Molecular Phylogenetics and Evolution
Vol. 20, No. 2, August, pp. 262–274, 2001
doi:10.1006/mpev.2001.0973, available online at http://www.idealibrary.com on IDEAL®

New Partial Sequences of Phosphoenolpyruvate Carboxylase as Molecular Phylogenetic Markers

Hans Gehrig,*,[1] Valentina Heute,† and Manfred Kluge†

*Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002-0948, Panama; and †Darmstadt University of Technology, Schnittspahnstr. 10, 64287 Darmstadt, Germany

Received October 10, 2000; revised March 12, 2001

To better understand the evolution of the enzyme phosphoenolpyruvate carboxylase (PEPC) and to test its versatility as a molecular character in phylogenetic and taxonomic studies, we have characterized and compared 70 new partial PEPC nucleotide and amino acid sequences (about 1100 bp of the 3′ side of the gene) from 50 plant species (24 species of Bryophyta, 1 of Pteridophyta, and 25 of Spermatophyta). Together with previously published data, the new set of sequences allowed us to construct the up to now most complete phylogenetic tree of PEPC, where the PEPC sequences cluster according to both the taxonomic positions of the donor plants and the assumed specific function of the PEPC isoforms. Altogether, the study further strengthens the view that PEPC sequences can provide interesting information for the reconstruction of phylogenetic relations between organisms and metabolic pathways. To avoid confusion in future discussion, we propose a new nomenclature for the denotation of PEPC isoforms. © 2001 Academic Press

Key Words: **crassulacean acid metabolism (CAM); photosynthesis (C_3, C_4); molecular taxonomy; molecular evolution; phosphoenolpyruvate carboxylase (PEPC).**

INTRODUCTION

The comparison of organisms on the level of molecular characters has become a powerful and now indispensable tool in taxonomic and phylogenetic research. However, the unequivocalness of results obtained by this approach depends essentially on the availability of versatile molecular markers. In plant sciences mainly four types of nucleotide sequences are used as such markers, namely the 18s rRNA (e.g., Bopp and Capesius, 1996; Qiu and Palmer, 1999), ITS regions (e.g., Bruns *et al.*, 1991; Bogler and Simpson, 1996), MADS-box genes (e.g., Winter *et al.*, 1999), and the rbcL genes coding for the large subunit of RUBISCO (e.g., Dressler and Chase, 1995; Yukawa *et al.*, 1996; Qiu and Palmer, 1999). Because the obvious limitation in the assortment of suitable markers may be one reason for controversial interpretations of obtained results, it is highly desired that more markers become available for taxonomic and phylogenetic studies in plant sciences. Searching for such markers, we investigated and compared sequences of phosphoenolpyruvate carboxylase (PEPC; EC 4.1.1.31).

PEPC catalyzes the β-carboxylation of phosphoenolpyruvate, with oxaloacetate and inorganic phosphate as products (Utter and Kolenbrander, 1972; Andreo *et al.*, 1987). The enzyme is ubiquitous in prokaryotic microorganisms and plants, and it is involved in many functions including photosynthetic and anaplerotic CO_2 fixation (e.g., Kluge and Ting, 1978; Winter, 1985; Cushman and Bohnert, 1999; Latzko and Kelly, 1983), production of carbon skeletons in symbiotic nitrogen fixation (Schuller *et al.*, 1990), modulation of turgor in stomatal guard cells, maintenance of ion balance, pH state mechanisms, and others (Latzko and Kelly, 1983; Melzer and O'Leary, 1987). In most bacteria and plants studied so far, physiological and molecular approaches showed the existence of PEPC multigene families (Cushman and Bohnert, 1989a,b; Crétin *et al.*, 1991; Poetsch *et al.*, 1991; Kawamura *et al.*, 1992; Lepiniec *et al.*, 1993, 1992; Gehrig *et al.*, 1995) encoding function- and tissue-specific isoforms of the enzyme (Lepiniec *et al.*, 1994; Toh *et al.*, 1994; Rajagopalan *et al.*, 1994; Gehrig *et al.*, 1998b). Because of the ubiquitous distribution of PEPC and the high diversity in its functions it has been proposed that the nucleotide sequences of the PEPC genes and the amino acid sequences of the gene products should provide powerful markers in molecular taxonomic and phylogenetic investigations (Gehrig *et al.*, 1998a).

First attempts to construct phylogenetic trees of PEPC were based on full-length sequences (Lepiniec *et al.*, 1993; Toh *et al.*, 1994; Cushman and Bohnert, 1996; Honda *et al.*, 1996). Gehrig *et al.* (1998a) showed that the comparison of partial PEPC sequences can provide valuable information on the phylogenetic in-

[1] To whom correspondence and reprint requests should be addressed. Fax: (00507)-212-8148. E-mail: hansgehrig@gmx.de.

1055-7903/01 $35.00

terrelationships of PEPC isoforms and the donor plants from which the genes were isolated. Although up to now full-length and partial PEPC nucleotide sequences of 11 prokaryotes (11 PEPC sequences), 1 alga (2 PEPC sequences), 5 ferns (5 PEPC sequences) and 25 higher plants (48 PEPC sequences) have been determined (Toh *et al.,* 1994; Chollet *et al.,* 1996; Gehrig *et al.,* 1998a) our present knowledge on the phylogeny of PEPC sequences and the isoforms in species is still quite fragmentary and urgently requires completion. For this reason we have analyzed PEPC sequences in numerous further plant species including Bryophyta, Pteridophyta, and Spermatophyta. The present paper reports on the obtained results.

MATERIALS AND METHODS

Plant Material

The experimental plants were cultivated in the Botanical Gardens of Heidelberg and Darmstadt (Germany). The plant material was thoroughly cleaned with sterilized water and by ultrasonic treatment. After the cleaning, tissue samples for RNA preparation were immediately frozen in liquid nitrogen and stored at −80°C until further processing.

RNA Extraction and PCR Amplification

Total RNA was extracted with the guanidine isothiocyanate method (Chirgwin *et al.,* 1979) or with the QIAGEN plant RNA isolation kit (Qiagen, Germany), depending on the plant material. RNA quality was examined by agarose gel electrophoresis, and after reverse transcription an approx 1100-bp PEPC fragment was amplified by RT-PCR. The RT-PCR was performed with two degenerated PEPC primers [PEPC1: TC-(CTA) GA(TC) TC(CAT) GG(AC) AA(AG) GA(TC) GC; PEPC2: GC(GAT) GC(GAT) AT(GCA) CC(CT) TTC AT(GT) G] under the following conditions: 35 cycles (Personal Cycler; Biometra, Germany) at 95°C for 30 s, 55°C for 30 s, 72°C for 3 min. The PCR products were cloned into the TA vector system of Invitrogen (Netherlands). Different PEPC isoform clones of the *Kalanchoe* and orchid species were identified by digestion with *Bam*HI, *Hind*III, *Pst*I, *Sal*I, and *Eco*RI and analyzed on 0.8% agarose gels.

Randomly selected transformants of each amplified PEPC cDNA clone were sequenced in both directions (SeqLab Co., Hannover, Germany). The nucleotide sequence data have been submitted to the EMBL and GenBank nucleotide sequence databases (see Table 1 for accession numbers).

Sequence Analysis

Amino acid sequences were obtained from 143 nucleotide PEPC sequences. The alignment and sequence identity were calculated for each pair with the Alignment Editor 3.7 (Hepperle, 1997). The alignment obtained was modified by visual inspection to increase the total alignment score (the alignment is available from the authors). Sequence data were evaluated by means of the PHYLIP package, version 3.5c (Felsenstein, 1993). Neighbor-joining analysis (Saitou and Nei, 1987) was employed as a distance method (PROTDIST) with 1000 resamplings with the Kimura formula for amino acid sequences (Kimura, 1983) of the PHYLIP package. This is a rough-and-ready distance formula for approximating PAM distance that simply measures the fraction of amino acids that differs between two sequences. Parsimony analyses with 100 resamplings were done with the PROTPARS program of the PHYLIP package. This program infers an unrooted phylogeny directly from protein sequences (for further explanations see Felsenstein, 1993). In all cases *Escherichia coli* was used as outgroup.

RESULTS AND DISCUSSION

As already mentioned in the Introduction, the PEPC is coded by multigene families, with isoforms being linked to a wide range of different functions. Up to now there has been no generally followed terminology of PEPC sequences suitable for relating a given PEPC isoform to a specific function. However, to avoid confusion, definition of a common nomenclature to denote PEPC isoforms is highly desired. Thus, we propose to distinguish and to denote PEPC isoforms as follows (Table 2): prokaryotic anaplerotic and other nonphotosynthetic isoforms (ppc-aP), eukaryotic anaplerotic and nonphotosynthetic isoforms (ppc-aX; with X standing for R = root, aerial root, root nodule, and for L = leaf), and photosynthetic isoforms catalyzing the primary CO_2 fixation in C4 photosynthesis and crassulacean acid metabolism (CAM) (respectively, ppc-C4 and ppc-CAM). The term “C3” isoform often used in the literature to describe anaplerotic or “housekeeping” PEPC isoforms in leaves of C_3 plants (e.g., Lepiniec *et al.,* 1994; Toh *et al.,* 1994; Rajagopalan *et al.,* 1994) should be avoided, because PEPC is not directly involved in the C_3 pathway of photosynthesis. In the present paper we will follow the terminology proposed here.

In the context of the question whether PEPC can serve as a useful molecular marker in taxonomic and phylogenetic investigations, in the present study we have analyzed numerous new partial PEPC sequences. The results not only increase the number of plants species that can be compared on the level of PEPC sequences but also include more taxa that are supposed to mark branching points of plant evolution. Altogether, in the present study 70 new PEPC sequences were analyzed and documented in the EMBL gene bank (Table 1). The new sequences represent 24 species of Bryophyta, 1 species of Pteridophyta, and 25 species of Spermatophyta.

As previously done (Gehrig *et al.,* 1998a), in the

TABLE 1

PEPC Partial Sequences Included in the Calculations of the Phylogenetic Trees Shown in Figs. 1–3

Organisms	Taxonomic unit	Accession No.	References
Corynebacterium glutamicum 1	Bacteria (α subdivision)	X14234	Eikmanns *et al.* (1989)
Corynebacterium glutamicum 2	Bacteria (α subdivision)	M25819	Viret and Lemoine (1989)
Mycobacterium leprae	Bacteria (α subdivision)	U00013	Robinson, K. (unpublished)
Rhodopseudomonas palustris	Bacteria (α subdivision)	D89668	Inui *et al.* (1997)
Rhodothermus obamensis	Bacteria (α subdivision)	X99379	Takai *et al.* (1998)
Thermus sp.	Bacteria (α subdivision)	D42166	Nakamura *et al.* (1995)
Escherichia coli	Bacteria (γ subdivision)	X05903	Fujita *et al.* (1984)
Haemophilus influenzae	Bacteria (γ subdivision)	U00086	Fleischmann *et al.* (1995)
Anabaena variabilis	Cyanophyaceae	M80541	Luinenburg and Coleman (1992)
Anacystis nidulans	Cyanophyaceae	M11198	Katagiri *et al.* (1995)
Synechocystis sp. *PCC6803*	Cyanophyaceae	AB001339	Kaneko *et al.* (1996)
Chara fragilis 1	Charophyceae	X95851	Gehrig *et al.* (1998a)
Chara fragilis 2	Charophyceae	X95857	Gehrig *et al.* (1998a)
Anthoceros agrestis	Anthocerotae	AJ231277	This study
Anthoceros punctatus	Anthocerotae	AJ231278	This study
Bucegia romanica	Hepaticae	AJ231280	This study
Conocephalum conicum	Hepaticae	X95853	Gehrig *et al.* (1998a)
Fossombronia pusilla	Hepaticae	AJ231304	This study
Jungermannia leiantha	Hepaticae	AJ231287	This study
Lunularia cruciata	Hepaticae	AJ231289	This study
Marchantia calcarata	Hepaticae	AJ231292	This study
Preissia quadrata	Hepaticae	AJ231297	This study
Scapania nemorea	Hepaticae	AJ231300	This study
Symphyogyna brongniartii	Hepaticae	AJ231299	This study
Batramia pomiformis	Musci	AJ231279	This study
Brachythecium salebrosum	Musci	AJ231281	This study
Calliergonella cuspidata	Musci	AJ231282	This study
Dicranella heteromalla	Musci	AJ231283	This study
Dicranum scoparium	Musci	AJ231284	This study
Funaria hygrometrica	Musci	AJ231285	This study
Hypnum cupressiforme	Musci	AJ231286	This study
Leptobryum pyriforme	Musci	AJ231291	This study
Leucobryum juniperiodeum	Musci	AJ231290	This study
Polytrichum commune	Musci	AJ231294	This study
Polytrichum formosum	Musci	AJ231296	This study
Rhytidiadelphus squarrosus	Musci	AJ231298	This study
Scleropodium purum	Musci	AJ231302	This study
Sphagnum sp.	Musci	X95852	Gehrig *et al.* (1998a)
Sphagnum palustre	Musci	AJ231301	This study
Isoetes histrix	Lycopodiatae	X95854	Gehrig *et al.* (1998a)
Isoetes duriei	Lycopodiatae	X95859	Gehrig *et al.* (1998a)
Lycopodium annotium	Lycopodiatae	X95858	Gehrig *et al.* (1998a)
Selaginella martinii	Lycopodiatae	AJ252913	This study
Psilotum nudum	Psilotatae	X91405	Gehrig *et al.* (1998a)
Equisetum hyemale	Equisetatae	X95855	Gehrig *et al.* (1998a)
Picea abies 1	Pinatae	X79090	Relle and Wild (1996)
Picea abies 2	Pinatae	P51063	Relle and Wild (1996)
Welwitschia mirabilis	Gnetatae	X91404	Gehrig *et al.* (1998a)
Saccharum hybride	Poaceae	M86661	Henrik *et al.* (1992)
Sorghum vulgare 1	Poaceae	X59925	Lepiniec *et al.* (1991)
Sorghum vulgare 2	Poaceae	X65137	Cretin *et al.* (1990)
Sorghum vulgare 3	Poaceae	X63756	Cretin *et al.* (1990)
Triticum aestivum	Poaceae	AJ007705	Gonzalez *et al.* (1998)
Zea mays 1	Poaceae	X03613	Izui *et al.* (1986)
Zea mays 2	Poaceae	X15239	Hudspeth and Grula (1989)
Zea mays 3	Poaceae	X15238	Hudspeth and Grula (1989)
Zea mays 4	Poaceae	X61489	Kawamura *et al.* (1992)
Zea mays 5	Poaceae	AB012228	Dong, L. (unpublished)
Zea mays 6	Poaceae	E01120	Katsuki, H. (unpublished)
Arabidopsis thaliana	Brassicaceae	AJ131710	Hartung, F. (unpublished)
Brassica juncea 1	Brassicaceae	AJ223496	Heiss, S. (unpublished)
Brassica juncea 2	Brassicaceae	AJ223497	Heiss, S. (unpublished)
Brassica napus	Brassicaceae	D13987	Yanai *et al.* (1994)

TABLE 1—*Continued*

Organisms	Taxonomic unit	Accession No.	References
Glycine max 1	Fabaceae	D13998	Tello *et al.* (1993)
Glycine max 2	Fabaceae	D10717	Sugimoto *et al.* (1992)
Glycine max 3	Fabaceae	AB008540	Hata *et al.* (1997)
Medicago sativa	Fabaceae	M83086	Pathirana *et al.* (1992)
Pisum sativum	Fabaceae	D64037	Suganuma *et al.* (1997)
Vicia faba 1	Fabaceae	AJ011302	Golombek, S. (unpublished)
Vicia faba 2	Fabaceae	AJ011303	Golombek, S. (unpublished)
Hydrilla verticillata 1	Hydrocharitaceae	U65226	Magnin *et al.* (1996)
Hydrilla verticillata 2	Hydrocharitaceae	U65227	Magnin *et al.* (1996)
Nicotiana tabacum 1	Solanaceae	X59016	Koizumi *et al.* (1991)
Nicotiana tabacum 2	Solanaceae	E03014	Yamada, Y. and Sato, F. (unpublished)
Solanum tuberosum 1	Solanaceae	X67053	Merkelbach *et al.* (1993)
Solanum tuberosum 2	Solanaceae	X90982	Panstruga, R. (unpublished)
Flaveria australasica	Asteraceae	Z25853	Bauwe, H. (unpublished)
Flaveria pringlei 1	Asteraceae	X64144	Hermans and Westhoff (1992)
Flaveria pringlei 2	Asteraceae	Z48966	Svensson *et al.* (1997)
Flaveria trinervia 1	Asteraceae	X61304	Poetsch *et al.* (1991)
Flaveria trinervia 2	Asteraceae	X64143	Hermans and Westhoff (1992)
Drosanthemum paxianum	Aizoaceae	Y17844	This study
Mesembryanthemum crystallinum 1	Aizoaceae	X14588	Cushman and Bohnert (1989a)
Mesembryanthemum crystallinum 2	Aizoaceae	X14587	Cushman and Bohnert (1989b)
Mesembryanthemum crystallinum 3	Aizoaceae	X13660	Rickers *et al.* (1989)
Pereskia aculeata	Cactaceae	X95860	Gehrig *et al.* (1998a)
Selenicereus vitii	Cactaceae	Y17843	This study
Aechmea filicaulis	Bromeliaceae	AJ252914	This study
Neoregelia ampullacea	Bromeliaceae	X95861	Gehrig *et al.* (1998a)
Tillandsia usneoides	Bromeliaceae	X91406	Gehrig *et al.* (1998a)
Aloe arborescens	Asphodelaceae	D83052	Honda *et al.* (1996)
Amaranthus hypochondriacus	Amarantaceae	Z68125	Rydzik and Berry (1996)
Gossypium hirsutum 1	Malvaceae	AF008939	Vodjani *et al.* (1997)
Gossypium hirsutum 2	Malvaceae	AF008940	Vodjani *et al.* (1997)
Kalanchoe blossfeldiana 1	Crassulaceae	X87818	Gehrig *et al.* (1995)
Kalanchoe blossfeldiana 2	Crassulaceae	X87819	Gehrig *et al.* (1995)
Kalanchoe blossfeldiana 3	Crassulaceae	X87820	Gehrig *et al.* (1995)
Kalanchoe blossfeldiana 4	Crassulaceae	X87821	Gehrig *et al.* (1995)
Kalanchoe fedtschenkoi	Crassulaceae	AJ0010	Menke, H. H. and H. Gehrig (unpublished)
Kalanchoe gracilipes	Crassulaceae	AJ231288	This study
Kalanchoe grandiflora 1	Crassulaceae	AJ252918	This study
Kalanchoe grandiflora 2	Crassulaceae	AJ252945	This study
Kalanchoe x kewensis 1	Crassulaceae	AJ252914	This study
Kalanchoe x kewensis 2	Crassulaceae	AJ252915	This study
Kalanchoe petitiana 1	Crassulaceae	AJ231295	This study
Kalanchoe petitiana 2	Crassulaceae	AJ252926	This study
Kalanchoe pinnata 1	Crassulaceae	AJ252919	This study
Kalanchoe pinnata 2	Crassulaceae	AJ252920	This study
Kalanchoe pinnata 3	Crassulaceae	AJ252921	This study
Kalanchoe pinnata 4	Crassulaceae	AJ252922	This study
Kalanchoe streptantha 1	Crassulaceae	AJ252923	This study
Kalanchoe streptantha 2	Crassulaceae	AJ252924	This study
Kalanchoe streptantha 3	Crassulaceae	AJ252925	This study
Kalanchoe tomentosa 1	Crassulaceae	AJ252916	This study
Kalanchoe tomentosa 2	Crassulaceae	AJ252917	This study
Angraecum eburneum 1	Orchidaceae	X91636	This study
Angraecum eburneum 2	Orchidaceae	X91631	This study
Chiloschista pusilla	Orchidaceae	X91633	This study
Dendrobium crumenatum	Orchidaceae	AJ252938	This study
Dendrobium delicatum	Orchidaceae	AJ252944	This study
Dendrobium farmeri 1	Orchidaceae	AJ252939	This study
Dendrobium farmeri 2	Orchidaceae	AJ252940	This study
Dendrobium fimbriatum 1	Orchidaceae	AJ252942	This study
Dendrobium fimbriatum 2	Orchidaceae	AJ252943	This study
Dendrobium loddigesie 1	Orchidaceae	AJ252933	This study
Dendrobium loddigesie 2	Orchidaceae	AJ252934	This study
Dendrobium mochentum	Orchidaceae	AJ252941	This study
Dendrobium thyrsifolium 1	Orchidaceae	AJ252935	This study

TABLE 1—*Continued*

Organisms	Taxonomic unit	Accession No.	References
Dendrobium thyrsifolium 2	Orchidaceae	AJ252936	This study
Dendrobium thyrsifolium 3	Orchidaceae	AJ252937	This study
Microcoelia exilis	Orchidaceae	X91635	This study
Solenangis aphylla	Orchidaceae	X91632	This study
Vanilla aphylla 1	Orchidaceae	X91634	This study
Vanilla aphylla 2	Orchidaceae	AJ252927	This study
Vanilla aphylla 3	Orchidaceae	AJ252928	This study
Vanilla phalaenopsis 1	Orchidaceae	AJ252948	This study
Vanilla phalaenopsis 2	Orchidaceae	AJ252930	This study
Vanilla phalaenopsis 3	Orchidaceae	AJ252931	This study
Vanilla phalaenopsis 4	Orchidaceae	AJ252932	This study
Vanilla planifolia 1	Orchidaceae	X87148	Gehrig *et al.* (1998b)
Vanilla planifolia 2	Orchidaceae	X87149	Gehrig *et al.* (1998b)
Vanilla planifolia 3	Orchidaceae	AJ249988/9	This study
Vanilla pompona	Orchidaceae	AJ252929	This study

present study we compared a distinct PEPC partial 1100-bp cDNA sequence and the amino acid sequence derived from it. The sequence is located on the 3′ side of the coding region of the PEPC gene and comprises the active center of the enzyme which is sufficiently conservative to reflect larger distances in taxonomic relations between the species. On the other hand, the fragment was found to be variable enough to allow statistically significant differentiation.

From the new and the already published sequences we obtained the most detailed phylogenetic trees of PEPC up to now available (Figs. 1–3). The trees were constructed by neighbor-joining calculations (Figs. 1A and 2; with statistics based on 1000 bootstrap resamplings) and by parsimony analyses (Figs. 1B and 3; 100 bootstrap resamplings). In both types of dendrograms, branches with bootstrap values below 50% were reduced to polytomies. Other distance calculations such as FITCH and KITSCH were also applied, but since the results showed the same main topology as that documented in the trees of Figs. 1–3, those data are not shown.

The comparison of the prokaryotic and eukaryotic PEPC sequences suggests the existence of an ancestral gene arising from the γ proteobacterial lineage. This supports the view of R. Kaemmerer (unpublished; cited in Cushman and Bohnert, 1999). Independent of the mode of calculation, the PEPC sequences of the prokaryotes form a distinct cluster (Fig. 1) with three separated branches (bootstrap support 93–100%) representing γ proteobacteria (*E. coli, Haemophilus influenza*), α proteobacteria (*Thermus* sp., *Rhodopseudomonas palustris, Rhodothermus obamensis, Mycobacterium leprae, Corynebacterium glutamicum*), and cyanobacteria (*Anacystis nidulans, Anabaena variabilis*). At the base of the cluster representing the eucaryotes, there is a small branch separating with high bootstrap support the two species of hornworts (Anthocerotae: *Anthoceros punctatus, A. agrestis*) from the branch comprising *Chara* and all the land plants. The finding that *Anthoceros* branches off before *Chara* from the lineage leading to the land plants (bootstrap support 100%) was unexpected because it is in contrast to the present widely held view that the land plants are monophyletic (e.g., Qui and Palmer, 1999). The Anthocerotae form a separate cluster in the trees presented in Fig 1. This result is contrary to those of Capesius (1995) and Bopp and Capesius (1996), which show the Anthocerotae located among the Jungermanniidae (Hepaticae), and to those of Mishler *et al.* (1994), Kenrick and Crane (1997a,b), and Graham and Wilcox (2001), which place the hornworts between the liver-

TABLE 2

Recommended Nomenclature for the Different Functional PEPC Isoforms: Previous and New Denotations

Origin of the isoform	Presumptive function	Previous denotation	Denotation after Toh *et al.* (1994)	New denotation
Prokaryotic cells	Anaplerotic	—	Bacterial	ppc-aP
Root	Anaplerotic	C_3R	C3-1	ppc-aR
Aerial root	Anaplerotic	C_3	—	ppc-aR
Root nodule	Anaplerotic	C_3	C3-2	ppc-aR
Leaf	Anaplerotic	C_3-	C3-3	ppc-aL
Cell culture	Anaplerotic	Housekeeping C_3	—	ppc-aL
Leaf	Primary carboxylation	C4 photosynthesis	C4	ppc-C4
Leaf	Primary carboxylation	CAM	CAM	ppc-CAM

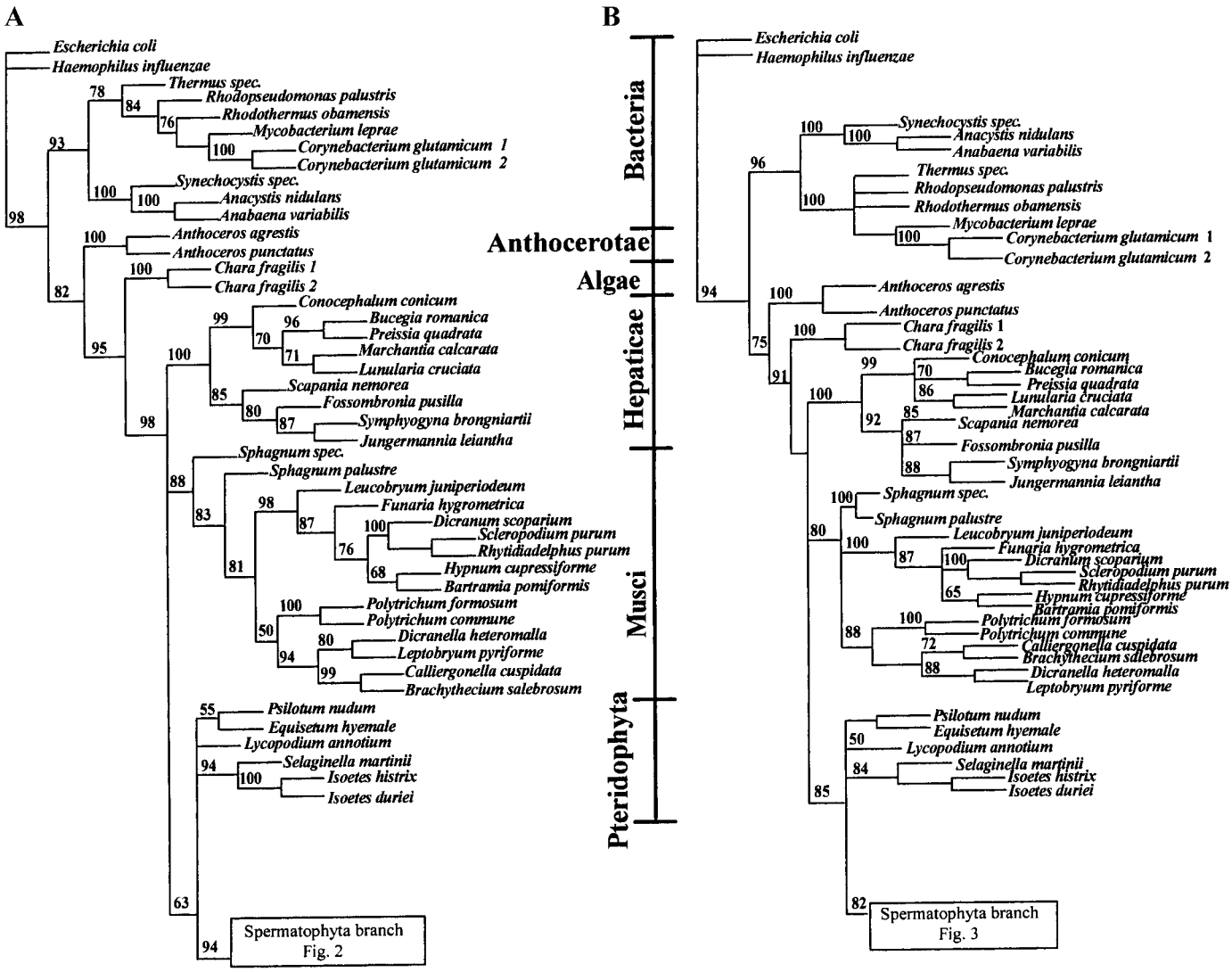

FIG. 1. (A) Neighbor-joining phylogenetic tree of 143 amino acid sequences based on an approx 1100-bp fragment (3′ side) of the PEPC gene. The tree comprises all PEPC sequences up to now known; among them, 70 were first analyzed in the present study. Numbers above the nodes indicate bootstrap values (values less than 50% are not shown). The box comprising the sequences of the Spermatophyta is shown in detail in Fig. 2. (B) Parsimony phylogenetic tree of the same PEPC amino acid sequences as shown in A (for details of the Spermatophyta box see Fig. 3).

worts and the mosses. On the other hand, studies of reproductive and structural innovations in the gametophyte and sporophyte generations of hornworts, liverworts, and mosses (Renzagali *et al.*, 2001) suggest that the hornworts represent the earliest divergent embryophyte clade, with the moss/liverwort clade as sister to tracheophytes. From our present results, hornworts can be regarded a basal group separated by high bootstrap values (100%). Our results are in good agreement with those of Waters *et al.* (1992), which show the Anthocerotae as a sister clade to the Musci. In addition, Anthocerotae show an amino acid composition of PEPC that is completely different from that of the remaining species investigated. Moreover, our results support the opinion by Sluiman (1985) that the hornworts represent an entirely independent derivation of land plants.

Our results fit with the view that the charophytes form a paraphyletic group relative to the land plants. In this context it would have been interesting also to compare other algae on the level of the PEPC partial sequence. However, with the PEPC primer pair used in the present studies, PCR amplification products were obtained only with *Chara* and not with other algae. Since our primers were derived from PEPC of higher plants this finding can be taken as a hint that the PEPC structure of *Chara* is related much closer to that of the land plants than to that of the algae. This is consistent with the phylogenetic position of the Charophyceae in relation to the land plants and algae as

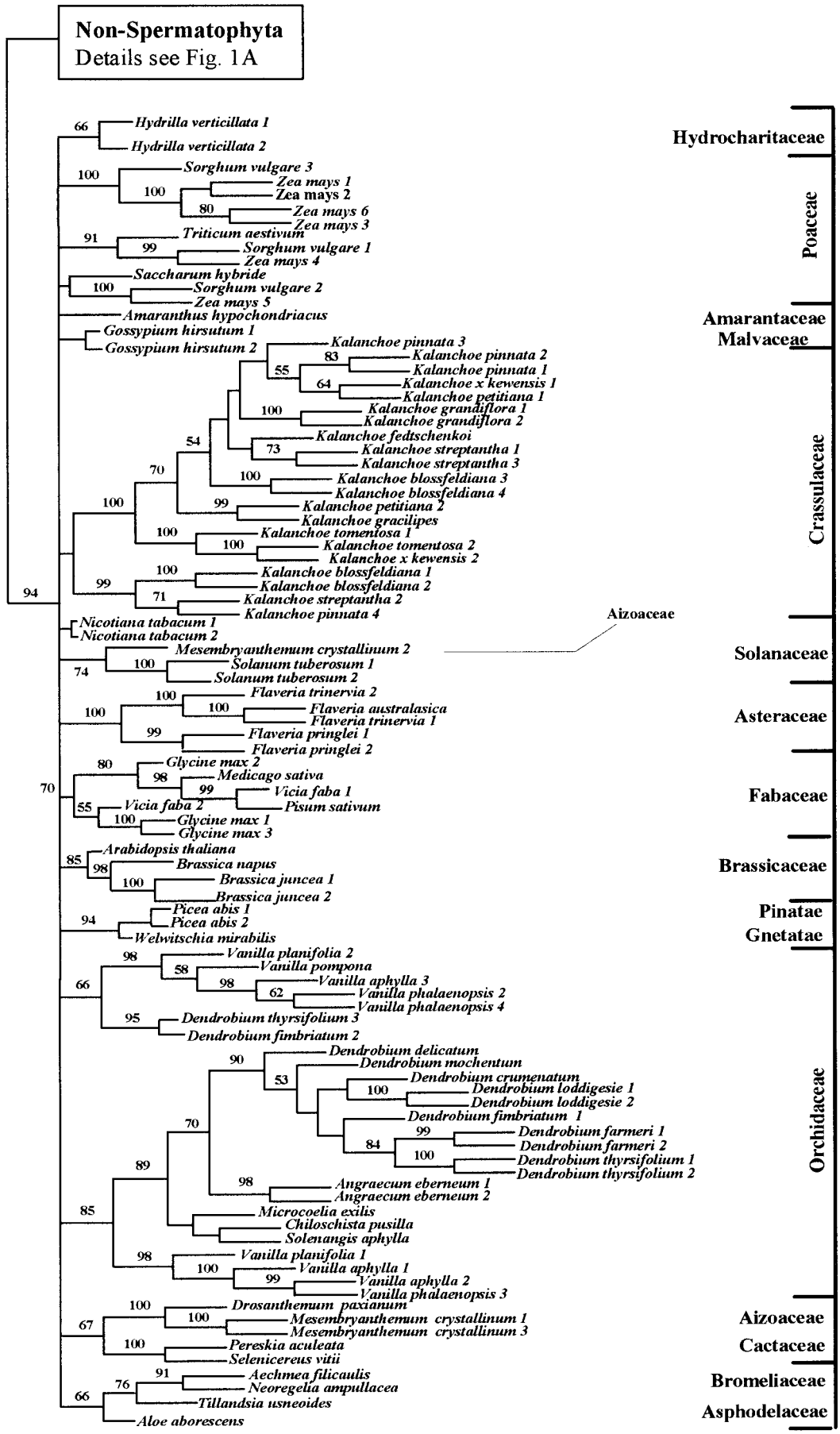

FIG. 2. Details of the Spermatophyta box of the neighbor-joining tree shown in Fig. 1A.

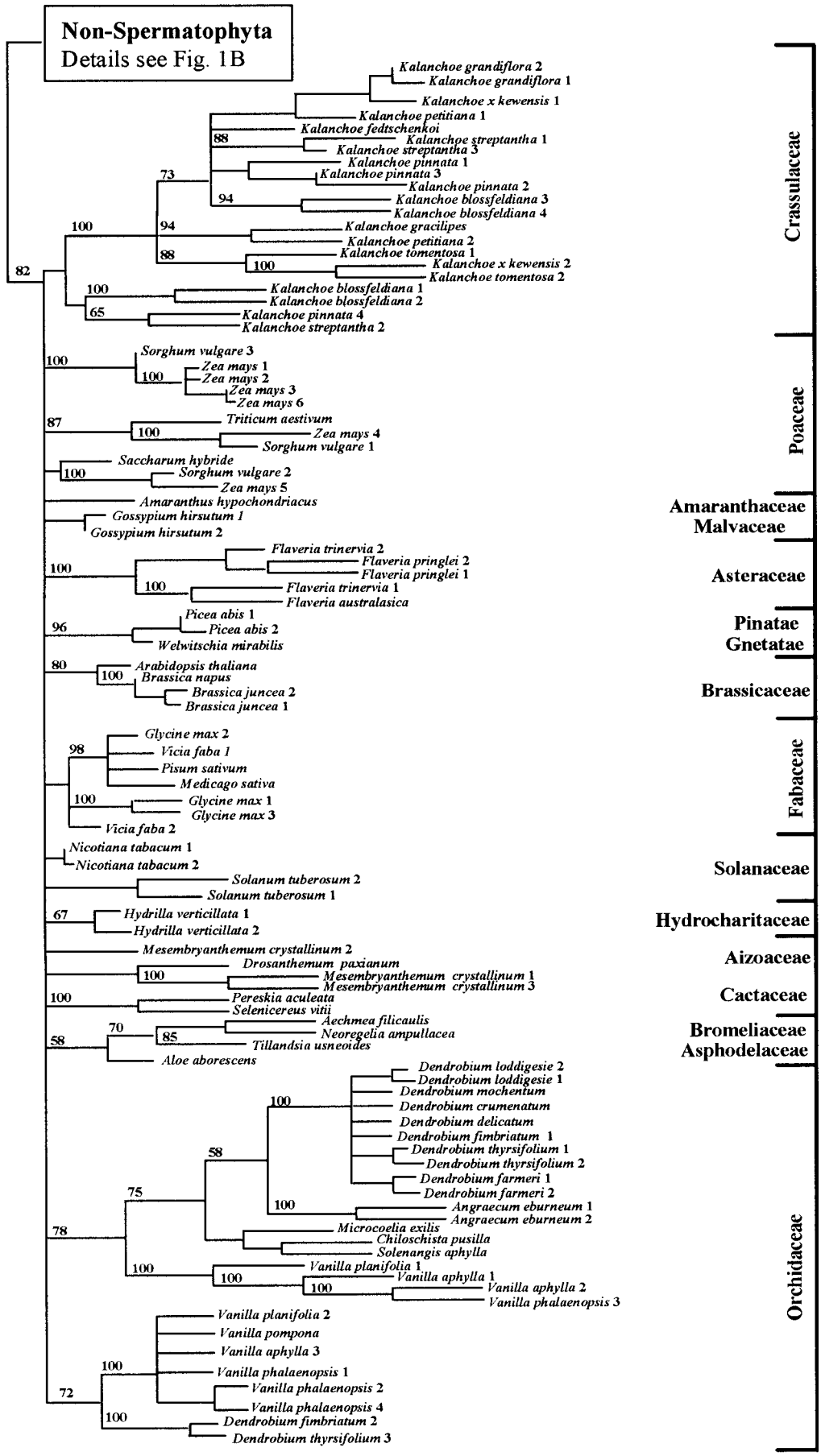

FIG. 3. Details of the Spermatophyta box of the parsimony tree shown in Fig. 1B.

proposed by Qui and Palmer (1999). To be able to analyze PEPC sequences in algae, we need to construct new primers.

Except for the hornworts, in the phylogenetic trees (Figs. 1A and 1B) the PEPC sequences of the other 24 bryophytes investigated so far form two clusters (bootstrap support 88 and 100%), with one branch comprising the Hepaticeae and the other the Musci. This suggests that these two groups of Bryophyta have evolved in parallel (Waters *et al.*, 1992; Mishler *et al.*, 1994; Capesius and Stech, 1997). It is worth mentioning that the amino acid sequences of the *Anthoceros* PEPC isoforms show a significantly lower homology to the other Hepaticeae and Musci (40 and 47%, respectively) than to the two PEPC isoforms of *Chara fragilis* (51 and 58%). This is further support of the view that the Anthocerotae have evolved separately from Hepaticae and Musci (Schuster, 1984).

The PEPC sequences of the 6 species of Pteridophyta investigated up to now form a sister group to the Spermatophyta, with three parallel polytomic branches (bootstrap values 49, 55, and 94%). The present data provide further evidence in favor of the view that, in contrast to former assumptions, the mosses but not the pteridophytes represent the first land plants. Independent of the mode of calculation, in the PEPC dendrograms shown in Figs. 1A and 1B, the Spermatophyta species form one large common cluster which is clearly separated from that of the archegoniates and prokaryotes (bootstrap support by 94 and 82%, respectively). The PEPC sequences of the Spermatophyta represent 53 plant species in 16 plant families (2 gymnosperms, 51 angiosperms, with 8 of the Dicotyledonae and 6 of the Monocotyledonae; Table 3). Within the large spermatophytean cluster the PEPC sequences form different branches (Figs. 2 and 3), although some of them have low bootstrap support. A particularly interesting finding concerns the PEPC sequences of the conifer *Picea abies* and the gnetophyt *Welwitschia mirabilis* (Fig. 2) showing that with high bootstrap support these two species cluster together. Classically the gnetophytes are considered the sister group of the angiosperms. However, there are now molecular data which imply that the gnetophytes are more closely related to the conifers than to the angiosperms (Winter *et al.*, 1999; Bowe *et al.*, 2000; Chaw *et al.*, 2000; Donoghue and Doyle, 2000). The PEPC tree shown in Fig. 2 provides further strong support in favor of the latter view.

Within the cluster representing the Spermatophyta the PEPC sequences not only arrange according to taxa but also within a taxon apparently arrange according to their assumed specific function. This phenomenon might reflect functional diversification during the generation of paralogous PEPC genes. However, a detailed analysis of the presumably complex relationships between paralogous and orthologous PEPC genes has to remain beyond the scope of the present treatise, because it would require much more information than at present is available on the existence of PEPC isoforms in the single plant species.

TABLE 3

Overview of the 143 PEPC Isoforms Considered in This Study and the Taxonomic Position of the Donor Organism

Taxonomic unit		Number of considered species	Numbers of PEPC sequences analyzed in the considered species
Prokaryota			
	α Subdivision	6	6
	γ Subdivision	2	2
	Cyanophyceae	3	3
Eukaryota			
	Charophyceae	1	2
Bryophyta	Anthocerotae	2	2
	Marchantiatae	9	9
	Bryatae	15	15
Pteridophyta	Psilotatae	1	1
	Lycopodiatae	4	4
	Equisetatae	1	1
Spermatophyta	Gymnospermae		
	Pinatae	1	2
	Gnetatae	1	1
	Angiospermae		
	Dicotyledoneae		
	Malvaceae	1	2
	Crassulaceae	9	21
	Solanaceae	2	4
	Aizoaceae	2	4
	Asteraceae	3	5
	Fabaceae	4	7
	Brassicacae	3	4
	Cactaceae	2	2
	Monocotyledoneae		
	Poaceae	4	11
	Bromeliaceae	3	3
	Orchidaceae	15	28
	Amarantaceae	1	1
	Asphodelaceae	1	1
	Hydrocharitaceae	1	2

The mentioned clustering of PEPC sequences according to the assumed specific function can be observed for instance in the genera *Flaveria* and *Kalanchoe* (data not shown in detail), but can be seen particularly clearly in the case of the Poaceae and Orchidaceae (Fig. 4). In the Poaceae the PEPC sequences are separated into three branches representing the functional isoforms ppc-C4, ppc-aL, and ppc-aR (denotation according to the nomenclature outlined in Table 2). Also in the Orchidaceae (28 PEPC sequences representing 15 species in 6 genera) there is evidence that the PEPC sequences branch according to their proposed functions (Fig. 4). In the orchids we found two major clusters, with the smaller cluster comprising all the isoforms

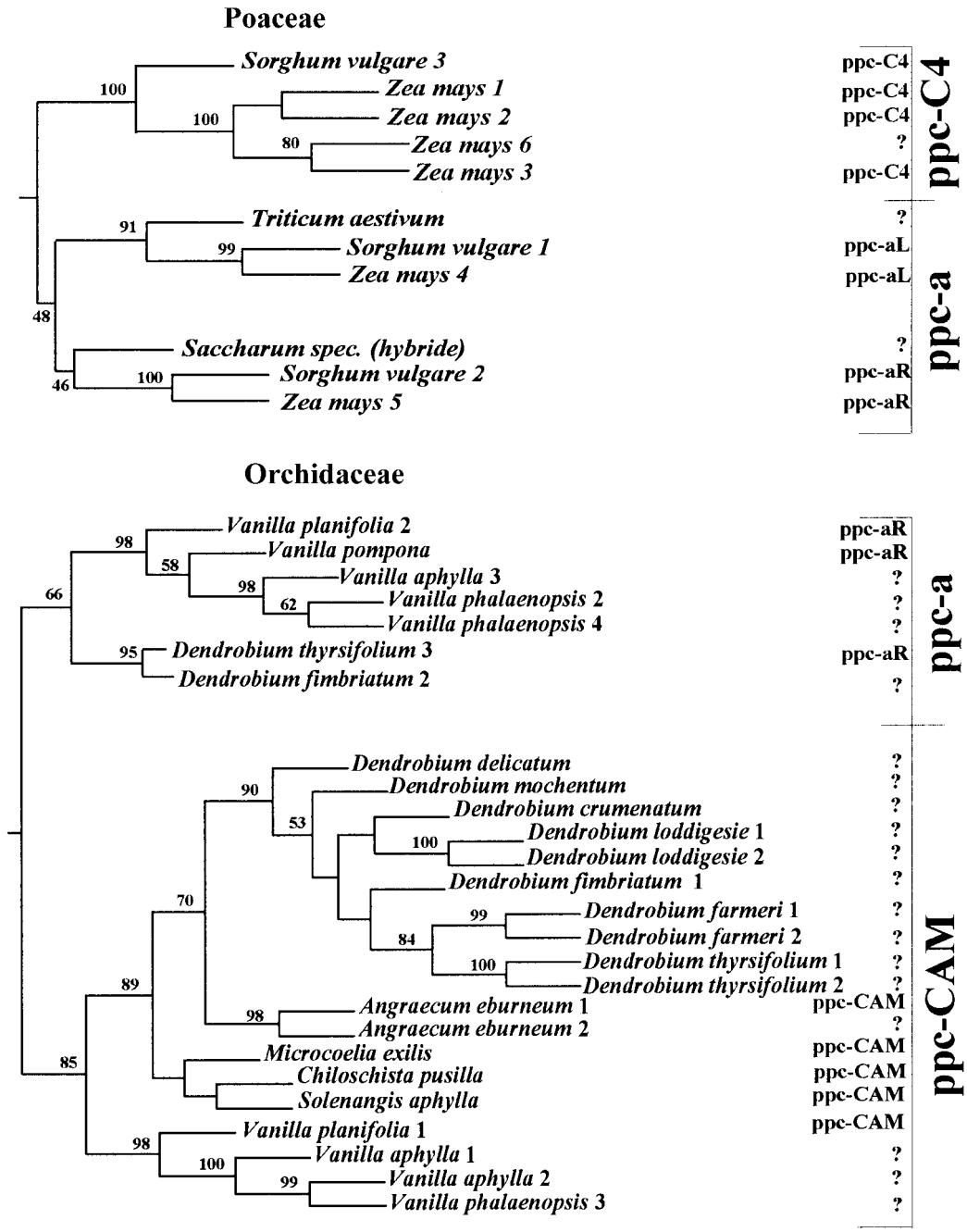

FIG. 4. Interrelationships between the position of PEPC sequences in the neighbor-joining tree (Fig. 2) and the likely specific functions of the concerned PEPC isoforms, exemplified for the Poaceae and Orchidaceae. In the Poaceae the attribution of the isoforms to a given function follows the suggestions published in the literature (see Table 1). For the Orchidaceae, attribution to CAM is based on the fact that the concerned isoform was the only or the mainly expressed isoform in plant material for which we have shown CAM performance when the material was extracted (data not shown). The denotation ppc-aR refers to the fact that the isoform was found in non-CAM-performing aerial roots. Question marks indicate unknown function of the concerned PEPC isoform.

that we have identified in nonphotosynthetic aerial roots and the larger cluster comprising isoforms expressed in photosynthetic organs. In those cases in which we have shown CAM performance when the plant tissue was extracted, we assumed that the found PEPC isoform was CAM related. For instance, in *Vanilla planifolia*, the "isoform 2" expressed in the aerial roots (ppc-aR) appears in a cluster other than that of the "isoform 1" cluster expressed in the CAM-performing leaves. Since in the case of *Dendrobium*, because of limitation in the plant material the mode of photosynthesis could not be investigated, in this genus the def-

inite attribution of PEPC isoforms to CAM remains open. However, because the *Dendrobium* species represent succulent-leaf epiphytes, and to our knowledge all species of that genus thereupon investigated are CAM plants, we believe that most of the concerned PEPC isoforms labeled in Fig. 4 by a question mark can be shown by future work to belong to the ppc-CAM type.

Among the investigated orchids there were three species (*Microcoelia exilis*, *Chilochista pusilla*, *Solenangis aphylla*) in which the photosynthetic organs consist of chloroplasts containing CAM-performing aerial roots (Winter, 1985), with leaves and shoot axes being largely reduced. As Fig. 4 shows, the PEPC isoforms of these "shootless" orchids do not cluster with the PEPC isoforms of the nonphotosynthetic aerial roots of the *Vanilla* and *Dendrobium* species but rather cluster with the PEPC isoforms found, in other species, in CAM-performing leaves. This interesting phenomenon suggests that the shootless orchids do not make use of the root-inherent ppc-aR isoform of PEPC to catalyze the initial β carboxylation in CAM. Rather, they obviously express for this function an additional isoform presumably specifically related to CAM.

The results of our present study support the previous findings by Gehrig *et al.* (1998a) that in the spermatophytes the PEPC isoforms assumed to be functionally related to CAM are widely dispersed over the different levels of taxa. This is in harmony with the view that the evolution of CAM is of polyphyletic origin (Cushman and Bohnert, 1997, 1999).

Altogether, the present study strengthens the view that PEPC sequences provide valuable molecular markers which may help to answer open questions in future phylogenetic studies of microorganisms and plants. They can also contribute to better knowledge of the evolution of metabolic pathways in which the PEPC is involved. Our study has also shown that, in the context of molecular phylogeny and taxonomy, it may be sufficient to compare suitable partial instead of full-length PEPC sequences. This helps to save time and financial resources, thus considerably increasing the value of PEPC nucleotide and amino acid sequences as widely applicable molecular markers.

ACKNOWLEDGMENTS

We thank Catharine von Glassow, Ute Reck, Kerstin Weitmann, and Fanka Rosic for practical assistance. The careful work by Monika Medina-España in preparing the manuscript is gratefully acknowledged. The investigations were supported by the Deutsche Forschungsgemeinschaft and by a grant from the FAZIT Foundation.

REFERENCES

Andreo, C. S., Gonzales, D. H., and Iglesias, A. A. (1987). Higher plant phosphoenolpyruvate carboxylase: Structure and regulation. *FEBS Lett.* **213:** 1–8.

Bogler, D., and Simpson, B. (1996). Phylogeny of Agavaceae based on IST rDNA sequence variation. *Oecologia* **83:** 1225–1235.

Bopp, M., and Capesius, I. (1996). New aspects of Bryophyte taxonomy provided by a molecular approach. *Bot. Acta* **109:** 368–372.

Bowe, L. M., Coat, G., and dePamphilis, C. W. (2000). Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophylethic and Gnetales closest relatives are conifers. *Proc. Natl. Acad. Sci. USA* **97:** 4092–4097.

Bruns, T., White, T., and Taylor, J. (1991). Fungal molecular systematics. *Annu. Rev. Ecol. Syst.* **22:** 525–564.

Capesius, I. (1995). A molecular phylogeny of bryophytes based on the nuclear encoded 18S rRNA genes. *J. Plant Physiol.* **146:** 59–63.

Capesius, I., and Stech, M. (1997). Molecular relationships within mosses based on 18S rRNA gene sequences. *Nova Hed.* **64:** 3–4.

Chaw, S., Parkinson, C. L., Cheng, Y., Vincent, T. M., and Palmer, J. D. (2000). Seed plant phylogeny from all three plant genomes: Monophyly of extant gymnospoerms and origin of Gnetales from conifers. *Proc. Natl. Acad. Sci. USA* **97:** 4086–4091.

Chirgwin, J. M., Przybyla, A. E., McDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribunucleases. *Biochemistry* **18:** 5294–5299.

Chollet, R., Vidal, J., and O'Leary, M. H. (1996). Phosphoenolpyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47:** 273–298.

Crétin, C., Keryrer, E., Tagu, D., Lepiniec, L., Vidal, J., and Gadal, P. (1990). Complete cDNA sequence of *Sorghum* phosphoenolpyruvate carboxylase involved in C4 photosynthesis. *Nucleic Acid Res.* **18:** 658.

Crétin, C., Santi, S., Keryrer, E., Lepiniec, L., Tagu, D., Vidal, J., and Gadal, P. (1991). The phosphoenolpyruvate carboxylase gene family in *Sorghum:* Promoter structure, amino acid sequences and expression of genes. *Gene* **99:** 87–94.

Cushman, J. C., and Bohnert, H. J. (1989a). Nucleotide sequence of the Ppc2 gene encoding a housekeeping isoform of PEPC from *Mesembryanthemum crystallinum*. *Nucleic Acid Res.* **17:** 6743–6744.

Cushman, J. C., and Bohnert, H. J. (1989b). Nucleotide sequence of the gene encoding a CAM specific isoform of PEPC from *Mesembryanthemum crystallinum*. *Nucleic Acid Res.* **17:** 6745–6746.

Cushman, J. C., and Bohnert, H. J. (1996). Transcriptional activation of CAM genes during development and environmental stress. *Ecol. Stud.* **114:** 135–158.

Cushman, J. C., and Bohnert, H. J. (1997). Molecular genetics of crassulacean acid metabolism. *Plant Physiol.* **113:** 667–676.

Cushman, J. C., and Bohnert, H. J. (1999). Crassulacean acid metabolism: Molecular genetics. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50:** 305–332.

Cushman, J. C., Meyer, G., Michalowski, C. B., Schmitt, J. M., and Bohnert, H. J. (1989). Salt stress leads to differential expression of two isogenes of phosphoenolpyruvate carboxylase during crassulacean acid metabolism induction in the common ice plant. *Plant Cell* **1:** 715–725.

Donoghue, M. J., and Doyle, J. A. (2000). Seed plant phylogeny: Demise of the anthophyte hypothesis? *Curr. Biol.* **10:** R106–R109.

Dressler, R. L., and Chase, M. W. (1995). Whence the orchids? *In* "Monocotyledons: Systematics and Evolution" (P. J. Rudall, D. F. Cutler, and C. J. F. Humphries, Eds.), pp. 217–226. Royal Botanic Gardens, Kew.

Eikmanns, B. J., Folletie, M. T., Griot, M. U., and Sinskey, A. J. (1989). The phosphoenolpyruvate carboxylase gene of *Corynebacterium glutamicum*. *Mol. Gen. Genet.* **218:** 330–339.

Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package) version 3.5c. Univ. of Washington.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., and Smith, H. O. (1995). Whole genome random sequencing and assembly of *Haemophilus influenzae* rd. *Science* **269:** 496–512.

Fujita, N., Miwa, T., Ishijima, K., and Katsuki, H. (1984). The primary structure of phosphoenolpyruvate carboxylase of *E. coli*. *J. Biochem.* **95:** 909–916.

Gehrig, H. H., Taybi, T., Kluge M., and Brulfert, J. (1995). Identification of multiple PEPC isogenes in leaves of the facultative crassulacean acid metabolism (CAM) plant *Kalanchoe blossfeldiana* Poelln. cv. Tom Thumb. *FEBS Lett.* **377:** 399–402.

Gehrig, H. H., Heute, V., and Kluge M. (1998a). Towards a better knowledge of the molecular evolution of phosphoenolpyruvate carboxylase by comparison of partial cDNA sequences. *J. Mol. Evol.* **46:** 107–114.

Gehrig, H. H., Faist, K., and Kluge, M. (1998b). Identification of phosphoenolpyruvate carboxylase isoforms in leaf, stem, and roots of the obligate CAM plant *Vanilla planifolia* Salib. (Orchidaceae): A physiological and molecular approach. *Plant Mol. Biol.* **38:** 1215–1223.

Gonzalez, M. C., Osuna, L., Echevarria, C., Vidal, J., and Cejudo, F. J. (1998). Expression and localization of phosphoenolpyruvate carboxylase in developing and germinating wheat grains. *Plant Physiol.* **116:** 1249–1258.

Graham, L. K. E., and Wilcox, L. W. (2001). The origin of alternation of generations in land plants: A focus on matrophy and hexose transporter. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **29:** 757–766.

Hata, S., Izui, K., and Kouchi, H. (1997). Expression of a soybean nodule-enhanced phosphoenolpyruvate carboxylase gene that shows striking similarity to another gene for a housekeeping isoform. *Plant J.* **7:** 198–201.

Henrik, A. H., Martin, T., and Sun, S. S. M. (1992). Structure and expression of a sugarcane gene encoding a housekeeping phosphoenolpyruvate carboxylase. *Plant Mol. Biol.* **20:** 663–671.

Hepperle, D. (1997). Alignment editor, version 3.5a. Heidelberg.

Hermans, J., and Westhoff, P. (1992). Homologous genes for the C4 isoform of phosphoenolpyruvate carboxylase in C3 and C4 *Flaveria* Species. *Mol. Gen. Genet.* **234:** 275–284.

Honda, H., Okamoto, T., and Shimada, H. (1996). Isolation of a cDNA for a phosphoenolpyruvate carboxylase from a monocot CAM-plant, *Aloe arborescens:* Structure and its gene expression. *Plant Cell Physiol.* **37:** 881–888.

Hudspeth, R. L., and Grula, J. W. (1989). Structure and expression of the maize gene encoding the phosphoenolpyruvate carboxylase isozyme involved in C4 photosynthesis. *Plant Mol. Biol.* **12:** 579–589.

Inui, M., Dumay, V., Zahn, K., Yamagata, H., and Yukawa, H. (1997). Structural and functional analysis of the phosphoenolpyruvate carboxylase gene from the purple nonsulfur bacterium *Rhodopseudomonas palustris* No. 7. *J. Bacteriol.* **179:** 4942–4945.

Izui, K., Ishijima, S., Yamaguchi, Y., Katagiri, F., Murata, T., Shigesada, K., Sugiyama, T., and Katsuki, H. (1986). Cloning and sequence analysis of cDNA encoding active PEPC of the C4 photosynthesis in maize. *Nuc. Acid Res.* **14:** 1615–1628.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. PCC6803: Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3:** 109–136.

Katagiri, F., Kodaki, T., Fuijta, N., Izui, K., and Katsuki, H. (1985). Nucleotide sequence of the phosphoenolpyruvate carboxylase gene of cyanobacterium *Anacystis nidulans. Gene* **38:** 265–269.

Kawamura, T., Shigesada, K., Toh, H., Okumura, S., Yanagisawa, S., and Izui, K. (1992). Molecular evolution of phosphoenolpyruvate carboxylase for C4 photosynthesis in maize: Comparison of its cDNA sequence with a newly isolated cDNA encoding an isozyme involved in the anaplerotic function. *J. Biochem.* **112:** 147–154.

Kenrick, P., and Crane, P. R. (1997a). The origin and early evolution of plants on land. *Nature* **389:** 33–39.

Kenrick, P., and Crane, P. R. (1997b). "The Origin and Early Diversification of Land Plants: A Cladistic Study." Smithsonian Institution Press, Washington, DC.

Kimura, M. (1983). "The Neutral Theory of Molecular Evolution." Cambridge Univ. Press, Cambridge, UK.

Kluge, M., and Ting, I. (1978). Crassulacean acid metabolism: Analysis of an ecological adaptation. *Ecol. Stud.* **30.**

Koizumi, N., Sato, F., Terano, Y., and Yamada, Y. (1991). Sequence analysis of cDNA encoding phosphoenolpyruvate carboxylase from cultured tobacco cells. *Plant Mol. Biol.* **17:** 535–539.

Latzko, E., and Kelly, G. J. (1983). The many-faceted function of phosphoenolpyruvate carboxylase in C3 plants. *Physiol. Veg.* **21:** 805–815.

Lepiniec, L., Santi, S., Keryer, E., Amiet, V., Vidal, J., Gadal, P., and Cretin, C. (1991). Complete nucleotide sequence of one member of the *sorghum* phosphoenolpyruvate carboxylase gene family. *Plant Mol. Biol.* **17:** 1077–1079.

Lepiniec, L., Keryer, E., Tagu, D., Gadal, P., and Cretin, C. (1992). Complete nucleotide sequence of a *sorghum* gene coding for the phosphoenolpyruvate carboxylase involved in C4 photosynthesis. *Plant Mol. Biol.* **19:** 339–342.

Lepiniec, L., Keryer, E., Philippe, H., Gadal, P., and Cretin, C. (1993). The phosphoenolpyruvate carboxylase gene family of *sorghum:* Structure, function and molecular evolution. *Plant Mol. Biol.* **21:** 487–502.

Lepiniec, L., Vidal, J., Chollet, R., Gadal, P., and Cretin, C. (1994). Phosphoenolpyruvate carboxylase: Structure, regulation and evolution. *Plant Sci.* **99:** 111–124.

Luinenburg, I., and Coleman, J. R. (1992). Identification, characterization and sequence analysis of the gene encoding phosphoenolpyruvate carboxylase in *Anabaena variabilis* PCC7120. *J. Gen. Microbiol.* **138:** 685–692.

Magnin, N., Reiskind, J. B., and Bowes, G. (1996). Identification of phosphoenolpyruvate carboxylase isoforms from an aquatic monocot with inducible C4-type photosynthesis. *Plant Physiol.* **8:** 72–72.

Melzer, E., and O'Leary, M. (1987). Anaplerotic fixation by phosphoenolpyruvate carboxylase in C3 plants. *Plant Physiol.* **84:** 58–60.

Merkelbach, S., Gehlen, J., Denecke, M., Hirsch, H. J., and Kreuztaler, F. (1993). Cloning, sequence analysis and expression of a cDNA encoding active phosphoenolpyruvate carboxylase of the C3 plant *Solanum tuberosum*. *Plant Mol. Biol.* **23:** 881–888.

Mishler, B. D., Lewis, L. A., Buchheim, M. A., Renzaglia, K. S., Garbary, D. J., Delwiche, C. F., Zechman, F. W., Kantz, T. S., and Chapman, R. L. (1994). Phylogenetic relationships of the green algae and bryophytes. *Ann. Missiouri Bot. Gard.* **81:** 451–483.

Nakamura, T., Yoshioka, I., Takahashi, M., and Izui, K. (1995). Cloning and sequence analysis of the gene for phosphoenolpyruvate carboxylase from an extreme thermophile, *Thermus* sp. *J. Biochem.* **118:** 319–324.

Pathirana, S. M., Vance, C. P., Miller, S. S., and Gantt, J. S. (1992). Alfalfa root nodule phosphoenolpyruvate carboxylase: Characterization of the cDNA and expression in effective and plant-controlled ineffective nodules. *Plant Mol. Biol.* **20:** 437–450.

Poetsch, W., Hermans, J., and Westhoff, P. (1991). Multiple cDNAs of phosphoenolpyruvate carboxylase in the C4 dicot *Flaveria trinervia*. *FEBS Lett.* **292:** 133–136.

Qiu, Y. L., and Palmer, J. D. (1999). Phylogeny of early land plants: Insights from genes and genomes. *Trends Plant Sci.* **4:** 26–30.

Rajagopalan, A. V., Tirumala, D. M., and Raghavendra, A. S. (1994). Molecular biology of C4 phosphoenolpyruvate carboxylase structure, regulation and genetic engineering (Review). *Photosynth. Res.* **39:** 115–135.

Relle, M., and Wild, A. (1996). Molecular characterization of a phosphoenolpyruvate carboxylase in the gymnosperm *Picea abies* (Norway spruce). *J. Plant Physiol.* **149:** 225–228.

Renzagali, K. S., Duff, R. J., Nickrent, D. L., and Garbary, D. J. (2001). Vegetative and reproductive innovations of early land plants: Implications for a unified phylogeny. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **29:** 769–793.

Rickers, J., Cushman, J., Michalowski, C., Schmitt, J., and Bohnert, H. J. (1989). Expression of the CAM-form of phosphoenolpyruvate carboxylase and nucleotide sequence of a full length cDNA from *Mesembryanthemum crystallinum*. *Mol. Gen. Genet.* **215:** 447–454.

Rydzik, E., and Berry, J. (1996). The C4 photosynthetic phosphoenolpyruvate carboxylase from grain amaranth. *Plant Physiol.* **110:** 713–715.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4:** 406–425.

Schuller, K. A., Turpin, T. H., and Plaxton, W. C. (1990). Metabolite regulation of partially purified soybean nodule phosphoenolpyruvate carboxylase. *Plant Physiol.* **94:** 1429–1435.

Schuster, R. M. (1984). *In* "New Manual of Bryology" (R. Schuster, Ed.), Vol. 2, p. 1071. Hattori Bot. Lab., Nichinan, Japan.

Sluiman, H. J. (1985). A cladistic evaluation of the lower and higher green plants. *Plant Syst. Evol.* **149:** 217–232.

Suganuma, N., Okada, Y., and Kanayama, Y. (1997). Isolation of a cDNA for nodule-enhanced phosphoenolpyruvate carboxylase from pea and its expression in effective and plant-determined ineffective pea nodules. *J. Exp. Bot.* **48:** 1165–1173.

Sugimoto, T., Kawasaki, T., Kato, T., Whittier, R. F., Shibata, D., and Kawamura, Y. (1992). cDNA sequence and expression of phosphoenolpyruvate carboxylase gene from soybean. *Plant Mol. Biol.* **20:** 743–747.

Svensson, P., Blaesing, O. E., and Westhoff, P. (1997). Evolution of the enzymatic characteristics of C4 phosphoenolpyruvate carboxylase: A comparison of the orthologous PPCA phosphoenolpyruvate carboxylases of *Flaveria trinervia* (C4) and *Flaveria pringlei* (C3). *Eur. J. Biochem.* **246:** 452–460.

Takai, K., Sako, Y., and Uchilda, A. (1998). The gene for phosphoenolpyruvate carboxylase from an extremely thermophilic bacterium, *Rhodothermus obamensis:* Cloning, sequencing and over expression in *Escherichia coli. Microbiology* **144:** 1423–1434.

Tello, A. V., Whittier, R. F., Kawasaki, T., Sugimoto, T., Kawamura, Y., and Shibata, D. (1993). Sequence of a Soybean (*Glycine max* L.) phosphoenolpyruvate carboxylase cDNA. *Plant Physiol.* **103:** 1025–1026.

Toh, H., Kawamura, T., and Izui, K. (1994). Molecular evolution of phosphoenolpyruvate carboxylase. *Plant Cell Environ.* **17:** 31–43.

Utter, M. F., and Kohlenbrander, H. M. (1972). Formation of oxalacetate by CO2-fixation on phosphoenolpyruvate. *In* "The Enzyme" (Boyer, Ed.), Vol. 4, pp. 117–170. Academic Press, New York.

Viret, J. F., and Lemoine, Y. (1989). Cloning and nucleotide sequence of the phosphoenolpyruvate carboxylase-coding gene of *Corynebacterium glutamicum* ATCC13032. *Gene* **77:** 237–251.

Vodjani, F., Kim, W., and Wilkins, T. A. (1997). Phosphoenolpyruvate carboxylase cDNA from developing cotton (*Gossypium hirsutum*) fibers. *Plant Physiol.* **115:** 315–315.

Waters, D. A., Buchheim, M. A., Dewery, R. A., and Chapman, R. L. (1992). Preliminary inferences of the phylogeny of bryophytes from nuclear-encoded ribosomal RNA sequences. *Am. J. Bot.* **79:** 459–466.

Winter, K. (1985). Crassulacean acid metabolism. *In* "Photosynthetic Mechanisms and Environment" (J. Barber, and N. R. Baker, Eds.), pp. 329–387. Elsevier, Amsterdam.

Winter, K. U., Becker, A., Muenster, T., Kim, J. T., Saedler, H., and Theissen, G. (1999). MADS-box genes reveal that gnetophytes are more closley related to conifers than to flowering plants. *Proc. Natl. Acad. Sci. USA* **96:** 7342–7347.

Yanai, Y., Okumura, S., and Shimada, H. (1994). Structure of *Brassica napus* phosphoenolpyruvate carboxylase genes: Missing introns causing polymorphisms among gene family members. *Biosci. Biotechnol. Biochem.* **58:** 950–953.

Yukawa, T., Ohba, H., Cameron, K. M., and Chase, M. W. (1996). Chloroplast DNA phylogeny of subtribe Dendrobiinae (Orchidaceae): Insights from combined analysis based on rbcL sequences and restriction site variation. *J. Plant Res.* **109:** 169–176.