Mol Gen Genet (1990) 224:459–468

MGG

Analysis of expression and evolutionary relationships of phosphoenolpyruvate carboxylase genes in *Flaveria trinervia* (C_4) and *F. pringlei* (C_3)

Jürgen Hermans and Peter Westhoff

Institut für Entwicklungs- und Molekularbiologie der Pflanzen, Heinrich-Heine Universität Düsseldorf, W-4000 Düsseldorf 1, Federal Republic of Germany

Received July 19, 1990

Summary. Phosphoenolpyruvate carboxylase (PEPCase) was shown to be encoded by a multigene family in various *Flaveria* species analysed. Several clones were isolated from genomic libraries of *F. pringlei* (C_3 species) and *F. trinervia* (C_4 species) and classified into four distinct groups according to their hybridization behaviour to a full-length cDNA clone encoding the PEPCase C_4 isozyme of *F. trinervia*. A detailed cross-hybridization analysis demonstrated that the closest relative of most of the PEPCase genes isolated from *F. trinervia* and *F. pringlei* was not found in the same but in the other species. Northern analysis, using stringent conditions, allowed discrimination of class-specific PEPCase transcripts and revealed characteristic organ-specific expression patterns.

Key words: *Flaveria* – C_3/C_4 plants – Phosphoenolpyruvate carboxylase – Gene family – Evolution

Introduction

Various higher plant species which inhabit desert or semi-arid tropical environments have a special type of photosynthetic carbon assimilation called the C_4 cycle. Photosynthesis in these plants relies upon the integrated metabolic activities of two morphologically and functionally distinct cell types, mesophyll and bundle sheath cells. As a consequence of this metabolic interaction, carbon dioxide is concentrated in the bundle sheath cells at the site of ribulose-1,5-bisphosphate carboxylase. Thus, photorespiration is largely abolished in plants exhibiting C_4 photosynthesis (e.g. Edwards and Walker 1983).

Since C_4 species are widely scattered among various families of mono- and dicotyledonous plants this special kind of metabolism has probably envolved several times independently in angiosperms (Moore 1982). This view is supported by the observation that the enzymes which permit C_4 photosynthesis are also found in the leaves of typical C_3 plants. Thus, the evolution of C_4 plants obviously took advantage of a set of already existing genes in ancestral plants (Moore 1982; Cockburn 1983). However, the temporal and spatial expression patterns of these genes as well as the regulatory properties of the enzymes encoded had to be changed to meet the special requirements of the C_4 photosynthesis pathway.

One of the key enzymes in C_4 metabolism is phosphoenolpyruvate carboxylase (PEPCase). This enzyme is located in the cytosol of the mesophyll cells and catalyzes the primary fixation of CO_2 into oxaloacetate (O'Leary 1982). In the NADP malic enzyme subgroup of C_4 species, oxaloacetate is subsequently reduced in the mesophyll chloroplasts to form malate, which is then transported into the bundle sheath cells. After decarboxylation, the resulting CO_2 is finally channeled into the Calvin-Benson cycle (Hatch and Osmond 1976; Edwards and Walker 1983; Hatch 1987).

Besides the C_4 specific PEPCase, several other isozymic forms have been detected in the various organs of C_4 as well as C_3 plants. These isoforms can be distinguished by chromatographic, immunological and kinetic properties (Ting and Osmond 1973; O'Leary 1982). It has been postulated that they are involved in diverse anapleurotic functions in basic plant metabolism such as gluconeogenesis, NADPH regeneration or the refixation of respired CO_2 (Andreo et al. 1987) and last, but not least, in the function of guard cells (Cockburn 1983).

A number of different genes which are differentially expressed during leaf development and in various plant organs (Vidal and Gadal 1983; Martineau and Taylor 1985; Harpster and Taylor 1986; Hudspeth and Grula 1989) have been identified in the monocotyledonous C_4 plant maize (Harpster and Taylor 1986; Hudspeth et al. 1986). Therefore it is reasonable to expect that the various isoforms of PEPCase are encoded by different

Offprint requests to: P. Westhoff

genes. Hence, one of the primary events in the evolution of C_4 plants may have been the development of a regulatory mechanism which generates novel patterns of PEPCase gene expression.

To gain an insight into the changes underlying the evolution of C_3 to C_4 plants we have started to analyse C_4 photosynthesis in the genus *Flaveria* (Asteraceae; Powell 1976). This genus is characterized by a wide range of C_3, C_4 and C_3-C_4 intermediate species, the latter exhibiting photosynthetic and anatomical properties ranging between those of C_3 and C_4 plants (Monson et al. 1984; Edwards and Ku 1987). It is a commonly accepted hypothesis that these C_3-C_4 *Flaveria* species may be regarded as representing transition forms from C_3 to C_4 photosynthesis (Monson and Moore 1989). Therefore, the genus *Flaveria* provides an attractive biological system for studying the evolution of C_4 plants.

As a prerequisite, a full-length cDNA (pcFtppc1-1) coding for the C_4 isoform of PEPCase in *F. trinervia* (Poetsch et al., in preparation) was isolated. To understand the genetic basis of evolution towards the C_4 syndrome, including the changes in regulatory mechanisms, PEPCase genes in *Flaveria* species were characterized. The isolation of genomic clones encoding PEPCase in a C_3 (*F. pringlei*) and a C_4 type *Flaveria* species (*F. trinervia*), their expression patterns and their putative evolutionary relationships are reported here.

Materials and methods

Growth of plant material and isolation of nucleic acids. Seeds of *F. trinervia* (C_4), *F. anomala* (C_3-C_4 intermediate), *F. pringlei* (C_3) and *F. brownii* (C_4-like) were obtained from H. Bauwe (ADW Gatersleben, GDR) or S. Holaday (Lubbock, Texas). The Tx 623 line of *Sorghum bicolor* (C_4) was a gift of Pioneer Hibred. Plants were grown in the greenhouse at 20–26° C (day) and 18–20° C (night). Supplementary light was provided by a combination of sodium and mercury high-pressure vapour lamps giving a photosynthetic photon flux density of 300 $\mu E/m^2$ per s at plant height. For the preparation of root tissue, cuttings were grown on Vermiculite under the same conditions.

Nuclear DNA was prepared from chromatin (Steinmüller and Apel 1986) isolated from fully expanded leaves. Total RNA was extracted from various *F. trinervia* and *F. pringlei* tissues following the procedure of Westhoff et al. (in preparation). Poly(A)$^+$ RNA was enriched on oligo(dT) cellulose according to Aviv and Leder (1972).

Isolation and labelling of plasmid DNA fragments. Plasmid DNA was isolated by alkaline lysis and restricted by standard procedures (Sambrook et al. 1989). DNA fragments were fractionated in agarose gels and purified by electroelution to NA45 membranes following the suggestions of the manufacturers (Schleicher and Schüll, Dassel, FRG). Hybridization probes were labelled by random priming (Feinberg and Vogelstein 1983) to an average specific activity of 1.7×10^9 dpm per µg DNA.

Southern analysis of genomic DNA. Nuclear DNAs of the various species were digested to completion using *Bam*HI, *Eco*RI, *Hin*dIII and *Pst*I restriction enzymes. Ten micrograms each of the fragmented DNAs were separated in 0.4% agarose gels and transferred to Biodyne B membranes as recommended by the manufacturer (Pall, Dreieich, FRG). Homologous hybridizations were carried out overnight in 50% (v/v) formamide, 1 × Denhardt's solution, 50 µg/ml heterologous DNA, 1% (w/v) SDS at 42° C. For heterologous hybridizations the formamide concentration and the temperature were decreased to 30% (v/v) and 37° C, respectively. The volume activity of the probe was 10^6 dpm per ml solution. After hybridization the filters were washed briefly several times in 2 × SSC, 0.1% (w/v) SDS at the hybridization temperature and, in the case of homologous hybridizations, finally for 1 h in 0.1 × SSC, 0.1% (w/v) SDS.

Construction and screening of genomic libraries. To generate fragments suitable for insertion into lambda vectors, the DNAs of *F. pringlei* and *F. trinervia* were treated with various restriction enzymes. *F. trinervia* DNA was digested to completion with *Eco*RI. In addition, partial digestions were carried out with *Eco*RI, *Hin*dIII and *Sau*3A using several enzyme to DNA ratios. In the case of *F. pringlei* only partial digestions with *Sau*3A were performed. The fragments obtained by the different assays were sized to 13–25 kb either by agarose gel electrophoresis followed by electroelution to NA45 membranes or by sucrose gradient centrifugation (Sambrook et al. 1989).

The genomic libraries were constructed in lambda DASH (Stratagene, La Jolla, USA) using standard procedures (Sambrook et al. 1989). Plating was done on *Escherichia coli* strain CES 200 (Wyman and Wertman 1987). A total of 2.5×10^6 (*F. trinervia*) and 5×10^5 (*F. pringlei*) phage were screened without amplification by plaque hybridization (Benton and Davis 1977). The hybridization was performed in a sodium phosphate/SDS/EDTA buffer (Church and Gilbert 1984) at 50° C. After hybridization the filters were incubated several times in 2 × SSC, 0.1% (w/v) SDS at the hybridization temperature.

Isolation of phage DNA and restriction site mapping. Phage were grown and the DNA was isolated by standard procedures (Sambrook et al. 1989) except for a minor modification: the intact phage were precipitated from the culture medium by the addition of ammonium sulfate to 50% saturation instead of polyethylene glycol.

For the localization of restriction sites the recombinant phage DNAs were first cut with an appropriate restriction enzyme to release the cloned fragment. Without further fragment isolation the DNA was partially digested by the enzyme to be mapped. The partial products were used for Southern blotting and hybridized with probes specific to the outermost regions of the cloned insert. To generate these probes the whole recombinant phage DNA was completely digested by a frequently cutting enzyme (*Rsa*I) and used as a template for in vitro transcriptions with T3 and T7 RNA polymerase.

The respective promoters in lambda DASH flank the multiple cloning sites, thus run-off transcription from both promoters yields RNAs extending up to the first *Rsa*I site. After exposure of the Southern blots, the resulting ladder of bands reflects the distances between the sites.

Slot-blot hybridization. DNA (0.5 µg) of each recombinant phage was denatured in 0.4 M NaOH, 0.6 M NaCl and bound to Biodyne B membranes using a slot-blot device (Schleicher and Schüll). Hybridization conditions were the same as described for screening of the genomic libraries. Different stringencies were achieved by performing the hybridization reactions at varying temperatures. The temperatures used are indicated in each particular experiment (see Results). The amount of probe DNA was chosen to give a molar ratio of membrane bound to probe DNA of about 200–400. To quantitate hybridization the individual slots were cut into equal pieces, placed into Readymix scintillation cocktail (Beckman, München, FRG), and the amount of probe bound to each slot was determined by liquid scintillation spectrometry. Background hybridization was determined by hybridization to vector DNA and was subtracted from each value.

Computer analysis. Phylogenetic tree analysis was performed with the program KITSCH of the Phylogeny Interference Package (PHYLIP, Vers. 3.2; J. Felsenstein, University of Washington).

RNA analysis. Northern blotting using glyoxylated RNA (McMaster and Carmichael 1977) has been previously described by Westhoff and Herrmann (1988). Hybridization was at 75° C with the sodium phosphate/SDS/EDTA buffer as above.

Results

Genomic Southern blot analysis

The leaf-specific *F. trinervia* PEPCase cDNA clone pcFtppc1-1 which was isolated from a lambda gt11 cDNA expression library (Poetsch et al., in preparation) was used to probe for PEPCase sequences in the nuclear DNA of *F. trinervia* (C_4), *F. pringlei* (C_3), *F. anomala* (C_3-C_4 intermediate) and *F. brownii* (C_4-like). For comparison, the monocotyledonous C_4 plant *S. bicolor* was included in the analysis, since a PEPCase cDNA clone

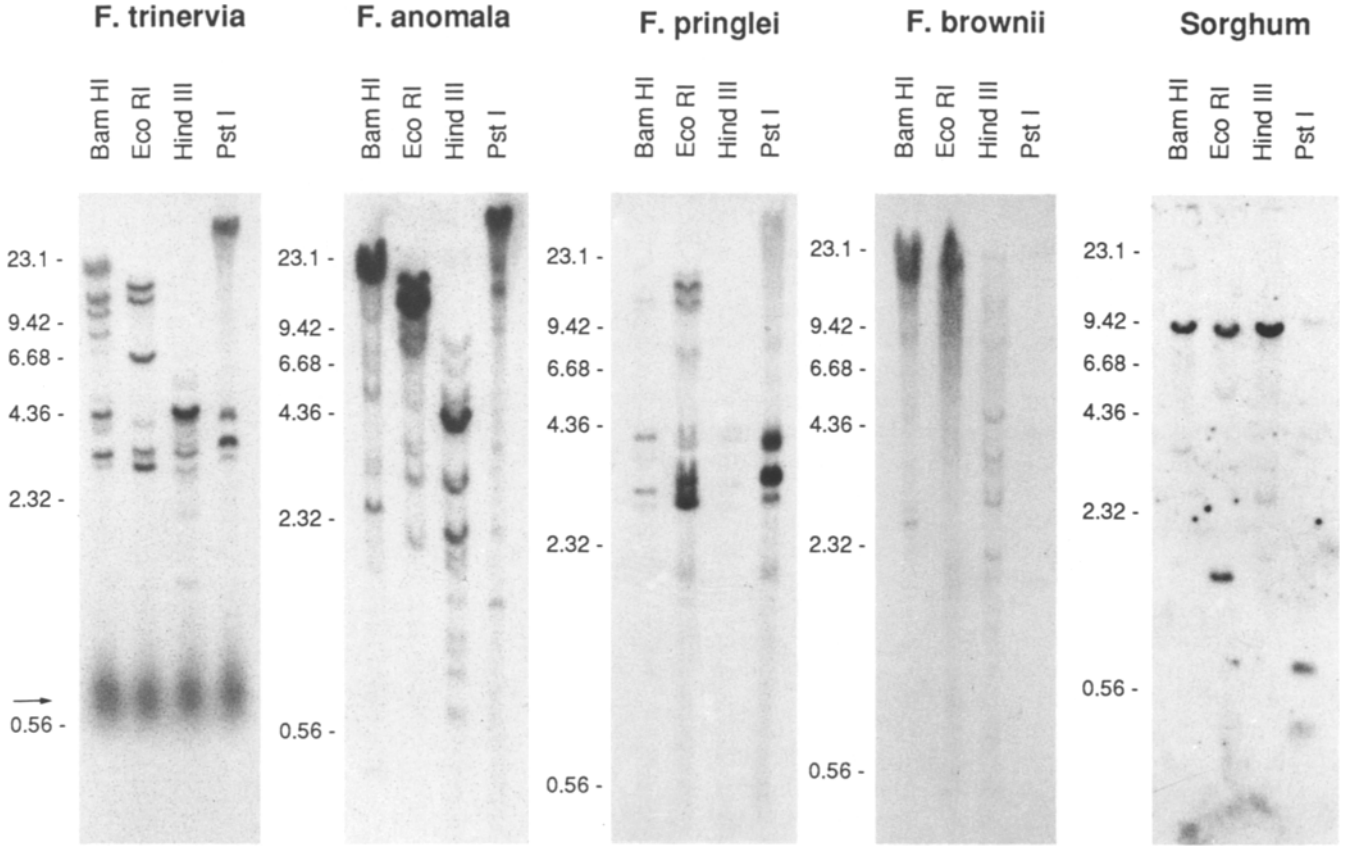

Fig. 1. Genomic Southern blot analysis of *Flaveria* species and *Sorghum bicolor*. Genomic DNAs (10 µg) were digested with different restriction enzymes as indicated above the lanes and analysed by Southern blotting. The random-primed labelled cDNAs pcFtppc1-1 from *F. trinervia* (for *Flaveria* species) and pcSbppcP1 from *S. bicolor* (for hybridization to the DNA from the same species; Oswald et al. 1990) were used as probes. The exposure time was 14 days. The *arrow* indicates hybridization to contaminating RNA. Sizes are given in kb. Among the investigated species, *F. brownii* is located most distantly from *F. trinervia*, in terms of phylogeny. Hence, the hybridization signals are weak under the given conditions

Table 1. Frequency of phosphoenolpyruvate carboxylase (PEPCase) clones in genomic lambda libraries

Species	Type of digestion	Size of library $\times 10^6$	Number of positive clones[a]	Positive clones per 10^6 phage
Flaveria trinervia	*Eco*RI total	1.20	4	4.4
	*Eco*RI partial	0.50	33	27.5
	*Hin*dIII partial	0.45	6	13.3
	*Sau*3A partial	0.75	2	2.7
F. pringlei	*Sau*3A partial	0.30	23	77.0

[a] Each library was screened with the *F. trinervia* cDNA pcFtppc1-1 at 50° C

specific for the C_4 isoform of PEPCase was available for that species (Oswald et al. 1990; Crétin et al. 1990).

A multi-banded hybridization pattern can be detected even under stringent conditions when the *F. trinervia* leaf-specific cDNA is hybridized to nuclear DNA from the same species (Fig. 1). The pattern indicates the presence of least three genes which show close relationship to the leaf-specific cDNA pcFtppc1-1. In contrast, only one or two bands are visible in *S. bicolor* after stringent washings (Fig. 1). This permits the conclusion that sequences highly homologous to the leaf-specific cDNA are probably represented by a single-copy gene, just as in *Zea mays* (Harpster and Taylor 1986; Matsuoka and Minami 1989; Hudspeth and Grula 1989).

Multiple hybridization signals are also obtained with the other *Flaveria* species investigated, again suggesting the existence of multigene families (Fig. 1).

Screening of genomic libraries

It is assumed that restriction enzyme sites are randomly distributed throughout the DNA and that a partial digest with a particular enzyme should give fragments of clonable sizes, allowing the construction of a representative library. In practice, however, there is often considerable heterogeneity for the distances between cleavage sites. This may lead to a severe underrepresentation of particular sequences in the final library (Seed et al. 1982). The multiplicity of bands revealed by Southern blot analysis indicated that it could be advantageous to construct a multiple set of libraries based on various enzymatic digestions at various stringencies, when it is attempted to clone the complete range of PEPCase genes in *F. trinervia* and *F. pringlei*. For this reason several libraries were prepared in lambda DASH (Table 1). Since the amount of nuclear DNA in *Flaveria* species is not known, rather large libraries ($0.3–1.2 \times 10^6$ independent clones) were constructed (Table 1).

By their ability to cross-hybridize to the cDNA pcFtppc1-1 at low stringency, 45 clones were isolated from the *F. trinervia* and 23 from the *F. pringlei* libraries. As expected, the number of positive clones isolated from the various libraries was quite different (Table 1). All clones were compared by restriction analysis using *Eco*RI, *Hin*dIII and *Xba*I to recognize identical or overlapping clones (data not included). Twelve clones from *F. trinervia* and 18 from *F. pringlei* proved to be clearly distinct and were selected for further characterization.

Classification of clones by hybridization analysis with the pcFtppc1-1 cDNA

To obtain information on the degree of homology between the various clones, phage DNA was isolated and initially subjected to slot-blot analysis using the complete pcFtppc1-1 cDNA as a probe. Discrimination between non-identical but similar DNA species can only be achieved if the probe DNA is limiting and thus saturation of the filter-bound DNA is avoided (Beltz et al. 1983). Therefore, at least 200-fold less probe than filter-bound DNA was used in hybridization experiments. Hybridizations were carried out at varying temperatures (75, 68, 60, 50 and 40° C) to estimate the melting temperatures of the hybrids. Based on melting behaviours, both the *F. trinervia* and *F. pringlei* clones could be roughly classified into four categories as shown in Table 2.

In separate experiments, a 240 bp *Bst*EII-*Eco*RI fragment from the 3′ non-translated region and a 135 bp *Eco*RI/*Hin*dIII fragment located at the 5′ end of pcFtppc1-1 were hybridized to the class I–IV clones of *F. trinervia* and *F. pringlei*. As expected, the 3′ and 5′ probes gave signals with decreasing intensity in the class I–III clones of *F. trinervia*, but failed to react with the

Table 2. Similarity of genomic PEPCase clones to the cDNA pcFtppc1-1

Species	Class	Melting temperature with the cDNA pcFtppc1-1	Clone designation
F. trinervia	I	>75° C	lnFtppc8, 7 (overlapping), 43, 47
	II	60–75° C	lnFtppc11, 44, 52, 55, 56
	III	50–60° C	lnFtppc12, 32, 54
	IV	<55° C	lnFtppc49
F. pringlei	I	>70° C	lnFpppc65, 80
	II	60–70° C	lnFpppc71, 59, 72, 76, 77, 79
	III	50–60° C	lnFpppc64, 62, 67
	IV	<50° C	lnFpppc66, 61, 63, 68, 69, 70, 74

DNA of each distinct recombinant phage was isolated and blotted onto Biodyne B membranes. Hybridization was carried out with the *F. trinervia* cDNA pcFtppc1-1 at 75, 68, 60, 50 and 40° C. The intensity of each signal was estimated by liquid scintillation counting. Melting curves were drawn to estimate the approximate melting points (data not shown). The designations of the clones used for further investigations are underlined

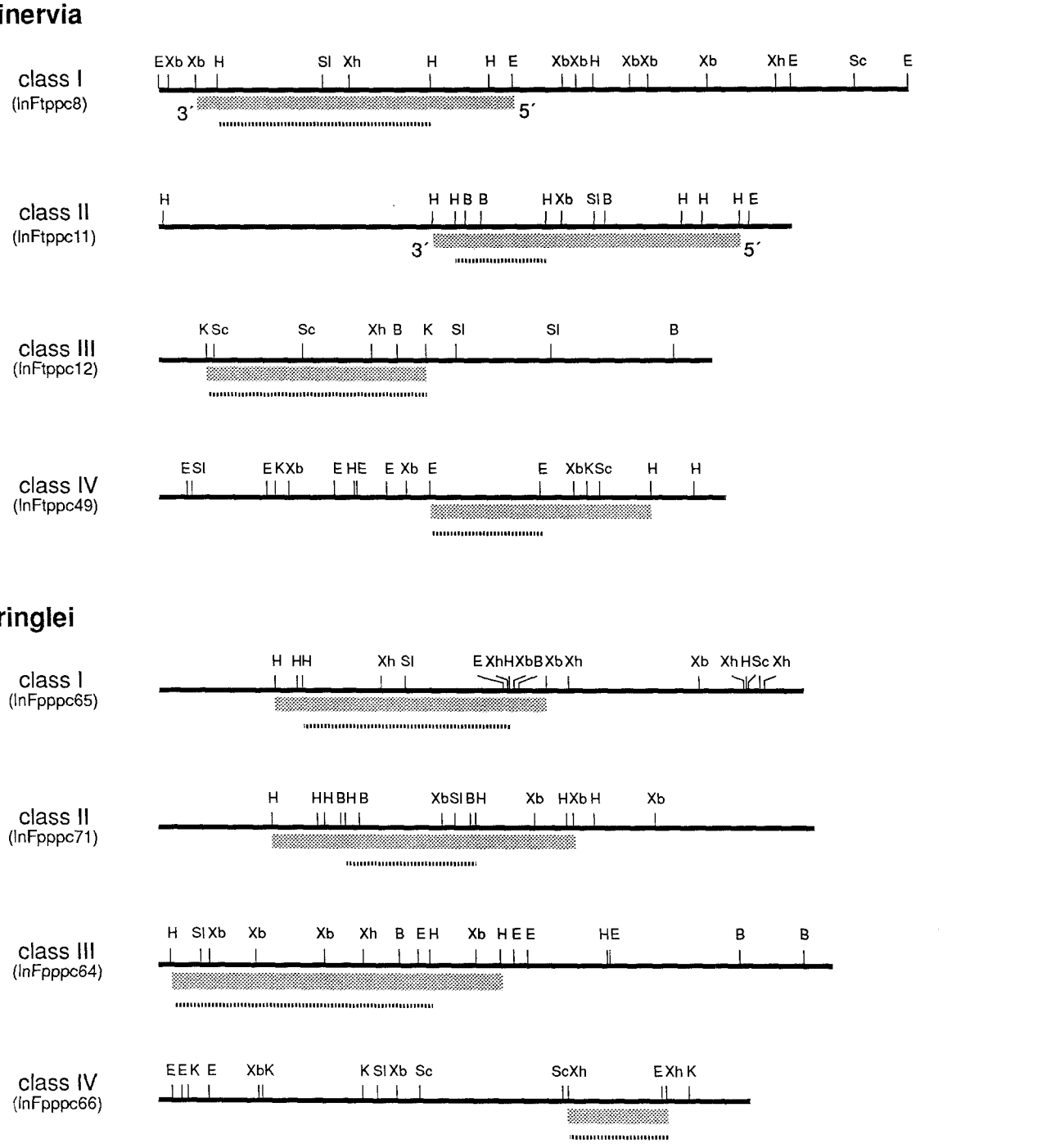

Fig. 2. Restriction maps of phosphoenolpyruvate carboxylase (PEPCase) phage clones. The DNA inserts are drawn in the orientation from left to right lambda arms, the locations of the genes are depicted as *grey bars*. Subclones for cross-hybridization and expression analysis are marked with *dotted lines*. Abbreviations for the designation of restriction endonucleases: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; Sl, *Sal*I; Sc, *Sac*I; Xb, *Xba*I; Xh, *Xho*I

F. trinervia class IV clone (data not shown). Surprisingly, the same probes also hybridized to the *F. pringlei* class I–III clones, in the case of the class I clone even at elevated stringency (68° C). The hybridization analysis with the 5′ and 3′ specific probes also allowed selection of those clones which were thought to contain the complete genes. One clone of each class (indicated in Table 2) was chosen for further analysis.

Restriction mapping and localization of PEPCase sequences in genomic clones

Physical maps were constructed using a modification of Smith and Birnstiel's method (1976) as outlined in Material and methods. This novel method circumvents the difficult preparation of single-sided labelled fragments. Figure 2 shows the restriction maps constructed using

the enzymes *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I, *Sac*I, *Sal*I, *Xba*I and *Xho*I.

To determine the approximate locations of the genes and their orientations, filters carrying Southern blots of *Eco*RI, *Hin*dIII and *Xba*I digests of each clone were hybridized with the complete cDNA (*F. trinervia* and *F. pringlei*) and additionally with the 3′ and 5′ specific fragments (*F. trinervia*). The results of these experiments are shown in Fig. 2.

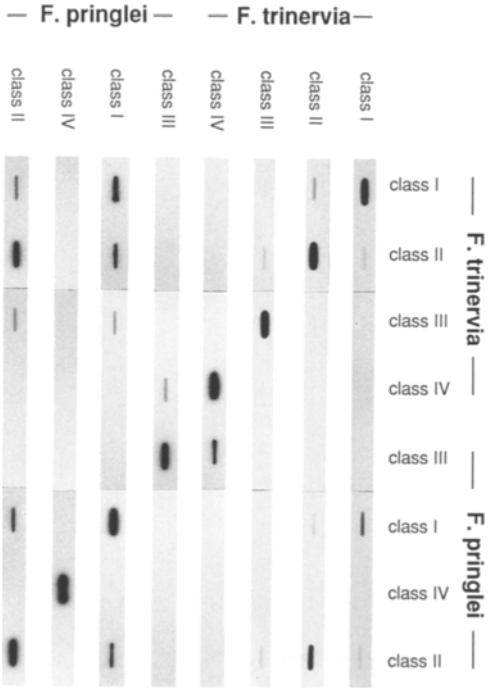

Fig. 3. Cross-hybridization analysis of genomic PEPCase clones. Recombinant phage DNAs of one representative member of each homology class (vertical axis) were immobilized on Biodyne B membranes using a slot-blot device. Hybridization probes of the respective clones (horizontal axis) are shown in Fig. 2. Hybridization was carried out at high stringency (75° C)

Hybridization analysis of similarity between F. trinervia *and* F. pringlei *class I–IV clones*

To estimate sequence similarities between the various classes of isolated genomic clones an extensive cross-hybridization analysis was performed. An internal fragment of each gene as assessed by the hybridization analysis described above (see Fig. 2 for details) was isolated and hybridized to the undigested recombinant phage DNAs at high (75° C; Fig. 3) and low (60° C; blots not shown) stringencies. The values obtained by hybridization with a particular fragment were standardized relative to its homologous hybridization (defined as 100%) and tabulated in a distance matrix (for values obtained at 60° C see Table 3). On the basis of an independent evolution of the PEPCase species in the *Flaveria* complex

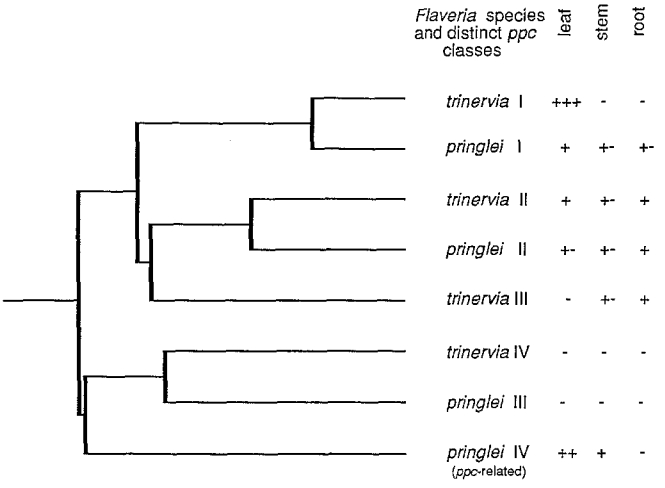

Fig. 4. Dendrogram of the relationships of PEPCase sequences in *F. trinervia* and *F. pringlei*. This tree is based on the distance data obtained by cross-hybridization of the genomic clones at 60° C. The species of origin and the respective homology classes are indicated at the right side. The expression patterns of the various genes are illustrated by +/− symbols as deduced after visual inspection of the fluorographs shown in Fig. 5

Table 3. Distance matrix of genomic PEPCase clones

	F. trinervia				*F. pringlei*			
	Class I	Class II	Class III	Class IV	Class III	Class I	Class IV	Class II
F. trinervia								
Class I	100.0	22.9	10.9	0.5	2.1	74.5	0.0	17.1
Class II	21.2	100.0	18.4	0.0	5.9	27.7	0.0	48.6
Class III	17.3	30.6	100.0	2.2	7.0	27.9	0.0	21.5
Class IV	1.0	1.2	0.8	100.0	36.2	1.3	0.0	1.0
F. pringlei								
Class III	2.4	2.4	1.0	13.0	100.0	0.6	0.0	1.6
Class I	37.4	13.8	11.7	0.7	2.4	100.0	0.0	18.7
Class IV	0.4	0.0	1.0	2.1	8.9	0.4	100.0	0.1
Class II	18.8	56.1	18.4	0.9	1.7	18.1	0.0	100.0

The membrane-bound and the probe DNAs as well as the hybridization conditions, with the exception of the temperature (60° C), were the same as described in Fig. 3. The hybridization signals were normalized to percent hybridization

a phylogenetic tree was calculated. The tree is based on the distance values obtained at 60° C (Table 3) and is depicted in Fig. 4.

Expression analysis of F. trinervia *and* F. pringlei *class I–IV clones*

In order to determine the mode in which the PEPCase genes found in the various classes are expressed in *F. trinervia* and *F. pringlei*, Northern blot analyses were carried out using poly(A)$^+$ RNA from fully expanded leaves, stems and roots of *F. pringlei* as well as of *F. trinervia*. From the latter species, RNA was also prepared from partially expanded leaves (about one-third of mature size) and complete flowers (just before opening). Hybridization was performed under stringent conditions (75° C) with the same probes that were used to estimate the similarities between the various clones. Thus, the conditions of hybridization were suited to minimize cross-hybridization between the various PEPCase RNA sequences. Although cross-hybridization could not be ruled out completely, the amounts could be assessed with the similarity analysis described above.

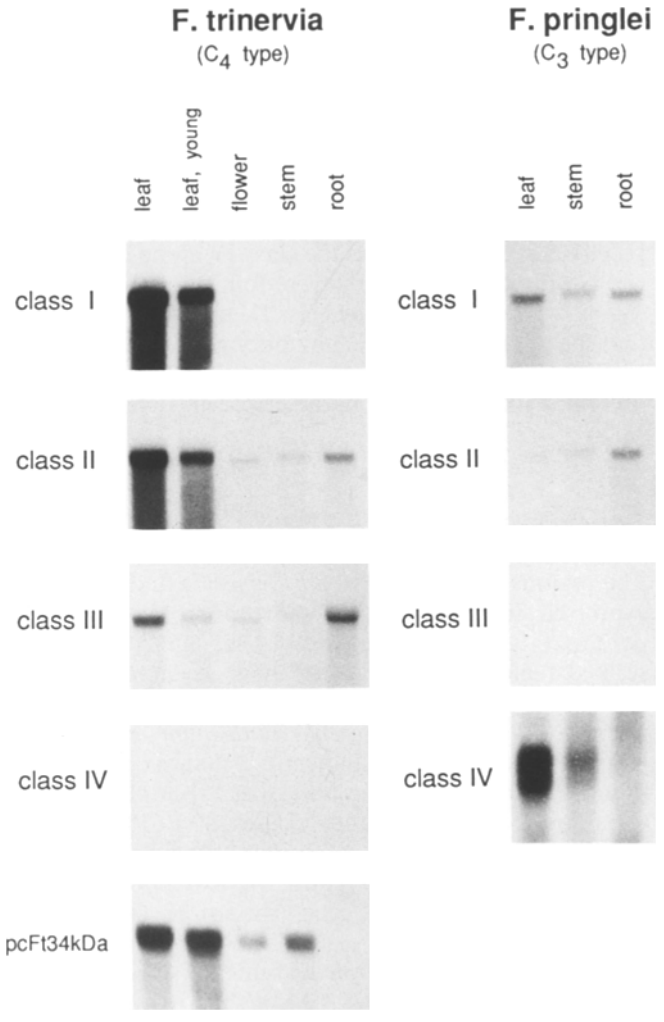

Fig. 5. Northern blot analysis of PEPCase genes of *F. trinervia* and *F. pringlei*. Four micrograms of glyoxylated poly(A)$^+$ RNA isolated from the tissues indicated above the lanes were separated in 1.0% agarose gels and blotted to Biodyne A membranes. Hybridization probes used are indicated on the left side of each panel (see Fig. 2 for the location of the class I–IV probes on the physical maps). Hybridization was carried out at high stringency (75° C). The exposure time was either 1 (*F. trinervia*, class I; 33 kDa protein of the water-splitting apparatus of photosystem II) or 12 days (the others)

The transcripts identified as specific for PEPCase were 3.3 kb. No significant difference was detectable in the sizes of the mRNAs between the different tissues and between the two species. Only the mRNA found in flowers of *F. trinervia* seemed to be about 50–100 bases shorter than the other transcripts. However, quite drastic differences in the strength and in the organ specificity of the expression of the various genes were observed.

When the *F. trinervia* class I clone was used as a probe an intense reaction with leaf RNA was detectable. Measurements with RNA slot-blot hybridizations (data not shown) gave a signal that was about twofold stronger with RNA from fully expanded in contrast to RNA from immature leaves. No hybridization could be detected with RNA from stems, roots and flowers (Fig. 5). The type of expression of this clone is thus identical to that of the cDNA pcFtppc1-1 (Poetsch et al., in preparation), indicating that this gene encodes a C_4 specific isoform of PEPCase.

With the *F. trinervia* class II clone a weak signal is detectable with root RNA, exceeding that obtained with RNA from flowers and stems (Fig. 5). However, hybridization is again most prominent with leaf RNA. Since this clone showed significant cross-reaction to the leaf-specific class I clone even under stringent hybridization conditions (see Fig. 3), the hybridization intensities were quantified by RNA slot-blot experiments at the same temperature (data not included). The ratio in the expression levels in leaf tissue of *F. trinervia* between class I and class II was about 8.0. This was significantly lower than the cross-hybridization ratio value (12.2) of the DNA probes. Therefore, a significant amount of expression of the class II clone occurs in leaves.

A preferential expression of the *F. trinervia* class III clone could be detected in root tissue (Fig. 5). The signal with leaf RNA ranges within the degree of cross-hybridization to other clones. No signal became visible in the tissues investigated when the *F. trinervia* class IV clone was used for hybridization.

A similar expression pattern was obtained with the *F. pringlei* clones (Fig. 5). The class I clone hybridizes preferentially with leaf RNA but also significantly with RNA from stems and roots. In contrast, the class II clone is expressed more strongly in roots than in the other tissues. Just as with the corresponding *F. trinervia* clone (class IV) no signals were obtained with the *F. pringlei* class III clone.

The *F. pringlei* class IV clone gave an unexpected result. Although hybridization analysis clearly revealed a relationship to the other PEPCase clones (see Tables 2, 3), Northern analysis identified only two 0.8 and 0.9 kb transcripts that are expressed in photosynthetic tissues.

Therefore, the identity of this clone remains unclear and requires further investigation.

As a control, Northern blots containing *F. trinervia* RNA were hybridized with a full-length cDNA encoding the 33 kDa protein of the water-splitting apparatus of photosystem II of *F. trinervia* (Höfer and Westhoff, in preparation). The corresponding gene should be expressed in all tissues active in photosynthesis. Figure 5 shows that mRNA encoding the 33 kDa protein (1.34 kb in size) can be detected in leaves, stems and young flowers, but not in root tissue.

Discussion

We are interested in understanding the molecular mechanisms that are responsible for the evolutionary changes in the development of the C_4 syndrome. The present study concentrates on phosphoenolpyruvate carboxylase, a key enzyme of C_4 metabolism.

It has been demonstrated that PEPCase is encoded by multigene families in several *Flaveria* species, representing various branches of the genus (Powell 1976). In contrast to *S. bicolor* (this paper) and *Z. mays* (Hudspeth and Grula 1989), in *F. trinervia* even the leaf-specific isoform is encoded by multiple genes. This finding is also supported by the isolation of different PEPCase cDNAs from a *F. trinervia* leaf-specific cDNA library (Poetsch et al., in preparation).

Unfortunately at present, no *F. trinervia* inbred lines are available which could have been used for the construction of cDNA and genomic libraries. Thus, it cannot be ruled out that the multiple bands observed in the Southern blots originate from alleles of a single locus in a heterogenous population. However, Southern blot analysis of *F. trinervia* genes coding for NADP-malic enzyme (D. Börsch, personal communication), pyruvate orthophosphate dikinase (Rosche and Westhoff, in preparation) or the 33 kDa protein of the water-splitting apparatus (Höfer and Westhoff, in preparation) revealed only single copies. These results suggest that the different genes for PEPCase truely represent different alleles in a $\pm$homogenous population. The multiplicity of leaf-specific PEPCase genes in *F. trinervia* compared to the monocot C_4 species investigated raises the question whether this is a common property of dicot C_4 plants or a unique exception.

Relationships of PEPCase genes

A large number of clones were isolated from genomic libraries of the *Flaveria* species *pringlei* (C_3) and *trinervia* (C_4 type) and were arranged into distinct classes according to their similarity to pcFtppc1-1 (Table 2). The hybridization data showed that some *F. pringlei* clones are more closely related to the leaf-specific C_4 cDNA than some *F. trinervia* clones. This was supported by the striking similarity of the 5′ and 3′ ends of the *F. pringlei* class I clone to the cDNA whereas several *F. trinervia* clones showed no or only weak reactions. Hence, the degree of heterogeneity in the PEPCase gene family within one species was found to be stronger than between the species.

To gain an insight into the relationship of the PEPCase genes within and between the two species, a cross-hybridization analysis was performed resulting in the construction of a phylogenetic tree (Fig. 4). The distance values were calculated from the hybridization data obtained at 60° C because all PEPCase genes showed a measurable rate of hybridization at this temperature (cf. Table 3). In a few cases, the results of the reciprocal approaches (the specular areas of the diagonal in Table 3) do not match. This could be due to the fact that the hybridization probes (see Fig. 2) have been selected by the convenience of restriction sites and thus may not necessarily represent homologous parts of the genes. Varying amounts and different locations of introns (Hudspeth and Grula 1989; Cushman et al. 1989) could also influence the melting behaviours of the hybrids.

Again, the hybridization analysis revealed a much higher degree of sequence divergence within than between the two species. The closest relative to most of the genes is not found within the same but in the other species. Thus, the class I and class II clones of both *F. trinervia* and *F. pringlei* contain pairs of closely related genes. A close relationship was also found for the class III clone of *F. pringlei* and the class IV clone of *F. trinervia*. No direct counterparts were found for the *F. trinervia* class III and *F. pringlei* class IV clones. It is possible that the related genes were not among the isolated clones or they may still be found among the clones which at present have not been characterized in detail (see Results).

Expression analysis

The regions of the clones of *F. trinervia* and *F. pringlei* expressed in leaves which show homology with the pcFtppc1-1 cDNA clone extend over 6.5–7.0 kb. Transcribed regions of similar size have been reported for *Z. mays* (Hudspeth and Grula 1989; Matsuoka and Minami 1989) and *Mesenbryanthemum crystallinum* (Cushman et al. 1989) indicating that a comparable number of introns (nine in both species) can be expected for *Flaveria* too. The other classes of PEPCase genes in both species showed no or only weak hybridization to the 3′ and 5′ probes. Thus, the extent of the transcribed regions of these genes cannot be estimated at present.

The length of *F. trinervia* leaf mRNA encoding PEPCase was estimated to be 3.3 kb, thus resembling that of *S. bicolor* (3.5 kb, Thomas et al. 1987) and *Z. mays* (3.4 kb, Harpster and Taylor 1986). Similar sizes were determined for the other PEPCase transcripts in *F. trinervia* as well as for *F. pringlei*.

The cross-hybridization data obtained at 75° C showed that the discrimination rate was sufficiently high to distinguish class-specific PEPCase transcripts by Northern blotting. In cases where a substantial cross-hybridization of genes occurs (e.g. *F. trinervia* class I

compared to class II) the distance data can be used to estimate the contribution of each particular gene. The results of the expression analysis together with the phylogenetic tree are summarized in Fig. 4.

Both the class I genes of *F. trinervia* and *F. pringlei* may be assigned as leaf-specific PEPCase genes. The class I gene of *F. trinervia* is expressed exclusively in leaf tissue while the corresponding *F. pringlei* gene also shows significant expression in roots and stems. Thus, the expression of the gene encoding the C_4 isoform of PEPCase is restricted to tissues exhibiting the C_4 pathway of photosynthesis. This is supported by hybridization analysis with a cDNA clone coding for the 33 kDa protein of the water-splitting apparatus of photosystem II. It could be expected that this gene is expressed in all photosynthetic tissues and without regard to either the C_3 or the C_4 mode of photosynthesis. Figure 5 shows that mRNA for the 33 kDa protein can be detected in leaves, stems and even young flowers, i.e. organs containing green tissues, but not in roots. There is no apparent difference in the abundance of transcripts in young and mature leaves. In contrast, the *F. trinervia* PEPCase class I clone, strictly correlated with C_4, revealed only half as much RNA in young as compared with expanded leaves. This may indicate that young leaves of *F. trinervia* do not fully express the C_4 cycle enzymes as has been found by Moore et al. (1986). The leaf-specific expression of the gene encoding the C_4 isoform of PEPCase in *F. trinervia* contrasts with *Z. mays*, where the expression of this gene is not restricted to leaves but occurs also in inner leaf sheaths, tassels and husks (Hudspeth and Grula 1989). These organs are fully or partially photosynthetic but do not possess Kranz anatomy (Hudspeth and Grula 1989). The different photosynthetic tissues in *F. trinervia* will offer an interesting system for the investigation of the development of the C_4 syndrome in a single plant species.

A correlation between the evolutionary relationship and expression patterns is also found in the class II genes. They are transcribed in all organs investigated. On the other hand the class III gene of *F. trinervia* is preferentially expressed in roots. The corresponding *F. pringlei* gene, if existing, needs to be identified. No expression can be detected for the *F. trinervia* class IV and *F. pringlei* class III genes, which form a third group of homologous genes. These clones may represent pseudogenes. However, in this case one could expect that their close similarity would not have been conserved during evolution of the two species. Therefore, the question is raised in which tissues, cells or developmental stages these genes are expressed. In this investigation the very young stages of plant development have not been analyzed for expression but it is known that in seedlings a PEPCase isozyme is expressed to ensure the refixation of CO_2 to avoid a loss of carbon (Latzko and Kelly 1983).

The function of the *F. pringlei* class IV gene remains unclear. No hybridization to a PEPCase mRNA-sized transcript occurred. Only a double signal of about 0.8 and 0.9 kb was obtained. It can be envisaged that these RNAs are transcribed from a gene which has a short homologous sequence in common with those that code for PEPCase. Alternatively, the hybridization probe used could contain a part of a PEPCase gene together with parts of a second nearby gene. The nature of this gene clearly requires further investigation.

Evolution of C_4 PEPCase

The multiplicity of PEPCase genes in both *Flaveria* species is intriguing as in maize only a small number of genes have been detected (Harpster and Taylor 1986; Hudspeth et al. 1986). In *F. trinervia* even the gene encoding the C_4 isoform of PEPCase appears to be represented by several genes. PEPCase genes which are closely related to the C_4 genes of *F. trinervia* and which are preferentially expressed in the leaf have been found in the C_3 species *F. pringlei*. According to morphological criteria (Powell 1976) all *Flaveria* C_4 species are more evolved than the C_3 species of this genus. Thus, *F. pringlei* probably reflects the ancient status in terms of evolution. This raises the question of how the C_4 PEPCase genes might have evolved from those genes which code for the leaf-expressed C_3 isozyme. It is highly unlikely that a gene to gene relationship can be expected, i.e. that each of the various C_4 genes has evolved independently from its C_3 ancestor. Hence, an evolutionary model should be favoured in which the primary event was the creation of the new, C_4-type gene followed by amplification to yield the present small C_4 PEPCase gene family in *F. trinervia*.

It will be of prime interest to identify those characteristics distinguishing a C_4-type promoter from its C_3 counterpart. Have distinct upstream regulatory elements in the C_3 gene been changed to result in the mesophyll cell-specific expression of the C_4 gene? Or have the promoters been completely revised to give this new pattern of expression? Sequence analysis of the corresponding regions should clarify these questions.

Acknowledgements. We are very grateful that the staff of the greenhouses, under the direction of Mr. W. Rogmann, provided such excellent C_4 plant material in our Northern latitudes. We are also indebted to Dr. K. Steinmüller, Dipl. Biol. D. Börsch and Dr. U. Santore for their helpful comments about this manuscript. This research was supported by the Deutsche Forschungsgemeinschaft (SFB 189) and in part by a grant from the Bundesministerium für Forschung und Technologie.

References

Andreo CS, Gonzales DH, Iglesias AA (1987) Higher plant phosphoenolpyruvate carboxylase. FEBS Lett 213:1–8

Aviv H, Leder P (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc Natl Acad Sci USA 69:1408–1412

Beltz GA, Jacobs KA, Eickbush TH, Cherbas PT, Kafatos FC (1983) Isolation of multigene families and determination of homologies by filter hybridization methods. Methods Enzymol 100:266–285

Benton WD, Davis RW (1977) Screening lambda gt recombinant clones by hybridization to single plaques in situ. Science 196:180–182

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81:1991–1995

Cockburn W (1983) Stomatal mechanism as the basis of the evolution of CAM and C_4 photosynthesis. Plant Cell Environ 6:275–279

Crétin C, Keryer E, Tagu D, Lepiniec L, Vidal J, Gadal P (1990) Complete cDNA sequence of *Sorghum* phosphoenolpyruvate carboxylase involved in C_4 photosynthesis. Nucleic Acids Res 18:658–658

Cushman JC, Meyer G, Michalowski CB, Schmitt JM, Bohnert HJ (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

Edwards GE, Ku MSB (1987) Biochemistry of C_3-C_4 intermediates. In: Hatch MD, Boardman NK (eds) The Biochemistry of Plants, vol 10. Academic Press, New York, pp 275–325

Edwards GE, Walker DA (1983) C_3, C_4: Mechanism, and Cellular and Environmental Regulation, of Photosynthesis. Blackwell Scientific Publications, Oxford London

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Harpster MH, Taylor WC (1986) Maize phosphoenolpyruvate carboxylase. Cloning and characterization of mRNAs encoding isozymic forms. J Biol Chem 261:6132–6136

Hatch MD (1987) C_4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. Biochim Biophys Acta 895:81–106

Hatch MD, Osmond CB (1976) Compartmentation and transport in C_4 photosynthesis. In: Stocking CR, Heber U (eds) Transport in Plants. III. Intracellular Interactions and Transport Processes. Encyclopedia of Plant Physiology, vol 3. Springer Verlag, Berlin, pp 144–184

Hudspeth RL, Grula JW (1989) Structure and expression of the maize gene encoding the phosphoenolpyruvate carboxylase isozyme involved in C_4 photosynthesis. Plant Mol Biol 12:579–589

Hudspeth RL, Glackin CA, Bonner J, Grula JW (1986) Genomic and cDNA clones for maize phosphoenolpyruvate carboxylase and pyruvate, orthophosphate dikinase: expression of different gene-family members in leaves and roots. Proc Natl Acad Sci USA 83:2884–2888

Latzko E, Kelly J (1983) The multi-faceted function of phosphoenolpyruvate carboxylase in C_3 plants. Physiol Veg 21:805–815

Martineau B, Taylor WC (1985) Photosynthetic gene expression and cellular differentiation in developing maize leaves. Plant Physiol 78:399–404

Matsuoka M, Minami E (1989) Complete structure of the gene for phosphoenolpyruvate carboxylase from maize. Eur J Biochem 181:593–598

McMaster GK, Carmichael GG (1977) Analysis of single and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc Natl Acad Sci USA 74:4835–4838

Monson RK, Moore Bd (1989) On the significance of C_3-C_4 intermediate photosynthesis to the evolution of C_4 photosynthesis. Plant Cell Environ 12:689–699

Monson RK, Edwards GE, Ku MSB (1984) C_3-C_4 intermediate photosynthesis in plants. Bioscience 34:563–574

Moore Bd, Cheng SH, Edwards GE (1986) The influence of leaf development on the expression of C_4 metabolism in *Flaveria trinervia*, a C_4 dicot. Plant Cell Physiol 27:1159–1167

Moore PD (1982) Evolution of photosynthetic pathways in flowering plants. Nature 295:647–648

O'Leary MH (1982) Phosphoenolpyruvate carboxylase: an enzymologist's view. Annu Rev Plant Physiol 33:297–315

Oswald A, Streubel M, Ljungberg U, Hermans J, Eskins K, Westhoff P (1990) Differential biogenesis of photosystem II in mesophyll and bundle sheath cells of NADP malic enzyme-type C_4 plants. A comparative protein and RNA analysis. Eur J Biochem 190:185–194

Powell AM (1976) Systematics of *Flaveria* (Flaveriinae-Asteraceae). Ann Missouri Bot Gard 65:590–636

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Seed B, Parker RC, Davidson N (1982) Representation of DNA sequences in recombinant DNA libraries prepared by restriction enzyme partial digestion. Gene 19:201–209

Smith HO, Birnstiel ML (1976) A simple method for DNA restriction site mapping. Nucleic Acids Res 3:2387–2398

Steinmüller K, Apel K (1986) A simple and efficient procedure for isolating plant chromatin which is suitable for studies of DNase I-sensitive domains and hypersensitive sites. Plant Mol Biol 7:87–94

Thomas M, Crétin C, Keryer E, Vidal J, Gadal P (1987) Photocontrol of *Sorghum* leaf phosphoenolpyruvate carboxylase. Characterization of messenger RNA and of photoreceptor. Plant Physiol 85:243–246

Ting IP, Osmond CB (1973) Multiple forms of plant phosphoenolpyruvate carboxylase associated with different metabolic pathways. Plant Physiol 51:448–453

Vidal J, Gadal P (1983) Influence of light on phosphoenol pyruvate carboxylase in *Sorghum* leaves. 1. Identification and properties of two isoforms. Physiol Plant 57:119–123

Westhoff P, Herrmann RG (1988) Complex RNA maturation in chloroplasts. The *psbB* operon from spinach. Eur J Biochem 171:551–564

Wyman AR, Wertman KF (1987) Host strains that alleviate and underrepresentation of specific sequences: overview. Methods Enzymol 152:173–180

Communicated by R.G. Herrmann