 bp in flowering plants, Baldwin *et al.*, 1995), but shorter than the chloroplast *matK* (a maturase-encoding gene, approximately 1500 bp, Johnson and Soltis, 1994), *ndhF* (encoding a component of a putative chloroplast respiratory chain, 2223 bp in tobacco, Olmstead and Sweere, 1994), and *rbcL* (encoding the large subunit of ribulose-1,5-bisphosphate carboxylase, 1431 bp in maize, Clegg, 1993), and the nuclear ribosomal 18S (1807 bp in soybean, Hamby and Zimmer, 1992) and 26S (3376 bp in rice, Takaiwa *et al.*, 1985). However, *betv1* homologues show a higher level of sequence divergence (Table 3) and thus generate enough differences to make phylogenetic hypotheses at the intrafamilial level as we have shown in Betulaceae and Fabaceae (Fig. 1). We are aware of the fact that the long and more conserved sequences usually have less homoplasy compared with the short and highly divergent sequences. We, however, value the phylogenetic information in both types of data.

Demonstration of species relationships. The ultimate goal of molecular systematists is to construct species relationships from gene phylogenies. However, gene trees may not reflect accurately species phylogenies due to biological processes such as introgression, lineage sorting, and gene duplication (Doyle, 1992). Gene duplication, divergence after duplication, speciation, divergence following speciation, and concerted evolution may yield gene trees with complex orthologous and paralogous homologies.

Simulation studies by Sanderson and Doyle (1992) suggested that correct species phylogenies can be inferred from genes which have undergone either very strong or little concerted evolution. As we have demonstrated, *betv1* homologues have undergone relatively strong concerted evolution. Empirical phylogenetic results of Betulaceae and Fabaceae inferred from *betv1* homologues were generally congruent with those inferred from morphology and other molecules.

Ease of sequencing. Most available sequences of *betv1* homologues were from cDNA. However, genomic DNA was used to amplify *betv1* and like genes by PCR in hazel (*Corylus avellana*) and apple (*Malus domestica*) (Hoffmann-Sommergruber *et al.*, 1997). If relatively conservative consensus primers can be designed, the small size of the *betv1* sequence and the relatively large divergence among taxa should be an advantage for systematists to effectively use it phylogenetically. In Table 4, we suggested family-specific primer pairs for the Fabaceae, the Fagales, and the Rosaceae. In addition, the single intron—in all cases we looked into—is small and its location is conserved.

Level of sequence divergence. The proportions of nonsynonymous differences (P_N) within Apiaceae, Betulaceae, and Fabaceae were 0–0.3360, 0–0.1630, and 0–0.2895, respectively, suggesting that *betv1* homologues can generate enough characters for phylogenetic comparisons at the intrafamilial level. Its utility at the very deep level (e.g., between monocots and dicots) is not recommended because a very high level of divergence between the monocot *Asparagus* and other dicots was found, ranging from 0.42 to 0.5430. The high level of divergence among monocots and dicots may prevent accurate sequence alignment across taxa and cause

TABLE 4

Suggested Amplification Primers for *Betv1* Homologues

Taxon	Primer sequence	Primer designation	Product region	Product size (bp) ^a
Fabaceae	5' primer CCT GCT ACA CTT TAC AAA GCT	fab1	46 to 460	415
	3' primer AAC GTA ACC CTC AAT GGC CTT G	fab2		
Apiaceae	5' primer GAG GTT ACT TCT TCA GTC TCA GC	api1	25 to 455	431
	3' primer AGA TAG GCC TCAATA GCC TTG AA	api2		
Fagales	5' primer ATG GGT GTT TTC AAT TAC GAR ^b	fag1	1 to 466	466
	3' primer GTG TGC CAA GAG GTA GCT CTC	fag2		
Rosaceae	5' primer: same as for Fagales	ros1	1 to 479	479
	3' primer: TTG TAG GCA TCG GAG TGT GCC	ros2		

^a PCR product length without intron.

^b R means A and G in equal amounts at this position.

high homoplasy in phylogenetic analyses. *Betv1*-like genes will be most useful at the intrafamilial level or among closely related plant families.

Ubiquity of *betv1* homologues. Members of the *betv1* multigene family, originally identified from *B. verrucosa*, have been found in 10 plant families: Apiaceae, Asteraceae, Betulaceae, Boraginaceae, Brassicaceae, Fabaceae, Fagaceae, Liliaceae, Rosaceae, and Solanaceae, representing Asteridae s.l., Rosidae s.l., and monocots (Chase *et al.*, 1993; Soltis *et al.*, 1997). Complete sequence data of *betv1* homologues were available for 7 of the 10 families except for Asteraceae, Brassicaceae, and Fagaceae. A *betv1* homologue in *Taraxacum officinale* (Asteraceae) was identified (J. D. Bewley, personal communication); the protein reacted with a polyclonal rabbit antiserum raised against Bet v 1, and the protein was isolated and N-terminally sequenced. The 30 amino acids obtained so far showed a high similarity with that of Api g 1 from Apiaceae. Most of the coding sequence of a *betv1* homologue from *Brassica oleracea* (Brassicaceae) has been obtained (M. Vanek-Krebitz, unpublished). An N-terminal amino acid sequence of Cas s 1, the major allergen from *C. sativa* (Fagaceae) was reported by Kos *et al.* (1993).

Potential Limitations

The phylogenetic utility of this multigene family at higher taxonomic level may not be appropriate due to the relatively high sequence divergence among distantly related plant families [e.g., 0.5430 between AocPR2A (*Asparagus*) and LEDI-1 (*Lithospermum*), Table 3], which may prevent accurate sequence alignment and cause a high level of homoplasy. We therefore need to be cautious in interpreting deep level phylogenetic relationships. One of the surprising results is the relationships among Betulaceae, Fabaceae, and Rosaceae. Traditional morphology (Cronquist, 1988) and the *rbcL* data (Chase *et al.*, 1993) suggested a closer relationship between Fabaceae and Rosaceae. The

nuclear 18S rDNA sequences showed a closer relationship between Betulaceae and Fabaceae (Soltis *et al.*, 1997). The *betv1* phylogeny supported a closer relationship between Betulaceae and Rosaceae (Fig. 1).

In addition the *betv1* sequences available were mostly from relatively advanced plant families of Asteridae s.l., Rosidae s.l., and monocots (Chase *et al.*, 1993; Soltis *et al.*, 1997). The root of the dicots (Fig. 1) identified by the outgroup monocot sequences from *Asparagus* falls in the middle of the sequences representing the highly derived Asteridae sequences (Apiaceae, Boraginaceae, and Solanaceae). This may be due to the limited sampling or the lack of resolution of the *betv1* gene family in constructing deep level phylogenies.

Furthermore the fact that nucleotide sequence differentiation was detected within species and occasional polyphyly was found for certain species (e.g., *P. sativum*) suggested that adequate sampling of clones from PCR products is necessary and we need to be cautious in interpreting "surprising" phylogenetic results.

In summary, phylogenies of Betulaceae and Fabaceae inferred from *betv1* gene trees were generally congruent with those based on morphology and other molecules. *Betv1* homologues have potential phylogenetic utilities at the intrafamilial level or among closely related families in flowering plants.

APPENDIX

Alignment of the nucleotide sequences of the *betv1* genes. All the available sequences of Betulaceae and one sequence of each of the other families were chosen to show the variation at different taxonomic levels. Dots indicate identity with the Cor a 1c (*Corylus*) sequence in the first row. Dashes indicate gaps at insertion/deletion sites and "N" represents missing data.

APPENDIX

Cor a 1c (<i>Corylus</i>)	ATGGGTGTTT	TCAATTACGA	GGCTGAGACC	ACCTCCGTTA	TCCCTGCGGC	AAGGCTGTTC	AAGTCCTATG	TCCTTGATGG	CGATAAGCTC	ATCCCAAAGG	100
Cor a 1a (<i>Corylus</i>)			..T.....T	C.....							100
Cor a 1d (<i>Corylus</i>)			..T.....T	C.....	..T.A.....						100
Cor a 1b (<i>Corylus</i>)			..T.....T	C.....	..A.....						100
Car b 1a (<i>Carpinus</i>)		..T.....	A..C..A..	C.....	..A.....	T.....					100
Car b 1b (<i>Carpinus</i>)			..A..A..	C.....	..A.....	T.....			..G.....		100
Car b 1/2 (<i>Carpinus</i>)		..T..C..T..	A..C..A..		..A..A..	T.....	..G...T.A		..A.....		100
Bet v 1f (<i>Betula</i>)		..ATA...G..	..T.....	..A.....	T.....A..	..G...TCA		..A...T..	T.T.....		100
Bet v 1i (<i>Betula</i>)		..ATA...G..	..T.....	..A.....	T.....A..	..G...TCA		..A...T..	T.T.....		100
Bet v 1j (<i>Betula</i>)		..A..A...G..	..T.....	..A.....	T.....A..	..G...TCA		..A...T..	T.T.....		100
Bet v 1e (<i>Betula</i>)		..A..A...G..	..T.....	..A.....	T.....A..	..G...TCA		..A...T..	T.T.....		100
Bet v 1d (<i>Betula</i>)		..AAT.....	..T.....	..A.....	TC.A.....	..G...T.A		..T.....	G.T.....		100
Bet v 1h (<i>Betula</i>)		..AAT.....	..T.....	..A.....	TC.A.....	..G...T.A	..T.....	..T.....	G.T.....		100
Bet v 1g (<i>Betula</i>)		..AG.....	..T.....	..A..A..	T.....	..G...T.A	..C..G..	..T.....	T.....		100
Bet v 1a (<i>Betula</i>)		..AA.....	..T.....	..A..A..	TC.A.....	..G...T.A		..T.....	T.T.....		100
Bet v 1-Sc1 (<i>Betula</i>)		..G.....	..G.....	..T.....	..A.....	..G...T.A		..T.....	T.....		100
Bet v 1-Sc2 (<i>Betula</i>)		..A.....	..G.....	..G.....	..C.....	..G...T.A		..T.....	T.....		100
Bet v 1m (<i>Betula</i>)		..AG.....	..T.....	..A..A..	T.....	..G...T.A	..C.....	..T.....	T.....		100
Bet v 1n (<i>Betula</i>)		..AG.....	..T.....	..A..A..	T.....	..G...T.A	..C.....	..T.....	T.....		100
Bet v 1c (<i>Betula</i>)		..AG.....	..T.....	..A..A..	T.....	..G...T.A	..C..G..	..CT...	T.....		100
Bet v 1k (<i>Betula</i>)		..AG.....	..T.....	..A..A..	T.....	..G...T.A	..C..G..	..CT...	T.....		100
Bet v 1l (<i>Betula</i>)		..AA...G...	..T.....	..A..A..	TC.AA...	..G...TCA		..A.....	G.T.....		100
Bet v 1b (<i>Betula</i>)		..AA...G...	..T.....	..A..A..	TC.A.....	..G...T.A	..C..G..	..T...CT...	T.....		100
Aln g 1 (<i>Alnus</i>)		..A..G..A..	C.....	..A.....	TC.....	..G...T.A			C.T.....		100
cagc10 (<i>Corylus</i>)		..AA...G...		..T..A..A..	T.....T	..G...T.A		..A...C...			100
Bet v 1-Sc3 (<i>Betula</i>)		..C.....	..A...G...		..TG...C...	..T...C...	..T...C...	..C...C...	..T...C...		100
cagc11 (<i>Corylus</i>)		..AA.....	..T..A...	..C.....	T.....	..AGG.T...	..A...TC	..C...C...			100
Mal d 1 (<i>Malus</i>)	..C..A..CA..TT..	..AAAC...TT..	..T..AG..	..T..AC..AT..	..AT.....	..G...T...		..C...C...	..C...A		100
LlR18A (<i>Lupinus</i>)	..A.....	..TGC...T...	..AAA...CAA	T...TAC..G	..GG...C..A..	T..AA...A..A..	..AG..TCTCA	CAAAA...TC	..T...G..AA	..G.....	100
LEDI-1 (<i>Lithospermum</i>)	..C...G..TG...	..T..A..T..TT..	..TAG..ACAG	..A...TC..C..A...	..A...	..G...GGT	..A..C..CT..	..T...CC..T	..C...T...	..A	100
Gea 20 (<i>Daucus</i>)	..CCC..AG..GCC..TTC	..ACTC...T...	..T...TTCAG	..T..C...A..A	..G..AAA..A..	..NCNG..AT...	..T...CAG..T	..T...C...			100
pSTH-2 (<i>Solanum</i>)	..CA..CT..GC..TAC	..ACA.....	..AA..ACCA..	..TG..CC..TA..	..T...T...	..AG..T..TG..	..TG...TC	..T...C...T..T	..T...T...T		100
AsPRI (<i>Asparagus</i>)	..A...TCAG	..GG..GC..GGAG	..CCAC...GT..	..G..TGT..AA..G	..G..C...A..G	..C...A...	..G..GGCAA	..G...C..CT..	..GC..C...C...	..GG...T...A	100

Cor a 1c (<i>Corylus</i>)	TTGCACCTCA	AGCTATTACC	AGCGTTGAAA	ACGTTGAAGG	AAATGGAGGG	CCTGGAACCA	TCAAGAATAT	CACCTTTGGC	GAA---GGCA	GCCGTTACAA	200
Cor a 1a (<i>Corylus</i>)									-----		200
Cor a 1d (<i>Corylus</i>)				..G...					-----		200
Cor a 1b (<i>Corylus</i>)									-----		200
Car b 1a (<i>Carpinus</i>)	..T.....	..G.....	..G.....					..C.....	..T...C...T...		200
Car b 1b (<i>Carpinus</i>)		..G.....	..G.....					..C.....	..C...T...		200
Car b 1/2 (<i>Carpinus</i>)	..T...C...	..G...G...	..T.....			..A.....	..TC.....	..C...GT...			200
Bet v 1f (<i>Betula</i>)	..G...G...	..C...G...	..T.....	..A.....		..G...G...	..G...CC...	..T...C...T...			200
Bet v 1i (<i>Betula</i>)	..G...G...	..C...G...	..T.....	..A.....		..G...G...	..G...CC...	..T...C...T...			200
Bet v 1j (<i>Betula</i>)	..G...G...	..C...G...	..T.....	..A.....		..G...G...	..G...CC...	..T...C...T...			200
Bet v 1e (<i>Betula</i>)	..G...G...	..C...G...	..T.....	..A.....		..G...G...	..G...CC...	..T...C...T...			200
Bet v 1d (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..T...C...T...			200
Bet v 1h (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..T...C...T...			200
Bet v 1g (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..T...C...T...			200
Bet v 1a (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..T...G...	..G...CCT...	..T...C...T...			200
Bet v 1-Sc1 (<i>Betula</i>)	..C...A..CG...	..G...T...T...	..A.....			..G...G...	..CC.....	..C...T...			200
Bet v 1-Sc2 (<i>Betula</i>)	..C...C...	..CG...G...	..T...T...	..A.....		..G...G...	..CC.....	..C...T...			200
Bet v 1m (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..C...T...			200
Bet v 1n (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..C...T...			200
Bet v 1c (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..C...T...			200
Bet v 1k (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..C...T...			200
Bet v 1l (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..T...CG..TT...			200
Bet v 1b (<i>Betula</i>)	..C...C...	..C...G...	..T.....	..A.....		..G...G...	..A...CCT...	..C...T...			200
Aln g 1 (<i>Alnus</i>)	..G...G...	..G...G...	..T...G...	..A.....		..G...G...	..CC.....	..C...TT...			200
cagc10 (<i>Corylus</i>)	..G...G...	..G...G...	..T.....		..A...A...	..G...G...	..CC.....	..T...C...T...			200
Bet v 1-Sc3 (<i>Betula</i>)	..G...G...	..GAACG...G...	..T...CC...	..TA.....	..C...A...	..G...G...	..CC.....	..A...A...T...			200
cagc11 (<i>Corylus</i>)	..T...AA..G...	..C...C..AG	..A.....	..T..A..C...	..T...C...	..T...G...	..TGT...C..AT	..G...CA..T...			200
Mal d 1 (<i>Malus</i>)	..C...G...A..C..AG	..CAA..C...	..T...C...	..C...T..C	..C...C...	..G...G...	..T...T...	..AG...GG			200
LlR18A (<i>Lupinus</i>)	..GAT---G...	..GC..A..CCAA	..T.....	..TT.....	..A...A...	..T...G...	..A..TTGC..ATT	..C...T...ATG	..T..ACACT..G		200
LEDI-1 (<i>Lithospermum</i>)	..C...G...CA..G..C..AA	..AGA..C...--	..T...C...	..C...T..C	..T...T...	..TGC...	..TCAT...A	..T...CGG	..T...CAGTG...		200
Gea 20 (<i>Daucus</i>)	..C...C..CGG	..TAC..AG	..T...C..TG	---A...	..G..C...T..A	..G...G...	..GA..T...	..C...CC...	..T...CAAT...C		200
pSTH-2 (<i>Solanum</i>)	..GATG...A...	-----	..A...TA...	..G..C	..TG..G...--	..A...G...	..A...G...	..G...TT	..TT..CA..CAAT...		200
AsPRI (<i>Asparagus</i>)	..TG...G...	..CTT...G...	..G..T..GCTC	..G...GTCT...	..G...TCT	..GTA...	..CGAG..G...	..AGA..CAA	..ATCCT..CT..TA..C...T...G		200

Cor a 1c (<i>Corylus</i>)	GTACGTGAAG	GAGAGGGTTG	ATGAGGTTGA	CAACACAAAC	TTCACATACA	GCTACACCGT	GATCGAGGGT	GATG-----	TCCTGGGTGA	CAAGCTGGAG	300
Cor a 1a (<i>Corylus</i>)								-----			300
Cor a 1d (<i>Corylus</i>)					A..T..			-----			300
Cor a 1b (<i>Corylus</i>)					A..T..			-----			300
Car b 1a (<i>Carpinus</i>)	.T.....			.G.....	A.....	A.....		-----	.C.....		300
Car b 1b (<i>Carpinus</i>)	.T.....			.G.....	A.....	A.....		-----	.C.....		300
Car b 1/2 (<i>Carpinus</i>)			A..A...	.C.....	A.....	A.....		-----	.C.....		300
Bet v 1f (<i>Betula</i>)		.C.....	.G.....	.C.....	A.....	T..G...	.C..G.C----	C.G...G...	.CAT...		300
Bet v 1i (<i>Betula</i>)		.C.....	.G.....	.C.....	A.....	T..G...	.C..G.C----	C.G...G...	.CAT...		300
Bet v 1j (<i>Betula</i>)		.C.....	.G.....	.C.....	A.....	T..G...	.C..G.C----	C.G...G...	.CAT...		300
Bet v 1e (<i>Betula</i>)		.GC.....	.G.....	.C.....	A.....	T..G...	.C..G.C----	C.G...G...	.CAT...		300
Bet v 1d (<i>Betula</i>)		.C..A...	.G.....	.C.....	A.....	AT..G...	.T...C..G.C----	C.G...C...	.CAT...		300
Bet v 1h (<i>Betula</i>)		.C..A...	.G.....	.C.....	A.....	AT..G...	.T...C..G.C----	C.G...C...	.CAT...		300
Bet v 1g (<i>Betula</i>)		.C..A...	.G.....	.C.....	A.....	AT..G...	.T...C..G.C----	C.G...C...	.CAT...		300
Bet v 1a (<i>Betula</i>)		.C..A...	.G.....	.C.....	A.....	AT..G...	.T...C..G.C----	C..A..C...	.CAT...		300
Bet v 1-Sc1 (<i>Betula</i>)		.T.....	.G.....	.C.TGT..	A.....	T..G...		CAG...C...	.CA...		300
Bet v 1-Sc2 (<i>Betula</i>)		.T.....	.G.....	.CGTGT..	A.....	T..G...		CAG...C...	.CA...		300
Bet v 1m (<i>Betula</i>)			.G.....	.C..G...	A.....	G..A...		CGT...C...	.CAT...		300
Bet v 1n (<i>Betula</i>)			.G.....	.C..G...	A.....	G..A...		CGT...C...	.CAT...		300
Bet v 1c (<i>Betula</i>)			.G.....	.C..G...	A.....	G..A...		CGT...C...	.CAT...		300
Bet v 1k (<i>Betula</i>)			.G.....	.C..G...	A.....	G..A...		CGT...C...	.CAT...		300
Bet v 1l (<i>Betula</i>)		.C..A...	.G.....	.C.....	A.....	AT..G...	.C...G.C----	C.G...C...	.CAT...		300
Bet v 1b (<i>Betula</i>)		A.....	.G.....	.C..G...	A.....	G..A...		CGT...C...	.CAT...		300
Aln g 1 (<i>Alnus</i>)		A.....		TOG.GT..	A.....	T..G...		C.G...C...	.GCA...		300
cagc10 (<i>Corylus</i>)			A.....		A.....		C.....	C.G.C----	C.G...A...	.AG...	300
Bet v 1-Sc3 (<i>Betula</i>)		A.....	C.C.....	.A.C...	.C..G...	A.....	T...G..A..A.....	C.G.C----	CAT...C...	.CA...	300
cagc11 (<i>Corylus</i>)	C...A.A...	C.A.A...	A..A...	.C.AG...	TT...TC		GT...T..A..C		CTT...TCC..	A.....	300
Mal d 1 (<i>Malus</i>)	C.....	C.C..A...	CTC.A...	.G.AG...G.	A.T...T	C.....	TT...T..A..A	-----	CTT...ACA...	.CCA.C...	300
LIR18A (<i>Lupinus</i>)	C.TT...CT.	C.C.AAC...	.CAA.A...	TG.AG.T...	.G.....	A.....	G..A..A..T.GA...	.A.....	GGT...A..A.	A.GTT.A...	300
LEDI-1 (<i>Lithospermum</i>)	.TG..T...	TTC.A.A...	.CAT.G...	.G.A.GT..	C.T...TG	CTG...T..	A...A...A...	.GC.AGTTAT	CAA.A.CG...	...AA.CTT.	300
Gea 20 (<i>Daucus</i>)	C.CAA...C	.TT...AC...	.CA..GA...	.GGAGGC...	.G.....	G..AT.C..A..	C..T..T..A	.CA-----	...TCTA...	ATTC.A.C..A	300
pSTH-2 (<i>Solanum</i>)	...T.....	C.C.A.A..C	.TT.....	TG...AG..T	.GGTGAC...	AA..TT.AA.	...T..A..A	-----	.T..T..A...	...A..T..A	300
AsPRI1 (<i>Asparagus</i>)	C..T.....	.AC.CC.G.	.TTC..A...	.C.TGAC..G	.GAGGTG...	AGC.G...C	CG.G..A...	.GA-----	GTT.A...A.	G.T.T.T..A	300

Cor a 1c (<i>Corylus</i>)	AAGGTCTGCC	ACGAGCTGAA	GATAGTGGCA	GCCCCGTGGT	GAGGATCCAT	CTTGAAGATC	AGCAGCAAGT	TCCACGCCAA	AGGTGACCAT	G---AGAITA	400
Cor a 1a (<i>Corylus</i>)									.C.....	-----	400
Cor a 1d (<i>Corylus</i>)	.A.....	A.G.....			.G.....	C.....				-----	400
Cor a 1b (<i>Corylus</i>)		A.G.....							.C.....	-----	400
Car b 1a (<i>Carpinus</i>)		A.....				.G.....			.CT.....	-----	400
Car b 1b (<i>Carpinus</i>)		A.....			T.....	.G.....			.CT.....	-----	400
Car b 1/2 (<i>Carpinus</i>)		A.....			T.....	.G.....			.CT.....	-----	400
Bet v 1f (<i>Betula</i>)	.A...CTA	.A.A...		A...AA..	.C.....		A..A..A..TA...	.A.....	---G.G.	400	
Bet v 1i (<i>Betula</i>)	.A...CTA	.A.A...		A...AA..	.C.....		A..A..A..TA...	.A.....	---G.G.	400	
Bet v 1j (<i>Betula</i>)	.A...CTA	.A.A...		A...AA..	.C.....		A..A..A..TA...	.A.....	---G.G.	400	
Bet v 1e (<i>Betula</i>)	.A...CTA	.A.A...		A...AA..	.C.....		A..A..A..TA...	.A.....	---G.G.	400	
Bet v 1d (<i>Betula</i>)	.A...C.A	.A.A...		A.....	.G.G...		A..A..A..A...	.CA.....	---G.G.	400	
Bet v 1h (<i>Betula</i>)	.A.AA..A	.A.A...		A.....	.G.G...		A..A..A..A...	.CA.....	---G.G.	400	
Bet v 1g (<i>Betula</i>)	.A.AA..A	.A.A...		A.....	.G.G...		A..A..A..A...	.CA.....	---G.G.	400	
Bet v 1a (<i>Betula</i>)	.A...C.A	.A.A...		A.....			A..A..A..A...		---G.G.	400	
Bet v 1-Sc1 (<i>Betula</i>)	.A...A	.A.A...	.C...				A..A..A..A...	.CA.....	---G.	400	
Bet v 1-Sc2 (<i>Betula</i>)	.A...A	.A.A...	.C...				A..A..A..A...	.CA.....	---G.	400	
Bet v 1m (<i>Betula</i>)	.A.T...A	.A.A...		A..A...			A..A..A..A...	.C.....	---G.	400	
Bet v 1n (<i>Betula</i>)	.A.T...A	.A.A...		A..A...			A..A..A..A...	.C.....	---G.	400	
Bet v 1c (<i>Betula</i>)	.A.T...A	.A.A...		A..A...			A..A..A..A...	.C...A	---G.	400	
Bet v 1k (<i>Betula</i>)	.A.T...A	.A.A...		A..A...			A..A..A..A...	.C.....	---G.	400	
Bet v 1l (<i>Betula</i>)	.A...C.A	.A.A...		A.....	.G.G...		A..A..A..A...	.CA.....	---G.G.	400	
Bet v 1b (<i>Betula</i>)	.A.T...A	.A.A...		A.....			A..A..A..A...	.C.....	---G.	400	
Aln g 1 (<i>Alnus</i>)	...TA	.A.C...		A.....			A..A..A..A...	.C.....	---A.	400	
cagc10 (<i>Corylus</i>)	.A...TA	.A.A...	.T...	A.....			TG.G.A...	A..A...	---G.G	400	
Bet v 1-Sc3 (<i>Betula</i>)	.A...CTT	.A.C...	.T...C				.C.....	A..A...	.C...AT	T---C.C.C.	400
cagc11 (<i>Corylus</i>)	.AA..AATT	.A.C...		T...CA...		T...	.G...	A..A..A..T	.A.....	---AC.C	400
Mal d 1 (<i>Malus</i>)	.AA.A.CTT	.ACC...	.T.G...	TGT---AA	.T.TG...C	A.C...G.	.T...TC.T.A...	.C.....	G..AA..AT	---A.C	400
LIR18A (<i>Lupinus</i>)	.AA...CAT	.T...AGT...	.TT..C.T	.G...A...	.T.....	TGGA.A..T	.ATGTT.A..T..TA...		---GT.C.AT	400	
LEDI-1 (<i>Lithospermum</i>)	.A..GACA	.T..AG...	A..T.A.T.	AG...AA		.G..A..C.T.CA.GT	T.TGTTG..A..T..CT...	G.---AGG.	ACAACTC..A	400	
Gea 20 (<i>Daucus</i>)	TOCA.TGAA	C.C.TA..GT	AG.T...C	A.TG...AC	...TAG...	TACC...C	.CTGC..TA	...A...	.C.---	.CCGT.G.C	400
pSTH-2 (<i>Solanum</i>)	TCCA.T.C.T	.T..T..C.	AT.T.AA..T	CATGAGAA..	...GTG.	T.GC...TC.T	.TA.CTG..A...	A..A.A...	...TT..	---T.T.G	400
AsPRI1 (<i>Asparagus</i>)	TG.T.C.AC.A	CTC.CT.C.	.T.C.A.C.C	T.GAGCAAC..	.T...G.C.	.G.C...G.G	.CTGCATCC..A..GATTCCT	CCCA.GTGTIC	.CCG.TGAG	400	

Cor a 1c (<i>Corylus</i>)	ATGCAGAGGA GATCAAGGGT GCCAAAGAAA TGSCCGAGAA ACTTTTAAGG GCGGTTGAGA CCTACCTATT GGCACACTCT GCTGAATACA ACTAA	495
Cor a 1a (<i>Corylus</i>)	G.....A.....	495
Cor a 1d (<i>Corylus</i>)	G.....	495
Cor a 1b (<i>Corylus</i>)	G.....	495
Car b 1a (<i>Carpinus</i>)	A.....G.....C.....A.....	495
Car b 1b (<i>Carpinus</i>)	.C.....G.....G.....C.....A.....	495
Car b 1/2 (<i>Carpinus</i>)	.C.....G.....G.....C.....A.....	495
Bet v 1f (<i>Betula</i>)	.G.....C.....T.....CA AGT.....GA.....C.....G.....C.....G.....C.....C.....A.....CC.....	495
Bet v 1i (<i>Betula</i>)	.G.....C.....T.....CA AGT.....GA.....C.....G.....C.....G.....C.....C.....A.....CC.....	495
Bet v 1j (<i>Betula</i>)	.G.....C.....T.....CA AGT.....GA.....C.....G.....C.....G.....C.....C.....A.....CC.....	495
Bet v 1e (<i>Betula</i>)	.G.....C.....T.....CA AGT.....GA.....C.....G.....C.....G.....C.....C.....A.....CC.....	495
Bet v 1d (<i>Betula</i>)	.G.....C.....G.T.....CA AGT.....G.....C.....G.....C.....G.....T.....C.....A.....CC.....	495
Bet v 1h (<i>Betula</i>)	.G.....C.....G.T.....CA AGT.....G.....C.....G.....C.....G.....T.....C.....A.....CC.....	495
Bet v 1g (<i>Betula</i>)	.G.....C.....G.T.....CA AGT.....G.....C.....G.....C.....G.....T.....C.....AC.CC.....	495
Bet v 1a (<i>Betula</i>)	.G.....C.....G.T.....CA AGT.....G.....C.....G.....C.....G.....C.....C.....A.....CC.....	495
Bet v 1-Sc1 (<i>Betula</i>)	.G.....C.....T.....C. AG.....A.....T.....GC.....C.....C.....G.....C.....A.....CC.T.....	495
Bet v 1-Sc2 (<i>Betula</i>)	.G.....C.....T.....C. AG.....A.....T.....GC.....C.....C.....G.....C.....A.....CC.T.....	495
Bet v 1m (<i>Betula</i>)	.G.....C.....T.....G.....C. ATT.....A.....G.....GC.....G.....C.....A.....G.....C.....A.....CC.....	495
Bet v 1n (<i>Betula</i>)	.G.....C.....T.....G.....C. ATT.....A.....G.....GC.....G.....C.....A.....G.....C.....A.....CC.....	495
Bet v 1c (<i>Betula</i>)	.G.....C.....T.....G.....C. ATT.....A.....G.....GC.....G.....C.....G.....C.....A.....CC.....	495
Bet v 1k (<i>Betula</i>)	.G.....C.....T.....G.....C. ATT.....A.....G.....GC.....G.....C.....G.....C.....A.....CC.....	495
Bet v 1l (<i>Betula</i>)	.G.....C.....G.T.....CA AGT.....G.....C.....G.....C.....G.....T.....C.....A.....CC.....	495
Bet v 1b (<i>Betula</i>)	.G.....C.....T.....G.....C. ATT.....A.....G.....GC.....G.....C.....G.....C.....A.....CC.....	495
Aln g 1 (<i>Alnus</i>)	C.....T.....AT.....AA.....A.....T.....GG.....C.C.A.....C.....G.....C.....A.....CC.....	495
cagc10 (<i>Corylus</i>)	C.....T.....T.....G.....A.....T.....GG.....T.....C.....G.....C.....C.....A.....CC.....	495
Bet v 1-Sc3 (<i>Betula</i>)	...AG.....T.....CC.....G.....A.....G.....CCGG.....C.A.....T.....G.....CG.....C.A.....AA.....CC.....	495
cagc11 (<i>Corylus</i>)	.G.AC.....C.....T.....C.....GA.....G.....TCAGG.....C.AA.....T.....G.....C.....C.....A.....CC.....	495
Mal d 1 (<i>Malus</i>)	.G.A.....C.....CG.....TC.....GA.....G.....GTC.TGG TT.G.....C.A.....CTTA.....G.....CAA.....GC.....C.G.....A.....CC.....	495
L1R18A (<i>Lupinus</i>)	CA.A.ACT.T.....CGTG.TCAA.....A.....TTC.....A.....GACTTGG.....T.A.....AA.....AG GT.....TG.T.....TC.....A.T.C.....G. NNNNN	495
LEDI-1 (<i>Lithospermum</i>)	C.AG.....T.....AG.T.....AG AG.....G.....G.....TCTTGG.....C.T.A.....T.....AG.....TC.TG.C.....CA.....AA NNNNNNNNN NNNNN	495
Gea 20 (<i>Daucus</i>)	C.AG.....A.....C.....TT.....AG.T.CTC A.AA.ACTGC T.....C.A.....TA.....G.....CA.....T.....TA.T.AA NNNNNNNNN NNNNN	495
pSTH-2 (<i>Solanum</i>)	.G.AT.....A.....ACA.....T.AA.....G.C.....A.....C.....AA.G.AT.G.....C.A.....ATT.....AG.....A.....CC.....C.....CA.TC.....T.....TC.....G CT...	495
AoPR1 (<i>Asparagus</i>)	G...GA...C.....AGG...A.....ATA.CCA.C- ----- -ACA.G.A.....AACC.....AG.....T.....CC.....A.....CA.....C.A.....CC...G TT...	495

ACKNOWLEDGMENTS

This work was supported in part by Grant P11410-MOB of the Austrian "Fonds zur Foerderung der wissenschaftlichen Forschung" (to O.S.) and a Career Enhancement Grant from the Colorado State University (to J.W.). We thank Dick Olmstead and Roberta Mason-Gamer for providing constructive and helpful comments on the manuscript.

REFERENCES

- Atkinson, R. G., Perry, J., Matsui, T., Ross, G. S., and Macrae, E. A. (1996). A stress-, pathogenesis-, and allergen-related cDNA in apple fruit is also ripening-related. *New Zealand J. Crop. Hort. Sci.* **24**: 103–107.
- Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., and Donoghue, M. J. (1995). The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Ann. Missouri Bot. Gard.* **82**: 247–277.
- Bousquet, J., Strauss, S. H., and Li, P. (1992). Complete congruence between morphological and rbcL-based molecular phylogenies in birches and related species (Betulaceae). *Mol. Biol. Evol.* **9**: 1076–1088.
- Breiteneder, H., Pettenburger, K., Bito, A., Valenta, R., Kraft, D., Rumpold, H., Scheiner, O., Breitenbach, M. (1989). The gene coding for the major birch pollen allergen Betv1, is highly homologous to a pea disease resistance response gene. *EMBO J.* **8**: 1935–1938.
- Breiteneder, H., Ferreira, F., Reikerstorfer, A., Duchene, M., Valenta, R., Hoffmann-Sommergruber, K., Ebner, C., Breitenbach, M., Kraft, D., and Scheiner, O. (1992). Complementary DNA cloning and expression in *Escherichia coli* of Aln g 1, the major allergen in pollen of alder (*Alnus glutinosa*). *J. Allergy Clin. Immunol.* **90**: 909–917.
- Breiteneder, H., Ferreira, F., Hoffmann-Sommergruber, K., Ebner, C., Breitenbach, M., Rumpold, H., Kraft, D., and Scheiner, O. (1993). Four recombinant isoforms of Cor a I, the major allergen of hazel pollen, show different IgE-binding properties. *Eur. J. Biochem.* **212**: 355–362.
- Breiteneder, H., Hoffmann-Sommergruber, K., O'Riordain, G., Susani, M., Ahorn, H., Ebner, C., Kraft, D., and Scheiner, O. (1995). Molecular characterization of Api g 1, the major allergen of celery (*Apium graveolens*), and its immunological and structural relationships to a group of 17-kDa tree pollen allergens. *Eur. J. Biochem.* **233**: 484–489.
- Bremer, K. (1988). The limits of amino acid sequence in angiosperm phylogenetic reconstruction. *Evolution* **42**: 795–803.
- Cervantes, E., Nicolas, C., Gonzales, B., and Iturriaga, E. A. (1995). Isolation and expression characteristics of CaPR-10a, an IPR coding cDNA from chickpea seeds. *Plant Physiol.* **110**: 335.
- Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., Duvall, M. R., Price, R. A., Hills, H. G., Qiu, Y.-L., Kron, K. A., Rettig, J. H., Conti, E., Palmer, J. D., Manhart, J. R., Sytsma, K. J., Michaels, H. J., Kress, J. W., Karol, K. G., Clark, W. D., Hedren, M., Gaut, B. S., Jansen, R. K., Kim, K.-J., Wimpee, C. F., Smith, J. F., Furnier, G. R., Strauss, S. H., Xiang, Q.-Y., Plunkett, G. M., Soltis, P. S., Swensen, S. M., Williams, S. E., Gadek, P. A., Quinn, C. J., Equarte, L. E., Golenberg, E., Learn, G. H., Jr., Graham, S. W., Barrett, S. C. H., Dayanandan, S., and Albert, V. A. (1993). Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann. Missouri Bot. Gard.* **80**: 528–580.
- Chiang, C. C., and Hadwiger, L. A. (1990). Cloning and characterization of a disease resistance response gene in pea inducible by *Fusarium solani*. *Mol. Plant-Microbe Interact.* **3**: 78–85.
- Clegg, M. T. (1993). Chloroplast gene sequences and the study of plant evolution. *Proc. Natl. Acad. Sci. USA* **90**: 363–367.
- Crane, P. R. (1989). Early fossil history and evolution of the Betulaceae. In "Evolution, Systematics, and Fossil History of the Hama-

- melidae" (P. R. Crane, and S. Blackmore, Eds.), Vol. 2, pp. 87–116, Clarendon, Oxford. [Systematics Association Special Volume No. 40B].
- Crisp, M. D., and Doyle, J. J. (Eds.) (1995). "Advances in Legume Systematics. Part 7. Phylogeny," The Royal Botanic Gardens, Kew.
- Cronquist, A. (1988). "The Evolution and Classification of Flowering Plants," 2nd ed., The New York Botanical Garden, Bronx, NY.
- Crowell, D. N., John, M. E., Russell, D., and Amasino, R. M. (1992). Characterization of a stress-induced, developmentally regulated gene family from soybean. *Plant Mol. Biol.* **18**: 459–466.
- Daniels, C. H., Fristensky, B., Wagoner, W., and Hadwiger, L. A. (1987). Pea genes associated with non-host disease resistance to *Fusarium* are also active in race-specific disease resistance to *Pseudomonas*. *Plant Mol. Biol.* **8**: 309–316.
- Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**: 387–395.
- Doyle, J. J. (1991). Evolution of higher plant glutamine synthetase genes: Regulatory specificity as a criterion for predicting orthology. *Mol. Biol. Evol.* **8**: 366–377.
- Doyle, J. J. (1992). Gene trees and species trees: Molecular systematics as one-character taxonomy. *Syst. Bot.* **17**: 144–163.
- Doyle, J. J. (1995). DNA data and legume phylogeny: A progress report. In "Advances in Legume Systematics, Part 7. Phylogeny" (M. Crisp, and J. J. Doyle, Eds.), pp. 11–30, The Royal Botanic Gardens, Kew.
- Doyle, J. J., Lavin, M., and Bruneau, A. (1992). Contributions of molecular data to Papilionoid legume systematics. In "Molecular Systematics of Plants" (P. S. Soltis, D. E. Soltis, and J. J. Doyle, Eds.), pp. 223–251, Chapman & Hall, NY.
- Esnault, R., Buffard, D., Breda, C., Salaud, C., el Turk, J., and Kondorosi, A. (1993). Pathological and molecular characterization of alfalfa interactions with compatible and incompatible bacteria, *Xanthomonas campestris* pv. *alfalfae* and *Pseudomonas syringae* pv. *pisii*. *Mol. Plant-Microbe Interact.* **6**: 655–664.
- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **17**: 368–376.
- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**: 783–791.
- Felsenstein, J. (1993). "PHYLIP (Phylogeny Inference Package), Version 3.5c." [distributed by the author, Department of Genetics, Univ. of Washington, Seattle].
- Fristensky, B., Riggelman, R., Wagoner, W., and Hadwiger, L. A. (1985). Gene expression in susceptible and disease resistant interactions of peas induced with *Fusarium solani* pathogens and chitosan. *Phys. Plant Pathol.* **27**: 15–28.
- Fristensky, B., Horovitz, D., and Hadwiger, L. A. (1988). cDNA sequences for pea disease resistance response genes. *Plant Mol. Biol.* **11**: 713–715.
- Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., van Neerven, R. J., Schou, C., Loewenstein, H., and Spangfort, M. (1996). X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nature Struct. Biol.* **3**: 1040–1045.
- Gottlieb, L. D., and Ford, V. S. (1996). Phylogenetic relationships among the sections of *Clarkia* Onagraceae inferred from the nucleotide sequences of PgiC. *Syst. Bot.* **21**: 45–62.
- Hamby, R. K., and Zimmer, E. A. (1992). Ribosomal RNA as a phylogenetic tool in plant systematics. In "Molecular Systematics of Plants" (P. S. Soltis, D. E. Soltis, and J. J. Doyle, Eds.), pp. 50–91, Chapman & Hall, NY.
- Hoffmann-Sommergruber, K., Vanek-Krebitz, M., Radauer, C., Wen, J., Ferreira, F., Scheiner, O., and Breiteneder, H. (1997). Genomic characterization of members of the Bet v 1 family: Genes coding for allergens and pathogenesis-related proteins share intron positions. *Gene*. [in press]
- Hsieh, L.-S., Martin, B., and Hadwiger, L. (1996). Immunological cross-reactivity of the apple 18 kD allergen and constituents of common vegetables. *J. Allergy Clin. Immunol.* **97**: 335.
- Iturriaga, E. A., Leech, M. J., Barrat, D. H. P., and Wang, T. L. (1994). Two ABA-responsive proteins from pea (*Pisum sativum* L.) are closely related to intracellular pathogenesis-related proteins. *Plant Mol. Biol.* **24**: 235–240.
- Johnson, L. A., and Soltis, D. E. (1994). *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Syst. Bot.* **19**: 143–156.
- Kos, T., Hoffmann-Sommergruber, K., Ferreira, F., Hirschwehr, R., Ahorn, H., Horak, F., Jager, S., Sperr, W., Kraft, D., and Scheiner, O. (1993). Purification, characterization and N-terminal amino acid sequence of a new major allergen from European chestnut pollen—Cas s 1. *Biochem. Biophys. Res. Commun.* **196**: 1086–1092.
- Kouchi, H., Sekine, M., and Hata, S. (1995). Distinct classes of mitotic cyclins are differentially expressed in the soybean shoot apex during the cell cycle. *Plant Cell* **7**: 1143–1155.
- Kumar, S., Tamura, K., and Nei, M. (1993). "MEGA: Molecular Evolutionary Genetics Analysis, Version 1.01," Pennsylvania State Univ., University Park, PA. [distributed by Institute of Molecular Evolutionary Genetics]
- Larsen, J. N., Stroman, P., and Ipsen, H. (1992). PCR based cloning and sequencing of isogenes encoding the tree pollen major allergen Car b 1 from *Carpinus betulus*, hornbeam. *Mol. Immunol.* **29**: 703–711.
- Li, W.-H., and Graur, D. (1991). "Fundamentals of Molecular Evolution," Sinauer, Sunderland, MA.
- Mason-Gamer, R. J., and Kellogg, E. A. (1996). Potential utility of the nuclear gene *waxy* for plant phylogenetic analyses. *Amer. J. Bot.* **83**: 505. [Abstract]
- Mathews, S., Lavin, M., and Sharrock, R. A. (1995). Evolution of the Phytochrome gene family and its utility for phylogenetic analyses of angiosperms. *Ann. Missouri Bot. Gard.* **82**: 296–321.
- Mathews, S., and Sharrock, R. A. (1996). The phytochrome gene family in grasses (Poaceae): A phylogeny and evidence that grasses have a subset of the loci found in dicot angiosperms. *Mol. Biol. Evol.* **13**: 1141–1150.
- Matton, D. P., and Brisson, N. (1989). Cloning, expression, and sequence conservation of pathogenesis-related gene transcripts of potato. *Mol. Plant-Microbe Interact.* **2**: 325–331.
- Matton, D. P., Prescott, G., Bertrand, C., Camirand, A., and Brisson, N. (1993). Identification of *cis*-acting elements involved in the regulation of the pathogenesis-related gene STH-2 in potato. *Plant Mol. Biol.* **22**: 279–291.
- McElroy, D., Rothenberg, M., Reece, K. S., and Wu, R. (1990). Characterization of the rice actin gene family. *Plant Mol. Biol.* **15**: 257–268.
- Meagher, R., Berry-Lowe, S., and Rice, K. (1989). Molecular evolution of the small subunit of ribulose biphosphate carboxylase: Nucleotide substitution and gene conversion. *Genetics* **123**: 845–863.
- Mylona, P., Moerman, M., Yang, W. C., Gloudemans, T., van de Kerckhove, J., van Kammen, A., Bisseling, T., and Franssen, H. J. (1994). The root epidermis-specific pea gene RH2 is homologous to a pathogenesis-related gene. *Plant Mol. Biol.* **26**: 39–50.
- Nei, M. (1987). "Molecular Evolutionary Genetics," Columbia Univ. Press, New York.
- Olmstead, R. G., and Sweere, J. A. (1994). Combining data in phylogenetic systematics: An empirical approach using three molecular data sets in the Solanaceae. *Syst. Biol.* **43**: 467–481.

- Polhill, R. M., and Raven, P. H. (Eds.) (1981). "Advances in Legume Systematics," Parts 1 and 2, The Royle Botanic Gardens, Kew.
- Riggleman, R. C., Fristensky, B., and Hadwiger, L. A. (1985). The disease resistance response in pea is associated with increased levels of specific mRNAs. *Plant Mol. Biol.* **4**: 81–86.
- Russo, C. A., Takezaki, N., and Nei, M. (1995). Molecular phylogeny and divergence times of Drosophilid species. *Mol. Biol. Evol.* **12**: 391–404.
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.
- Sanderson, M. J., and Doyle, J. J. (1992). Reconstruction of organismal and gene phylogenies from data on multigene families: Concerted evolution, homoplasy, and confidence. *Syst. Biol.* **41**: 4–17.
- Schoening, B., Vieths, S., Petersen, A., and Baltes, W. (1995). Identification and characterization of allergens related to Bet v I, the major birch pollen allergen, in apple, cherry, celery and carrot by two-dimensional immunoblotting and N-terminal microsequencing. *J. Sci. Food Agric.* **67**: 431–440.
- Sikorski, M. M., Szlagowska, A. E., and Legocki, A. B. (1995). cDNA sequences encoding for two homologues of *Lupinus luteus* (L.) IPR-like proteins. *Plant Physiol.* **110**: 335.
- Soltis, D. E., Soltis, P. S., Nickrent, D. L., Johnson, L. A., Hahn, W. J., Hoot, S. B., Sweere, J. A., Kuzoff, R. K., Kron, K. A., Chase, M. W., Swensen, S. M., Zimmer, E. A., Chaw, S.-M., Gillespie, L. J., Kress, W. J., and Sytsma, K. J. (1997). Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Ann. Missouri Bot. Gard.* **84**: 1–49.
- Somssich, I. E., Schmelzer, E., Bollmann, J., and Hahlbrock, K. (1986). Rapid activation by fungal elicitor of genes encoding 'pathogenesis-related' proteins in cultured parsley cells. *Proc. Natl. Acad. Sci. USA* **83**: 2427–2430.
- Somssich, I. E., Schmelzer, E., Kawalleck, P., and Hahlbrock, K. (1988). Gene structure and in situ transcript localization of pathogenesis-related protein 1 in parsley. *Mol. Gen. Genet.* **213**: 93–98.
- Swoboda, I., Jilek, A., Ferreira, F., Engel, E., Hoffmann-Sommergruber, K., Scheiner, O., Kraft, D., Breiteneder, H., Pittenauer, E., Schmid, E., Vicente, O., Heberle-Bors, E., Ahorn, H., and Breitenbach, M. (1995a). Isoforms of Bet v 1, the major birch pollen allergen, analyzed by liquid chromatography, mass spectrometry, and cDNA cloning. *J. Biol. Chem.* **270**: 2607–2613.
- Swoboda, I., Scheiner, O., Heberle-Bors, E., and Vicente, O. (1995b). cDNA cloning and characterization of three genes in the Bet v 1 gene family that encode pathogenesis-related proteins. *Plant Cell Environ.* **18**: 865–874.
- Swofford, D. L. (1993). "PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1." [distributed by the Illinois Natural History Survey, Champaign, IL]
- Takaiwa, F., Oono, K., Iida, Y., and Sugiura, M. (1985). The complete nucleotide sequence of a rice 25S ribosomal RNA gene. *Gene* **37**: 255–289.
- van de Loecht, U., Meier, I., Hahlbrock, K., and Somssich, I. (1990). A 125 bp promoter fragment is sufficient for strong elicitor-mediated gene activation in parsley. *EMBO J.* **9**: 2945–2950.
- Vanek-Krebitz, M., Hoffmann-Sommergruber, K., Laimer da Camara Machado, M., Susani, M., Ebner, C., Kraft, D., Scheiner, O., and Breiteneder, H. (1995). Cloning and sequencing of Mal d 1, the major allergen from apple (*Malus domestica*), and its immunological relationship to Bet v 1, the major birch pollen allergen. *Biochem. Biophys. Res. Commun.* **214**: 538–551.
- Walter, M. H., Liu, J. W., Grand, C., Lamb, C. J., and Hess, D. (1990). Bean pathogenesis-related (PR) proteins deduced from elicitor-induced transcripts are members of a ubiquitous new class of conserved PR proteins including pollen allergens. *Mol. Gen. Genet.* **222**: 353–360.
- Walter, M. H., Liu, J. W., Wuenn, J., and Hess, D. (1996). Bean ribonuclease-like pathogenesis-related protein genes (Ypr10) display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. *Eur. J. Biochem.* **239**: 281–293.
- Warner, S. A. J., Scott, R., and Draper, J. (1992). Characterisation of a wound-induced transcript from the monocot asparagus that shares similarity with a class of intracellular pathogenesis-related (PR) proteins. *Plant Mol. Biol.* **19**: 555–651.
- Warner, S. A. J., Scott, R., and Draper, J. (1993). Isolation of an asparagus intracellular PR gene (AoPR1) wound-responsive promoter by the inverse polymerase chain reaction and its characterization in transgenic tobacco. *Plant J.* **3**: 191–201.
- Warner, S. A. J., Gill, A., and Draper, J. (1994). The developmental expression of the asparagus intracellular PR protein (AoPR1) gene correlates with sites of phenylpropanoid synthesis. *Plant J.* **6**: 31–43.
- Waters, E. R. (1995). An evaluation of the usefulness of the small heat shock genes for phylogenetic analysis in plants. *Ann. Missouri Bot. Gard.* **82**: 278–295.
- Yazaki, K., Bechthold, A., and Tabata, M. (1995). Nucleotide sequence of a cDNA from *Lithospermum erythrorhizon* homologous to PR-1 of parsley. *Plant Physiol.* **108**: 1331–1332.
- Zimmer, E. A., Martin, S. L., Beverly, S. M., Kan, Y. W., and Wilson, A. C. (1980). Rapid duplication and loss of genes coding for the chains of hemoglobin. *Proc. Natl. Acad. Sci. USA* **77**: 2158–2162.