Plant, Cell and Environment (2006) **29**, 2113–2123 doi: 10.1111/j.1365-3040.2006.01583.x

Phospho*enol*pyruvate carboxylase genes in C_3, crassulacean acid metabolism (CAM) and C_3/CAM intermediate species of the genus *Clusia*: rapid reversible C_3/CAM switches are based on the C_3 housekeeping gene

ANJA VAASEN[1], DOMINIK BEGEROW[2] & RÜDIGER HAMPP[1]

[1]*Physiological Ecology of Plants and* [2]*Plant Systematics/Mycology, Botanical Institute, University of Tübingen, Auf der Morgenstelle 1, 72076 Tübingen, Germany*

ABSTRACT

The genus *Clusia* includes species that exhibit either the C_3 or crassulacean acid metabolism (CAM) mode of photosynthesis, or those that are able to switch between both modes according to water availability. In order to screen for species-specific genetic variability, we investigated the key carboxylase for CAM, phospho*enol*pyruvate carboxylase (PEPC). Sequence analysis of DNA isolated from the obligate CAM species, *Clusia hilariana*, the obligate C_3 species, *Clusia multiflora*, and an intermediate species that can switch between C_3 and CAM photosynthesis, *Clusia minor*, revealed three different isoforms for *C. hilariana* and one each for the other two species. Sequence alignments indicated that PEPC from the intermediate species had high homology with the C_3 protein and with one of CAM plant proteins. These were assumed to constitute 'housekeeping' proteins, which can also support CAM in intermediate species. The other two isoforms of the CAM plant *C. hilariana* were either CAM-specific or showed homologies with PEPC from roots. Phylogenetic trees derived from neighbour-joining analysis of amino acid sequences from 13 different *Clusia* species resulted in two distinct groups of plants with either 'housekeeping' PEPC only, or additionally CAM-related isoforms. Only *C. hilariana* showed the third, probably root-specific isoform. The high homology of the PEPC from the intermediate species with the C_3 protein indicates that for the reversible transition from the C_3 to CAM mode of photosynthesis, the C_3 type of PEPC is sufficient. Its expression, however, is strongly increased under CAM-inducing conditions. The use of the C_3 isoform could have facilitated the evolution of CAM within the genus, which occurred independently for several times.

Key-words: *Clusia hilariana*; *Clusia minor*; *Clusia multiflora*; CAM erolution; CAM phylogeny; PEPC isoforms.

Correspondence: R. Hampp. E-mail: ruediger.hampp@uni-tuebingen.de

INTRODUCTION

Clusia L. is a genus with about 350 species (Willis 1973; Pipoly, Kearns & Berry 1998), which occur mainly in the tropical parts of South America, Madagascar and New Caledonia in a wide range of different life forms and habitats (Lüttge 1991). These shrubs and trees belong to the few dicotyledonous trees species that are able to perform crassulacean acid metabolism (CAM), an ecophysiological adaptation of plants to arid environments (Winter & Smith 1996). CAM plants reduce respiratory water loss by taking up CO_2 at night (stomates open, low leaf-to-air water vapour pressure deficit), intermediate storage in the vacuole in the form of malic acid and ordinary refixation of CO_2 through the Calvin cycle after decarboxylation of malic acid during the day (e.g. Grams, Herzog & Lüttge 1998). Most of the *Clusia* species studied so far are either obligate CAM species or C_3/CAM intermediates. The latter have the ability to adjust to increasing aridity by switching to the CAM mode of photosynthesis, but also back to the C_3 mode, when water availability is improved again. This reversibility of adaptation is unique to this genus. Only few species are obligate C_3 performers (Borland *et al*. 1992; Zotz & Winter 1993; Haag-Kerwer *et al*. 1996; Grams, Herzog & Lüttge (1998); Herzog *et al*. 1999).

Phospho*enol*pyruvate carboxylase (PEPC) is the key enzyme of CAM and C_4 photosynthesis. In CAM plants, it is the most studied enzyme. (Vazquez-Tello *et al*. 1993; Grams *et al*. 1998). In addition, PEPC is involved in a variety of metabolic processes in plants regardless of their type of photosynthesis. Non-photosynthesis-related PEPC isoforms are important for basic functions such as anaplerosis/nitrogen assimilation, or stomatal movement (Chollet, Vidal & O'Leary 1996; Ernst & Westhoff 1997). Because the different isoforms are regulated in different ways, the existence of different genes were postulated (see Lepiniec *et al*. 1993; Toh, Kawamura & Izui 1994). According to current knowledge, various PEPC isoforms as identified in C_4, CAM and C_3 plants are encoded by a small gene family (Lepiniec *et al*. 1993, 1994; Ernst & Westhoff 1997; Gehrig, Heute & Kluge 2001; Besnard *et al*. 2003; Engelmann *et al*. 2003).

In the C_4 plants *Zea mays*, *Sorghum vulgare* and *Flaveria trinervia* (Kawamura *et al*. 1990; Lepiniec *et al*. 1993; Ernst & Westhoff 1997; Tsuchida *et al*. 2001; Kai, Matsumura & Izui 2003), three genes exist, which were classified as 'housekeeping', root-inherent and light-inducible photosynthetic PEPC isoforms, respectively (Gehrig *et al*. 1995).

In the case of CAM photosynthesis, there are two plant genera mostly used for investigations, *Mesembryanthemum* and *Kalanchoë*. They include plants that perform the C_3 and CAM types of photosynthesis, as well as species that are able to switch from the C_3 to the CAM modes, but not back again (Gehrig *et al*. 1995; Borland & Griffiths 1997; Herppich & Herppich 1997). In *Mesembryanthemum crystallinum*, the CAM-related PEPC derives from a single member of a small gene family (Cushman *et al*. 1989; Ermolova *et al*. 2003).

CAM occurs among all major taxa of vascular plants. Its polyphyletic evolution was facilitated because there are no unique enzymes and metabolic reactions specifically required for CAM. A rearrangement and appropriately regulated complement of enzyme reactions present for basic functions in any green plant tissue are sufficient for performing CAM (Lüttge 2004). However, CAM-specific isoforms of key enzymes have evolved. Studies on the facultative CAM plants *M. crystallinum* and *Kalanchoë blossfeldiana* led to the conclusion that during the induction of CAM, in addition to the existing housekeeping isoform, a CAM-specific PEPC isoform is expressed, which is responsible for primary CO_2 fixation of this photosynthetic pathway (Cushman *et al*. 1989; Cushman & Bohnert 1996). Support for the existence of a CAM-specific PEPC isoform came from the comparison of transcripts obtained in the C_3 and the CAM state (Gehrig, Heute & Kluge 1998b).

As the genus *Clusia*, in contrast, covers species that are either C_3 or CAM plants, or are able to reversibly switch between the C_3 and CAM modes of photosynthesis, we wanted to know whether these properties are also related to the existence of CAM-specific PEPC genes. This also has implications with respect to the polyphyletic evolution of CAM among the vascular plants in general and within the genus *Clusia* in particular (Vaasen *et al*. 2002; Gehrig *et al*. 2003). We show that C_3/CAM intermediate types have a gene complement similar to C_3 but different from CAM plants.

MATERIAL AND METHODS

The studies were performed with the obligate C_3 species *Clusia multiflora*, the obligate CAM plant *Clusia hilariana* and the intermediate species *Clusia alata*, *Clusia aripoensis*, *Clusia articulata*, *Clusia criuva*, *Clusia major*, *Clusia minor*, *Clusia nemorosa*, *Clusia rosea*, *Clusia schomburgkiana*, *Clusia venosa* and *Clusia obovata* (synonymous: *Oedematopus obovatus*).

The different *Clusia* species were cultivated from cuttings obtained from the Botanical Garden, Department of Biology, Technical University of Darmstadt, where plants were grown for many years in greenhouses in pot culture. *C. minor* (TH Darmstadt) as well as *C. rosea*, *C. schomburgkiana* and *C. venosa* (University of Tübingen) were induced to change to the CAM mode of photosynthesis by not watering the plants for up to 3 weeks.

Extraction of DNA and RNA

For extraction of total genomic DNA and total RNA, 150–200 mg of leaf tissue were homogenized in liquid nitrogen. DNA was extracted according to the manufacturer's protocol (DNeasy Plant mini kit, Qiagen GmbH, Hilden, Germany) and eluted in a 200 μL elution buffer. For the extraction of RNA, we used the TRIZOL reagent (TRIR; Invitrogen BV, Groningen, the Netherlands) method according to Gehrig *et al*. (2000). RNA was eluted in 40 μL of diethyl pyrocarbonate (DEPC)-treated water.

PCR

Degenerated oligonucleotide primers were designed for one conserved region near the 3′ end of the gene among previously cloned PEPCs of CAM plants (Gehrig *et al*. 1995). The fragment obtained with the primers TCG GAY TCG GGG AAR GAY GC (forward) and GCR GCR ATR CCY TTC ATG GT (reverse) was about 1100 bp in size. For other parts of the gene, primers were designed from known PEPC sequences. The respective primer sequences are shown in Table 1.

PCRs were performed in 25 μL reaction mixtures, containing 2.5 μL of 10 × Qiagen PCR buffer (Qiagen, Hilden, Germany), 0.5 μL of dNTP mix (10 mM each of dATP,

Table 1. Primer name, sequence and sequence region of specific amplification used for PCR-based cloning of two cDNA fragments coding for phospho*enol*pyruvate carboxylase (PEPC) and for RACE-PCR reactions

Primer name	5′→ 3′ sequence	Region of amplification
PEPCfor1	TCG GAY TCG GGG AAR GAY GC	3′ forward primer
PEPCrev1	GCR GCR ATR CCY TTC ATG GT	3′ reverse primer
PEPCfor2	CDG TDG AYY TDG TYY TDA CTG CTC ATC C	Central fragment region Forward primer
PEPCrev2	GTA CAA CTG CCA AGC TGC AGA AAG TCG TCC	Central fragment region Reverse primer
5′RACE-PEPC1	GCA ACG CCA CAT GGA TAA TTC A	5′ race primer
3′RACE-PEPC1	CCA GGA TTG GAG GAC ACC CTC ATC TTG ACC	3′ race primer

dCTP, dGTP, dTTP), 0.15 μL of Taq DNA polymerase (0.75 units; Taq PCR Core Kit, Qiagen), 0.5 μL of each primer (10 pmol μL^{-1}) and 0.5 μL of genomic DNA (~ 100 ng).

Reverse transcriptase (RT)-PCR reactions were also carried out in 25 μL reaction mixtures, containing 5 μL of 5 × Qiagen OneStep RT-PCR buffer (Qiagen), 1 μL of dNTP mix, 1 μL of (Qiagen) OneStep RT-PCR Enzyme Mix (Qiagen OneStep RT-PCR Kit, Qiagen), 0.5 μL of each primer as used earlier and 0.5 μL of RNA.

PCR was performed using a Peltier Thermal Cycler (Biozym Diagnostics, Hess. Oldendorf, Germany) under the following conditions: initial denaturation at 94 °C for 5 min, followed by 35 cycles of template denaturation at 94 °C for 30 s, primer annealling at 57 °C for 1 min, primer extension at 72 °C for 1 min and a final extension at 72 °C for 5 min. RT-PCR was initiated by a reverse transcription step at 50 °C for 30 min and an initial PCR activation step of 95 °C for 15 min.

Rapid amplification of cDNA-ends (RACE)-PCR was achieved with SMART RACE cDNA Amplification Kit (BD Biosciences, Palo Alto, CA, USA) using 0.5–2.0 μL of RNA following the manufacturer's protocol.

Primer design

Several PEPC-specific, degenerated primers were designed based on the highly conserved regions of amino acid sequences of published PEPC isoforms in order to amplify the coding region of PEPC (Table 1). This resulted in the amplification of different gene fragments (from the 5′ end towards the 3′ end) with sizes of about 1350, 1090, 1000 and 400 bp. The RACE-PCR technique was used to obtain the 5′ and the 3′ ends of the sequences.

Purification and transformation of PCR products

Products of PCR amplifications were purified with the MinElute PCR Purification Kit (Qiagen). The purified fragments (3.5 μL) were ligated into the pCR2.1-TOPO vector (Invitrogen) and used to transform TOP 10 cells (Invitrogen). Successful transformation was tested by PCR with standard M13 primers with concentrations and conditions as mentioned earlier, and with an initial step of 98 °C for 2 min to denature the bacterial cells.

Sequence analysis

About fifteen purified clones of each *Clusia* species were sequenced by cycle sequencing (modified after Sanger, Nicklen & Coulson 1977) with M13 primers and the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, Wassington, UK) according to the manufacturer's protocol.

The complete gene sequence was established according to about 200-bp-long overlapping parts of the fragments, which showed enough differences to assign the sequences to the different isoforms.

Table 2. Accession numbers

Species	Accession no.
Clusia hilariana isoform 1	DQ320102
C. hilariana isoform 2	DQ320103
C. hilariana isoform 3	DQ320104
Clusia minor	DQ320105
Clusia multiflora	DQ320106
Clusia articulate	DQ320107
Clusia criuva	DQ320108
Clusia major	DQ320109
Oedematopus obovatus Synonym: *Clusia obovata*	DQ320110
Clusia schomburgkiana isoform 1	DQ320111
C. schomburgkiana isoform 2	DQ320112
C. schomburgkiana isoform 3	DQ320113
Clusia nemorosa	DQ320114
Clusia aripoensis	DQ320115
Clusia alata	DQ320116
Clusia venosa isoform 2	DQ320117
C. venosa isoform 1	DQ320118
Clusia rosea isoform 2	DQ320119
C. rosea isoform 1	DQ320120

For phylogenetic reconstruction, we considered either complete cDNA sequences or only the 1100 coding nucleotides in the 3′ terminal part. The sequences were proofread with the help of Sequencher 4.1 software (Gene Codes Corp., Ann Arbor, MI, USA). The DNA sequences determined for this study were deposited in GenBank. Accession numbers are given in Table 2. The sequences were aligned with ClustalX software (version 1.8; National Center of Biotechnology Information (NCBI), Bethseda, MD, USA; Thompson *et al*. 1997). Neighbour-joining analysis of protein sequences was carried out using phylogenetic analysis using parsimony (PAUP*) (Swofford 1999) with standard settings for amino acids.

RESULTS

The aim of the study was to test if C_3, CAM and intermediate *Clusia* species possess different members of the PEPC gene family and how many do exist. By comparing sequences, we wanted to distinguish possible housekeeping genes from those related to specific modes of photosynthesis.

Leaf samples of *C. minor*, which were taken during the C_3 and CAM modes of photosynthesis, and of *C. multiflora* (obligate C_3 plant), showed only one form of the PEPC gene.

In *C. hilariana*, we could identify three different isoforms. In contrast to the analysis of DNA, only two of the isoforms were detected by RT-PCR from mRNA preparations, which indicate that the respective gene is not expressed in leaves. The size of the full-length PEPC gene is about 3400 bp, with an open reading frame of 2892 bp and a 3′ untranslated part (data not shown). The respective protein comprises 962 amino acids (Fig. 1). In *C. hilariana*, also three different isoforms were clearly distinguishable, whereas in *C. minor* and *C. multiflora*, only one form occurs. We denoted the proteins as pepc-*Chi1*, pepc-*Chi2*

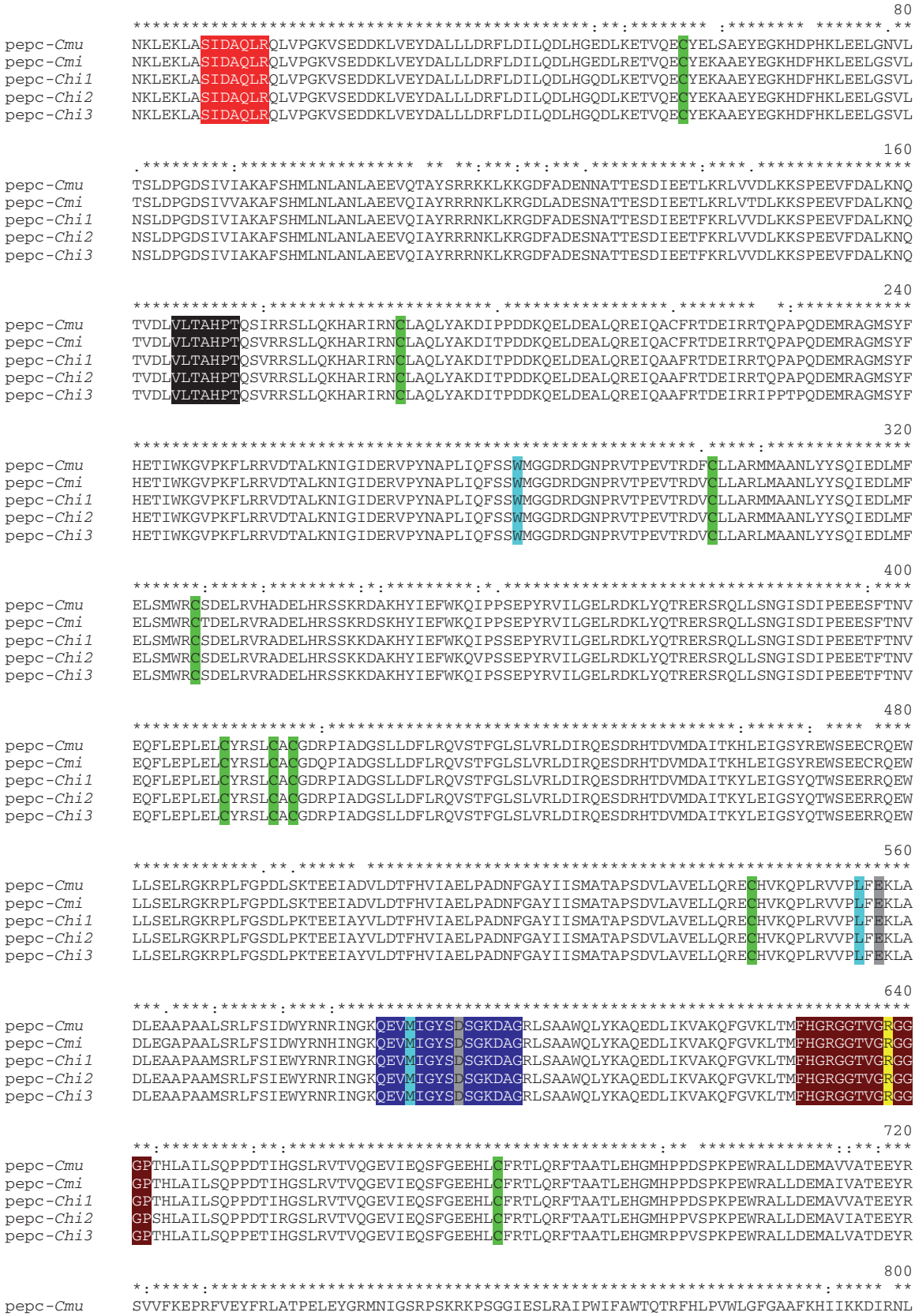

Figure 1. Alignment of the amino acid sequence of the PEPC proteins of *C. multiflora* (pepc-*Cmu*, C3), *C. minor* (pepc-*Cmi*, C3/CAM) und *C. hilariana* (pepc-*Chi*, CAM). For the alignment ClustalX was used. Identical amino acids are marked with stars (*) above the lines, colons indicate different amino acids with strongly similar characteristics, dots indicate different amino acids with a weak similarity. Highly conserved regions in all known PEPC's in bacteria and plants are coloured. Shown is colour are also regions with known function such as a phosphorylation site (pos. 8–14, bright red), the carboxylation site (pos. 165–171, black), the active centre (pos. 586–600, dark blue), the substrate binding site (pos. 629–642, dark red), the PED binding site (pos. 280, 555, 589, bright blue), the Mg^{2+} binding site (pos. 557, 594, grey), the aspartate binding site (pos. 638, 826, 885, 960, yellow), and conserved cysteine residues (Pos. 57, 188, 300, 327, 410, 415, 417, 544, 678, bright green) (after Besnard *et al*. 2003; Engelmann *et al*. 2003 und Lepiniec *et al*. 1994).

and pepc-*Chi3* (isoforms of *C. hilariana*), pepc-*Cmu* (*C. multiflora*) and pepc-*Cmi* (*C. minor*).

An alignment of the translated amino acid sequences (Fig. 1) shows that 91.3% of the amino acid positions in the consensus sequence of all species are identical. The sequences of *C. minor* and *C. multiflora* are mostly identical (97.5%) and also have a high identity with pepc-*Chi1* of *C. hilariana* (96.9 and 96.7%, respectively). In contrast, the other two isoforms of *C. hilariana* differ considerably more from the other two species (sequence homologies 95.3 and 94.1% with *C. minor*; 95.1 and 93.9% with *C. multiflora*). Within the isoforms of *C. hilariana*, pepc-*Chi1* has a higher homology to pepc-*Chi2* (98.0%) than to pepc-*Chi3* (96.8%); similarly, pepc-*Chi2* and pepc-*Chi3* are more distant from each other (96.5%). We thus assume that *pepc-Chi1* is the anaplerotic enzyme, present in all plants, whereas *pepc-Chi2* is related to CAM photosynthesis in *C. hilariana*. The third isoform, which was only identified in genomic DNA, is obviously not expressed in the leaves and could fulfil anaplerotic functions in the roots (Latzko & Kelly 1983).

An inter-species comparison of the three PEPC amino acid sequences (Fig. 1) indicated that 17 amino acids are different in the *C. hilariana* sequence (nos. 48, 81, 139, 216, 344, 358, 396, 463, 470, 471, 476, 494, 497, 504, 569, 576 and 855), 11 in the *C. minor* sequence (nos. 91, 123, 144, 328, 346, 420, 564, 581, 713, 832 and 940) and 13 in the *C. multiflora* sequence (nos. 60, 61, 70, 78, 110, 113, 116, 120, 127, 174, 198, 299 and 334).

A comparison of the isoforms of *C. hilariana* shows that the assumed 'housekeeping' isoform, pepc-*Chi1*, differs in two additional amino acids from the other two *C. hilariana* isoforms, as well as from the other two species (nos. 850 and 918). In contrast, pepc-*Chi1* amino acids in positions 698, 722, 861, 876 and 912 differ from the other two isoforms of *C. hilariana* (pepc-*Chi2* and pepc-*Chi3*), while they are identical to the other two species. Furthermore, pepc-*Chi2* has 11 positions (nos. 356, 643, 656, 714, 728, 807, 820, 846, 860, 886 and 897) and pepc-*Chi3* has 23 positions (nos. 225, 226, 228, 305, 653, 695, 713, 718, 792, 798, 801, 802, 805, 809, 843, 846, 869, 877, 909, 911, 924, 928 and 936), which contain amino acids different from pepc-*Chi1*.

We exposed three of the additionally studied species, *C. rosea*, *C. venosa* and *C. schomburgkiana*, to drought stress in order to get more information about *pepc* gene expression during the transition from C_3 and CAM photosynthesis. Physiologically, the transition was followed by leaf gas exchange (i.e. onset of nocturnal CO_2 fixation, data not shown).

For RNA extraction, we took samples at the beginning of the experiment in a well-watered stage and after 3 weeks of drying. We analysed cDNA fragments of 727 bp at the 3′ end of the *ppc* gene and compared them with the respective sequences of *C. hilariana*, *C. minor*, *C. multiflora* and the other seven species, *C. alata*, *C. aripoensis*, *C. articulata*, *C. criuva*, *C. major*, *C. nemorosa* and *C. obovata.*

An alignment of the sequences showed an identity of 78.8% of the nucleotides with 66 variations in only one sequence and 88 modifications in more than one sequence. The three species, *C. rosea*, *C. venosa* and *C. schomburgkiana*, developed a second isoform of PEPC under water shortage. These second isoforms, as well as the second isoform of *C. hilariana*, showed 24 analogous differences to the first isoforms.

The different isoforms can be identified even in the transcribed amino acid sequences. An alignment (Fig. 2) identifies seven amino acids that separate the two groups of isoforms from each other. In general, 79% of the amino acids of the C-terminal part of the protein are identical, 44% of those differing have a strong similarity and 17% have a weak similarity with regard to their biochemical properties.

The PEPC sequence of the C_4 isoform of *Saccharum officinarum* was taken as out-group (Accession no. AJ 293346; Besnard *et al*. 2002) because of a clear difference to the C_3 and CAM isoforms.

We then compared the sequences with the 'neighbour joining' method using the program PAUP (Swofford 1999) with standard settings for amino acids (Fig. 3).

This analysis resulted in two major clusters, formed by two of the PEPC isoforms. A third isoform (*C. hilariana*) is located at the bottom of the phylogenetic tree and is not expressed in leaves.

DISCUSSION

PEPC isoforms

Phospho*enol*pyruvate carboxylase has been characterized at the genomic level extensively in C_3 and C_4 plants, but much less is known about this enzyme in CAM plants (Gehrig *et al.* 1998b).

In this context, it was the aim of our investigation to characterize isoforms of PEPC that are specific for crassulacean acid metabolism as well as to screen for the presence of members of the small PEPC gene family in the genus *Clusia*, which consists of species being either obligate C_3, obligate CAM or able to reversibly switch between the two modes of photosynthesis.

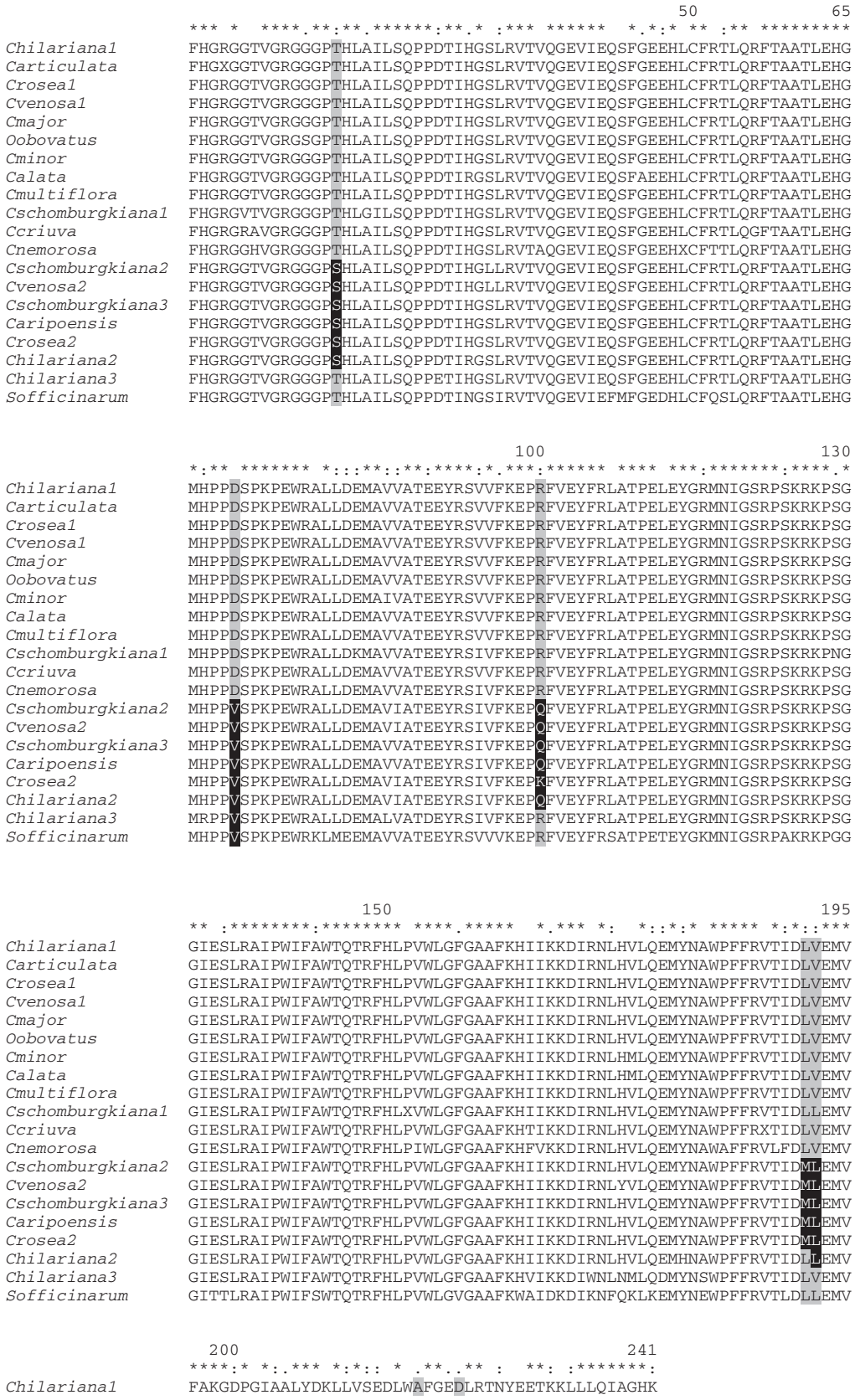

Figure 2. Alignment of a C-terminal part of the PEPC protein with 241 amino acids of 13 *Clusia* species and of *Saccharum officinarum* as outgroup (ClustalW). Different isoforms of PEPC's from *C. hilariana*, *C. rosea*, *C. schomburgkiana* and *C. venosa* are marked with numbers behind the species name. Identical amino acids are identified with stars (*) above the lines, colons indicate different amino acids with strongly similar characteristics, dots indicate different amino acids with a weak similarity. Differences between the assumed C_3 form (grey) and CAM form (black) are coloured.

Sequence analysis of *C. hilariana* (obligate CAM species) yielded three different isoforms, indicating the presence of three genes. In contrast, only one isoform was detectable in *C. minor* (intermediate species) and *C. multiflora* (obligate C_3 species), which both had a high sequence identity with one of the *C. hilariana* genes. From this observation, we conclude that the latter isoforms, which are highly homologous to leaf-specific (anaplerotic) ones, are probably represented by a single-copy gene, as shown for *Z. mays* (Harpster & Taylor 1986). The three different isoforms of *C. hilariana* and four different isoforms from NCBI of *C. uvitana* (Accession nos. AJ312636–39) are in support of the existence of a multigene family.

C. minor, the intermediate *Clusia* species, displayed only one isoform even after switching to CAM, which is more related to C_3-type PEPCs. These isoforms also possess a phosphorylation site at the N-terminal end of the protein and thus, can also be regulated by phosphorylation and dephosphorylation (Hermans & Westhoff 1992). Accordingly, we assume that after switching from C_3 to CAM, the housekeeping form of PEPC could become the key carboxylase for CAM photosynthesis at least shortly after starting with CAM until the second isoform is expressed. This could be especially important for plants such as *C. minor*, which can switch modes of photosynthesis very rapidly (Grams & Thiel 2002). Such a dual function could have supported the independent evolution of CAM in many species within the genus *Clusia* (Vaasen *et al*. 2002).

Similar analyses with C_4 plants such as *Z. mays*, *S. vulgare*, *Sorghum bicolor* and *F. trinervia* also identified three genes, which were classified as 'housekeeping', root-inherent and light-inducible photosynthetic PEPC isoforms

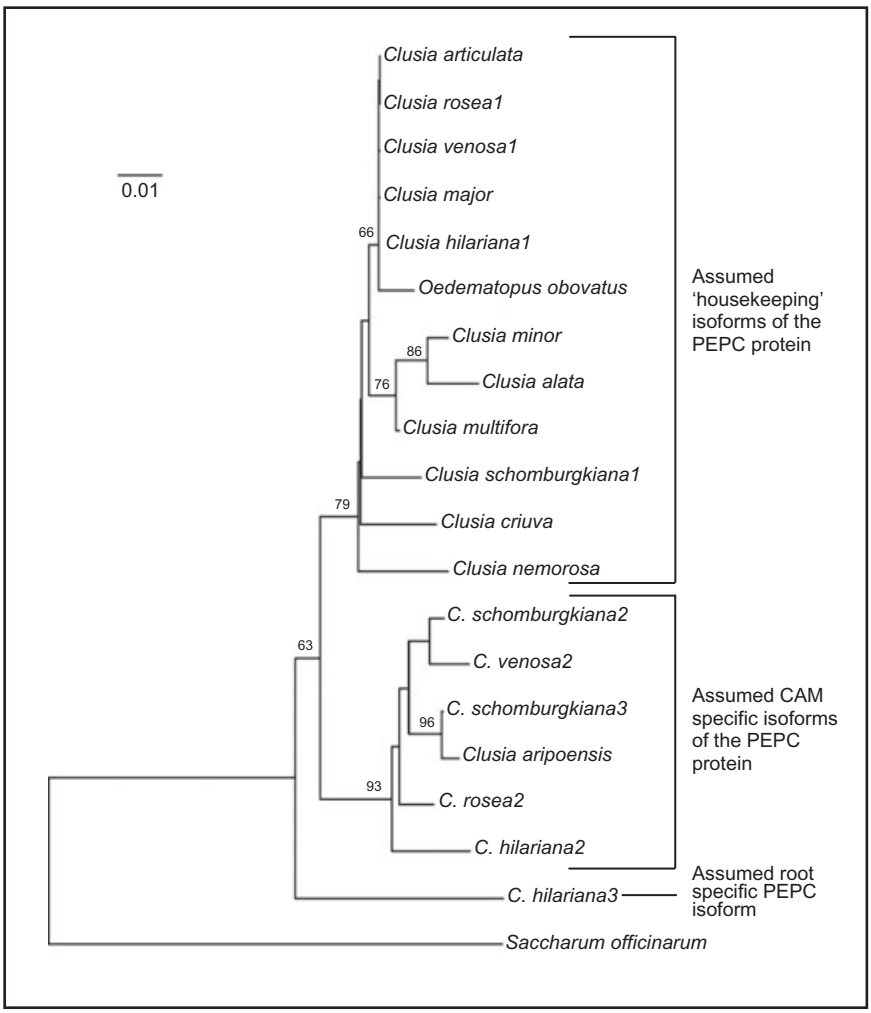

Figure 3. Phylogenetic hypothesis derived from neighbor joining analysis of a fragment of 241 amino acids at the C-terminal end of the PEPC protein derived from 13 different *Clusia* species, and rooted with *Saccharum officinarum*. Different isoforms of *C. hilariana*, *C. rosea*, *C. schomburgkiana* and *C. venosa* are identified with numbers behind the species name. Bar = 0.01 expected changes per site.

(Crétin *et al*. 1991; Lepiniec *et al*. 1993; Ernst & Westhoff 1997; Tsuchida *et al*. 2001). Furthermore, the use of housekeeping PEPCs for transition to C_4 has been suggested for the genus *Alternanthera* (Gowik *et al*. 2006). Molecular analysis showed that *Alternanthera pungens* (C_4) exhibited a typical C_4 PEPC isoform, while the PEPCs from *Alternanthera sessilis* (C_3) and *Alternanthera tenella* (C_3/C_4 intermediate) were found to be typical C_3 PEPC isozymes.

Comparing the *Clusia* sequences with other PEPCs of CAM plants such as *Mesembryanthemum* (Cushman *et al*. 1989; Slocombe, Whitelam & Cockburn 1993), *Kalanchoë* (Gehrig *et al*. 1995, 2005), *Aloe* (Honda, Okamoto & Shimada 1996), or *Vanilla* (Gehrig *et al.* 1998b), we screened for amino acids that distinguish CAM PEPC isoforms from other PEPC isoforms. In contrast to C_4-specific amino acid sequences (Bläsing, Westhoff & Svensson 2000; Besnard *et al*. 2002, 2003; Engelmann *et al*. 2003), we could not identify any conserved amino acid motifs in order to differentiate the housekeeping and root isoforms. While C_4 isoforms can be separated from non-C_4 isoforms (Honda *et al*. 1996), the isoforms in CAM plants group only within one genus and not within several genera (Honda *et al*. 1996; Gehrig *et al*. 1998b, 2001, 2005).

The 13 *Clusia* species investigated here form two main clusters (Fig. 3). *C. hilariana* is an obligate CAM species, *C. multiflora* is an obligate C_3 species, while all the other species are intermediate species, which can switch between the two modes of photosynthesis. The first cluster comprises the first isoforms of *C. hilariana*, *C. rosea*, *C. venosa*, *C. schomburgkiana* and the single isoforms of *C. minor*, *C. multiflora*, *C. articulata*, *C. major*, *C. alata*, *C. obovata*, *C. nemorosa* and *C. criuva*. The second group includes the second isoforms of *C. hilariana*, *C. rosea*, *C. venosa*, *C. schomburgkiana* and the single isoform of *C. aripoensis.* Because of this grouping, the first isoforms seem to be the 'housekeeping' isoforms, while the second isoforms are probably related to CAM photosynthesis. The presence of a third isoform in *C. schomburgkiana* could be attributed to duplication of the ancestral gene as it was also suggested for *K. blossfeldiana* and *Kalanchoë pinnata* (Gehrig *et al*. 1995, 2005). The third isoform of *C. hilariana* does not group and is located at the bottom of the tree. We thus assume that this isoform is root-specific, as also shown for *Sorghum*, *Zea* and *Vanilla* before (Gehrig *et al*. 1995; Dong *et al*. 1998; Besnard *et al*. 2003).

Each of the second isoforms has obvious differences to the housekeeping isoforms and can be clearly distinguished. It is thus very likely that the intermediate species will all express such a second isoform after switching to CAM for a longer period.

Our assumptions in regard to the functions of the different members of the *Clusia* PEPC gene family are supported by data on *K. pinnata* (Gehrig *et al.* 2005). These authors recovered seven distinct PEPC isogenes (four in leaves, three in roots). Sequence similarity comparison together with distance neighbourhood-joining calculations separated these isogenes in two clades, one formed by the root (C_3 isoforms) and the other formed by the leaf isogenes. The latter could be further divided into two branches, one of these containing typical CAM isoforms. Only one of these CAM isoforms was abundantly expressed under CAM conditions.

Characteristics of full-length *pepc* gene sequences

Comparing the consensus sequences (part of total gene) of all species analysed revealed identical positions for 91.3% of the amino acids. In addition, the nucleotide sequences in the coding region are highly homologous to those of the known *pepc* sequences, whereas a lower level of homology was found in the 3′ non-coding region, which comprises 300–400 bp. Amino acid sequences deduced from the nucleotide sequence of *Clusia* were highly conserved as compared to those from *Arabidopsis thaliana* (75.3%; Kaneko *et al*. 2000), *M. crystallinum* (75.2%; Cushman *et al*. 1989; Cushman & Bohnert 1989; Rickers *et al*. 1989), *F. trinervia* (79.4%; Poetsch, Hermans & Westhoff 1991; Hermans & Westhoff 1992), *Solanum tuberosum* (85.1%; Merkelbach *et al*. 1993), *Vanilla planifolia* (72.9%; Gehrig, Faist & Kluge 1998a), *S. bicolor* (80.8%; Lepiniec *et al*. 1992) and *Z. mays* (70.7%; Hudspeth & Grula 1989; Matsuoka & Minami 1989; Kawamura *et al*. 1992).

The sequence alignments given in Fig. 1 show that the first 600 amino acids exhibit species-specific sequence differences, while isoform-specific variations occur primarily within the last 360 amino acids. It is known that the N-terminal parts of the enzymes are responsible for their respective kinetic properties (Svensson, Bläsing & Westhoff 1997). It is thus not surprising that this part of the protein shows less difference between the isoforms.

The four conserved amino acid motifs of PEPCs and the several cysteine residues that are characteristic for PEPCs from all plants are present (Fig. 1): (1) SIDAQLR (nos. 8–14) is found only in plant proteins and the serine at position 8 is responsible for the phosphorylation in CAM and C_4 PEPC; (2) VxTAHPT (nos. 165–171) contains a histidine residue, which is important for the carboxylating activity of PEPC; (3) QEVMIGYSDSGKDAG (nos. 586–600) represents a highly conserved region containing the lysine residue implicated in the active site; and (4) the glycine-rich motif FHGRGGTVGRGGGP (nos. 629–642) is part of the substrate binding site (Lepiniec *et al*. 1994). Disulphide bridges as well as cysteine residues seem to be involved in the redox regulation of the PEPC activity. Lepiniec *et al*. (1994) described seven conserved cysteine residues (nos. 188, 300, 327, 410, 415, 417 and 678) and Besnard *et al*. (2003) described additional two (nos. 57 and 544), which are present in all plant PEPC proteins and also in *Clusia*. Furthermore, the three PEP binding sites (nos. 280, 555 and 589), two Mg^{2+} binding sites (nos. 557 and 594) and four aspartate binding sites (nos. 638, 826, 885 and 960) described for *Flaveria* (Engelmann *et al*. 2003) are present.

The amino acids upstream of the serine residue (phosphorylation site) were characterized as isoform-specific in

maize (Dong *et al*. 1998), but this is not true for *Clusia* because all isoforms show the same sequence NKLEKLA (nos. 1–7). Another part of the maize protein (in *Clusia* nos. 331–348) was also declared as isoform-specific, but shows no such relationship in *Clusia* and is not strongly conserved. One amino acid assumed to be C_4-specific is also present in *Clusia* (V at 333), as well as two (R at 340, S at 342), which were identified as root-specific (Dong *et al*. 1998). These could be identified in all isoforms in *Clusia*. Three other motifs suggested as C_4-specific (Dong *et al*. 1998) do not exist in *Clusia*. *Clusia* sequences have, however, many similarities with isoforms from roots, or C_3 plants in general. We thus assume that the CAM isoforms of PEPC are more similar to the isoforms with anaplerotic functions (nos. 112–130, 561–570 and 876–887) than with those involved in the C_4 type of photosynthesis.

In summary, our data indicate that as far as the genus *Clusia* is concerned, intermediate species such as *C. minor* can switch between the C_3 and CAM modes of photosynthesis by using the C_3 isoform of PEPC. There is still the possibility that intermediate *Clusia* species contain a CAM-isoform of PEPC, which we did not detect. This is, however, rather unlikely as in plants, containing the CAM-isoform, the latter made up between 25 and 50% of all isoforms identified and should thus not have been missed.

Taken together, we thus conclude that for transition to CAM, the C_3-type protein is obviously sufficient.

ACKNOWLEDGMENTS

The authors are indebted to Prof. Dr. Ulrich Lüttge for the donation of the *Clusia* plants, for critical reading of the manuscript and helpful suggestions, as well as for his general support, Dr. Fabio Scarano (University of Rio de Janeiro, Brazil) for sharing his great knowledge about the genus *Clusia*, Heitor Duarte and Corinna Schubert for the help with the switching experiments, and the Landesgraduierten-Förderung Baden-Württemberg and the Reinhold-und-Maria-Teufel foundation for financial support.

REFERENCES

Besnard G., Offmann B., Robert C., Rouch C. & Cadet F. (2002) Assessment of the C4 phospho*enol*pyruvate carboxylase gene diversity in grasses (Poaceae). *Theoretical and Applied Genetics* **105,** 404–412.

Besnard G., Pincon G., D'Hont A., Hoarau J.-Y., Cadet F. & Offmann B. (2003) Characterisation of the phospho*enol*pyruvate carboxylase gene family in sugarcane (*Saccharum* spp.). *Theoretical and Applied Genetics* **107,** 407–478.

Bläsing O.E., Westhoff P. & Svensson P. (2000) Evolution of C4 phospho*enol*pyruvate carboxylase in *Flaveria*, a conserved serine residue in the carboxyl-terminal part of the enzyme is a major determinant for C4-specific characteristics. *Journal of Biological Chemistry* **275,** 27 917–27 923.

Borland A.M. & Griffiths H. (1997) A comparative study on the regulation of C_3 and C_4 carboxylation processes in the constitutive crassulacean acid metabolism (CAM) plant *Kalanchoe daigremontiana* and the C_3-CAM intermediate *Clusia minor*. *Planta* **201,** 368–378.

Borland A.M., Griffith H., Maxwell C., Broadmeadow M.S.J., Griffith N.M. & Barnes J.D. (1992) On the ecophysiology of Clusiaceae in Trinidad: expression of CAM in *Clusia minor* L. during the transition from wet to dry season and characterization of three endemic species. *New Phytologist* **122,** 349–357.

Chollet R., Vidal J. & O'Leary M.H. (1996) Phospho*enol*pyruvate carboxylase: a ubiquitous, highly regulated enzyme in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47,** 273–298.

Crétin C., Santi S., Keryer E., Lepiniec L., Tagu D., Vidal J. & Gadal P. (1991) The phospho*enol*pyruvate carboxylase gene family of *Sorghum*: promoter structures, amino acid sequences and expression of genes. *Gene* **99,** 87–94.

Cushman J.C. & Bohnert H.J. (1989) Nucleotide sequence of the *Ppc*2 gene encoding a housekeeping isoform of phospho*enol*pyruvate carboxylase from *Mesembryanthemum crystallinum*. *Nucleic Acids Research* **17,** 6743–6744.

Cushman J.C. & Bohnert H.J. (1996) Transcriptional activation of CAM genes during development and environmental stress. In *Crassulacean Acid Metabolism: Biochemistry, Ecophysiology and Evolution* (eds K. Winter & J.A.C. Smith) pp. 135–158. Springer-Verlag, Berlin, Heidelberg, Germany.

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

Dong L.-Y., Masuda T., Kawamura T., Hata S. & Izui K. (1998) Cloning, expression, and characterization of a root-form phospho*enol*pyruvate carboxylase from *Zea mays*: comparison with the C_4-form enzyme. *Plant and Cell Physiology* **39,** 865–873.

Engelmann S., Bläsing O.E., Gowik U., Svensson P. & Westhoff P. (2003) Molecular evolution of C_4 phospho*enol*pyruvate carboxylase in the genus *Flaveria* – a gradual increase from C_3 to C_4 characteristics. *Planta* **217,** 717–725.

Ermolova N.V., Cushman M.A., Taybi T., Condon S.A., Cushman J.C. & Chollet R. (2003) Expression, purification and initial characterization of a recombinant form of plant PEP-carboxylase kinase from CAM-induced *Mesembryanthemum crystallinum* with enhanced solubility in *Escherichia coli*. *Protein Expression and Purification* **29,** 123–131.

Ernst K. & Westhoff P. (1997) The phospho*enol*pyruvate carboxylase (*ppc*) gene family of *Flaveria trinervia* (C_4) and *F. pringlei* (C_3): molecular characterization and expression analysis for the *ppcB* and *ppcC* genes. *Plant Molecular Biology* **34,** 427–443.

Gehrig H., Taybi T., Kluge M. & 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 Letters* **377,** 399–402.

Gehrig H., Faist K. & Kluge M. (1998a) Identification of phospho*enol*pyruvate carboxylase isoforms in leaf, stem and roots of the obligate CAM plant *Vanilla planifolia* Salib. (Orchidaceae): a physiological and molecular approach. *Plant Molecular Biology* **38,** 1215–1223.

Gehrig H.H., Heute V. & Kluge M. (1998b) Toward a better knowledge of the molecular evolution of phosphoenolpyruvate carboxylase by comparison of partial cDNA sequences. *Journal of Molecular Evolution* **46,** 107–114.

Gehrig H.H., Winter K., Cushman J., Borland A. & Taybi T. (2000) An improved RNA isolation method for succulent plant species rich in polyphenols and polysaccharides. *Plant Molecular Biology Reporter* **18,** 369–376.

Gehrig H., Heute V. & Kluge M. (2001) New partial sequences of

phospho*enol*pyruvate carboxylase as molecular phylogenetic markers. *Molecular Phylogenetics and Evolution* **20,** 262–274.

Gehrig H.H., Aranda J., Cushman M.A., Virgo A., Cushman J.C., Hammel B.E. & Winter K. (2003) Cladogramm of Panamanian *Clusia* based on nuclear DNA: implications of the origins of crassulacean acid metabolism. *Plant Biology* **5,** 59–70.

Gehrig H.H., Wood J.A., Cushman M.A., Virgo A., Cushman J.C. & Winter K. (2005) Large gene family of phospho*enol*pyruvate carboxylase in the crassulaceen acid metabolism plant *Kalanchoe pinnata* (Crassulaceae) characterised by partal cDNA sequence analysis. *Functional Plant Biology* **32,** 467–472.

Gowik U., Engelmann S., Bläsing O.E., Raghavendra A.S. & Westhoff P. (2006) Evolution of C_4 phosphoenolpyruvate carboxylase in the genus *Alternanthera*: gene families and the enzymatic characteristics of the C_4 isozyme and its orthologues in C_3 and C_3/C4 Alternantheras. *Planta* **223,** 359–368.

Grams T.E.E. & Thiel S. (2002) High light-induced switch from C3-photosynthesis to crassulacean acid metabolism is mediated by UV-A/blue light. *Journal of Experimental Botany* **53,** 1475–1483.

Grams T.E.E., Herzog B. & Lüttge U. (1998) Are there species in the genus *Clusia* with obligate C_3-photosynthesis? *Journal of Plant Physiology* **152,** 1–9.

Haag-Kerwer A., Grams T.E.E., Olivares E., Ball E., Arndt S, Popp M., Medina E. & Lüttge U. (1996) Comparative measurements of gas-exchange, acid accumulation and chlorophyll a fluorescence of different species of *Clusia* showing C3 photosynthesis, or crassulacean acid metabolism, at the same field site in Venezuela. *New Phytologist* **134,** 215–226.

Harpster M.H. & Taylor W.C. (1986) Maize phosphoenolpyruvate carboxylase. *The Journal of Biological Chemistry* **261,** 6132–6136.

Hermans J. & Westhoff P. (1992) Homologous genes for the C_4 isoform of phospho*enol*pyruvate carboxylase in a C_3 and a C_4 *Flaveria* species. *Molecular and General Genetics* **234,** 275–284.

Herppich W.B. & Herppich M. (1997) The interrelationship between changes in PEPC activity and organic acid accumulation during the C_3-CAM shift in *Mesembryanthemum crystallinum*. *Journal of Plant Physiology* **151,** 373–378.

Herzog B., Hübner C., Ball E., Bastos R.D.N., Franco A.C., Scarano F.R. & Lüttge U. (1999) Comparative study of the C_3/CAM intermediate species *Clusia parviflora* Saldanha et Engl. and the obligate CAM species *Clusia hilariana* Schlecht. growing sympatrically exposed and shaded in the coastal restinga of Brazil. *Plant Biology* **1,** 453–459.

Honda H., Okamoto T. & Shimada H. (1996) Isolation of a cDNA for a phospho*enol*pyruvate carboxylase from a monocot CAM-plant, *Aloe arborescens*: structure and its gene expression. *Plant and Cell Physiology* **37,** 881–888.

Hudspeth R.L. & Grula J.W. (1989) Structure and expression of the maize gene encoding the phospho*enol*pyruvate carboxylase isozyme involved in C_4 photosynthesis. *Plant Molecular Biology* **12,** 579–589.

Kai Y., Matsumura H. & Izui K. (2003) Phospho*enol*pyruvate carboxylase: three-dimensional structure and molecular mechanisms. *Archives of Biochemistry and Biophysics* **414,** 170–179.

Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E. & Tabata S. (2000) Structural analysis of *Arabidopsis thaliana* chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC and BAC clones. *DNA Research* **7,** 217–221.

Kawamura T., Shigesada K., Yanagisawa S. & Izui K. (1990) Phospho*enol*pyruvate carboxylase prevalent in maize roots: isolation of a cDNA clone and its use for analyses of the gene and gene expression. *Journal of Biochemistry* **107,** 165–168.

Kawamura T., Shigesada K., Toh H., Okumura S., Yanagisawa S. & Izui K. (1992) Molecular evolution of phospho*enol*pyruvate carboxylase for C_4 photosynthesis in maize: comparison of its cDNA sequence with a newly isolated cDNA encoding an isozyme involved in the anaplerotic function. *Journal of Biochemistry* **112,** 147–154.

Latzko E. & Kelly G.J. (1983) The many-faceted function of phospho*enol*pyruvate carboxylase in C3 plants. *Physiologie Végétale* **21,** 805–815.

Lepiniec L., Keryer E., Tagu D., Gadal P. & Cretin C. (1992) Complete nucleotide sequence of a *Sorghum* gene coding for the phospho*enol*pyruvate carboxylase involved in C_4 photosynthesis. *Plant Molecular Biology* **19,** 339–342.

Lepiniec L., Keryer E., Philippe H., Gadal P. & Crétin C. (1993) *Sorghum* phospho*enol*pyruvate carboxylase gene family: structure, function and molecular evolution. *Plant Molecular Biology* **21,** 487–502.

Lepiniec L., Vidal J., Chollet R., Gadal P. & Cretin C. (1994) Phospho*enol*pyruvate carboxylase: structure, regulation and evolution. *Plant Science* **99,** 111–124.

Lüttge U. (1991) Clusia – morphogenetische, physiologische und biochemische Strategien von Baumwürgern im tropischen Wald. *Naturwissenschaften* **78,** 49–58.

Lüttge U. (2004) Ecophysiology of crassulacean acid metabolism. *Annals of Botany* **93,** 629–652.

Matsuoka M. & Minami E. (1989) Complete structure of the gene for phospho*enol*pyruvate carboxylase from maize. *European Journal of Biochemistry* **181,** 593–598.

Merkelbach S., Gehlen J., Denecke M., Hirsch H.-J. & Kreuzaler F. (1993) Cloning, sequence analysis and expression of a cDNA encoding active phospho*enol*pyruvate carboxylase of the C_3 plant *Solanum tuberosum*. *Plant Molecular Biology* **23,** 881–888.

Pipoly J.J., Kearns D.M. & Berry P.E. (1998) *Clusia* L. In *Flora of the Venezuelan Guayana* (eds P.E. Berry, B.K. Holst & K. Yatskievych) pp. 260–294. Missouri Botanical Garden Press, St. Louis, MO, USA.

Poetsch W., Hermans J. & Westhoff P. (1991) Multiple cDNAs of phospho*enol*pyruvate carboxylase in the C_4 dicot *Flaveria trinervia*. *FEBS Letters* **292,** 133–136.

Rickers J., Cushman J.C., Michalowski C.B., Schmitt J.M. & Bohnert H.J. (1989) Expression of the CAM-form of phospho*enol*pyruvate carboxylase and nucleotide sequence of a full length cDNA from *Mesembryanthemum crystallinum*. *Molecular and General Genetics* **215,** 447–454.

Sanger F., Nicklen S. & Coulson A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74,** 5463–5467.

Slocombe S.P., Whitelam G.C. & Cockburn W. (1993) Investigation of phospho*enol*pyruvate carboxylase (PEPCase) in *Mesembryanthemum crystallinum* L. in C_3 and CAM photosynthetic states. *Plant, Cell & Environment* **16,** 403–411.

Svensson P., Bläsing O.E. & Westhoff P. (1997) Evolution of the enzymatic characteristics of C_4 phospho*enol*pyruvate carboxylase. *European Journal of Biochemistry* **246,** 452–460.

Swofford D.L. (1999) *PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods)*. Sinauer Associates, Sunderland, MA, USA.

Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F. & Higgins D.G. (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25,** 4876–4882.

Toh H., Kawamura T. & Izui K. (1994) Molecular evolution of phospho*enol*pyruvate carboxylase. *Plant, Cell & Environment* **17,** 31–43.

Tsuchida Y., Furumoto T., Izumida A., Hata S. & Izui K. (2001) Phospho*enol*pyruvate carboxylase kinase involved in C_4 photosynthesis in *Flaveria trinervia*: cDNA cloning and characterization. *FEBS Letters* **507,** 318–322.

Vaasen A., Begerow D., LÜttge U. & Hampp R. (2002) The genus

Clusia L.: molecular evidence for independent evolution of photosynthetic flexibility. *Plant Biology* **4,** 86–93.

Vazquez-Tello A., Whittier R.F., Kawasaki T., Sugimoto T., Kawamura Y. & Shibata D. (1993) Sequence of a soybean (*Glycine max* L.) phospho*enol*pyruvate carboxylase cDNA. *Plant Physiology* **103,** 1025–1026.

Willis J.C. (1973) *A Dictionary of the Flowering Plants and Ferns*, 8th edn, University Press, Cambridge, UK.

Winter K. & Smith J.A.C. (1996) Crassulacean acid metabolism: current status and perspectives. In *Crassulacean Acid Metabolism: Biochemistry, Ecophysiology and Evolution* (eds K. Winter & J.A,C. Smith) pp. 389–426. Ecological Studies 114, Springer-Verlag, Berlin, Heidelberg, Germany.

Zotz G. & Winter K. (1993) Short-term regulation of crassulacean acid metabolism activity in a tropical hemiepiphyte, *Clusia uvitana*. *Plant Physiology* **102,** 835–841.

Received 10 June 2006; received in revised form 7 July 2006; accepted for publication 7 July 2006

