Plant Molecular Biology **26**: 1961–1973, 1994.

Chloroplast and cytosolic triosephosphate isomerases from spinach: purification, microsequencing and cDNA cloning of the chloroplast enzyme

Katrin Henze [1], Claus Schnarrenberger [2], Josef Kellermann [3] and William Martin [1,*]
[1] *Institut für Genetik, Technische Universität Braunschweig, Spielmannstrasse 7, D-38106 Braunschweig, FRG (* author for correspondence);* [2] *Institut für Pflanzenphysiologie und Mikrobiologie, Freie Universität Berlin, Königin-Luise-Strasse 12–16a, D-14195 Berlin, Germany;* [3] *Max Planck Institut für Biochemie, Am Klopferspitz, D-82152 Martinsried, Germany*

Received 27 September 1994; accepted 25 October 1994

Key words: triosephosphate isomerase, isoenzyme purification, Calvin cycle enzyme, cDNA cloning, gene phylogeny

Abstract

Chloroplast and cytosolic triosephosphate isomerases from spinach were separated and purified to homogeneity. Both enzymes were partially sequenced by Edman degradation. Using degenerate primers designed against the amino acid sequences, a homologous probe for the chloroplast enzyme was amplified and used to isolate several full-size cDNA clones. Chloroplast triosephosphate isomerase is encoded by a single gene in spinach. Analysis of the chloroplast cDNA sequence in the context of its homologues from eukaryotes and eubacteria reveals that the gene arose through duplication of its pre-existing nuclear counterpart for the cytosolic enzyme during plant evolution.

Abbreviations: TPI, triosephosphate isomerase; PEG, polyethylene glycol; cp, plastid; c, cytosolic; SDS, sodium dodecyl sulphate; PAGE, polyacrylamide gel electrophoresis; PVP, polyvinylpyrrolidone; PCR, polymerase chain reaction; PGK, 3-phosphoglycerate kinase.

Introduction

Triosephosphate isomerase (TPI) (EC 5.3.1.1) catalyses the interconversion of dihydroxyacetone phosphate and D-glyceraldehyde-3-phosphate. In higher plants and *Euglena*, TPI exists as two distinct isoforms, a glycolytic enzyme located in the cytosol and a Calvin cycle enzyme compartmentalized in chloroplasts [15, 37]. By contrast, some green algae possess only the chloroplast isoenzyme and lack the cytosolic form altogether [51, 52]. Like its counterparts in other eukaryotes, cytosolic TPI (cTPI) of higher plants is a homodimer composed of 27 kDa, subunits [25], molecular sequences for which have been reported from several sources, including maize [30], rice [64] and *Arabidopsis* [54]. Chloroplast TPI (cpTPI) has also been studied at the biochemical level and the enzyme has been purified from rye, spinach and lettuce [25, 26, 43]. But, in contrast

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number L36387.

to its cytosolic counterpart, molecular probes for the chloroplast enzyme which would permit investigation of expression, structure and evolution of the gene(s) have not been reported.

The complete sequences of several plastid genomes have revealed that, with the exceptions of *rbcL* in chlorophytes [62] and *rbcL*, *rbcS* and phosphoglycerate mutase in rhodophytes [45], all enzymes of sugar phosphate metabolism in plastids are encoded in nuclear DNA and imported into the organelle with the aid of a transit peptide. In general terms, genes for plastid enzymes may have arisen during evolution through duplication of pre-existing nuclear genes, as in the case of glutamine synthase [58, 52], through transfer of the gene from the chloroplast genome to the nucleus as in the case of chloroplast GAPDH [31, 20] and *rbcS* [63, 35] or through gene transfer emanating from mitochondrial DNA, as may be the case for *rpl21* [32]. Regardless of their origin, contemporary nuclear genes for enzymes of plastid sugar phosphate metabolism are regulated by the eukaryotic transcriptional apparatus. Study of these genes can provide valuable insights into the evolution of the isoenzymes and their regulation in plants.

We have separated the chloroplast and cytosolic triosephosphate isomerases from spinach leaves and have purified them to apparent homogeneity. Partial amino acid sequences were determined from both enzymes. Using PCR primers designed against oligopeptide sequences, we amplified specific probes for the chloroplast enzyme and isolated several independent full-size cDNA clones for cpTPI from a spinach cDNA library. Here we show that the chloroplast enzyme is encoded by a single-copy gene, *TpiP1*, which arose during plant evolution through duplication of a pre-existing nuclear gene for the cytosolic enzyme.

Material and methods

Plant material

Leaves of *Spinacia oleracea* L. cv. Monnopa for enzyme isolation were grown in the field. Spinach seedlings for mRNA isolation were grown for one week under a 14–10 h light/dark regime at 25 °C, harvested in liquid nitrogen and stored at −80 °C.

Enzyme separation

All procedures were carried out at 0–4 °C. A total of 500 g spinach leaves were washed and their midribs removed. The leaves were homogenized in a Waring blender in 600 ml of buffer A (10 mM potassium phosphate pH 8.6, 20 mM 2-mercaptoethanol, 1 mM leupeptin; 0.5 mM PMSF could be substituted for leupeptin). The homogenate was filtered through cheesecloth and centrifuged for 45 min at 15000 × *g*. The supernatant was ad justed to pH 5.5 with 85% phosphoric acid, gently stirred for 5 min on ice and centrifuged for 20 min at 15000 × *g*. The pH of the supernatant was adjusted to pH 8.6 with 5 M KOH. Ammonium sulphate was added to 40% saturation while the pH was held at 8.6 by addition of 5 M KOH. After 15 min of stirring, the mixture was centrifuged for 30 min at 15000 × *g*. Proteins of the supernatant were precipitated with 80% ammonium sulphate as described above. The pellet was dissolved in a small volume of buffer A. The solution was dialysed overnight against 2 l of buffer A. The dialysate was centrifuged for 20 min at 23000 × *g*, diluted with distilled water until its conductivity was less than 2 mS/cm and applied to a 2 cm × 10 cm DEAE-Fractogel 650S (Merck) column equilibrated with buffer A. The column was washed with two volumes of buffer A and proteins were eluted with a 300 ml gradient of 0 to 0.2 M KCl gradient in buffer A. Fractions of 1 ml were collected. Fractions with TPI activity were pooled, diluted with water to a conductivity of less than 3 mS/cm, and applied at 3 ml/min to a Synchropak AX 300 column (250 mm × 8 mm, Molnar, Berlin) equilibrated with buffer B (10 mM potassium phosphate pH 7.5, 20 mM 2-mercaptoethanol). The column was washed with buffer B until the extinction at 280 nm fell to zero. Proteins were eluted with a 30 min gradient of 0.05–0.15 M KCl in buffer B. One ml fractions were collected and tested for TPI and PGK activity. Fractions con-

taining chloroplast TPI (cpTPI) and cytosolic TPI (cTPI) activity, respectively, were separately pooled and dialysed overnight against 2 l of buffer B.

Isoenzyme purification

For the isolation of chloroplast triosephosphate isomerase (cpTPI), the above enzyme solution was rechromatographed on the AX 300 column by a 0.06–0.15 M KCl gradient in buffer B. The cpTPI fractions were pooled, concentrated to 1–2 ml by dialysis against solid PEG 4,000 and dialysed overnight against 2 l of buffer B. Proteins of this solution were applied at 1 ml/min to a 300 mm × 7.5 mm Biosil TSK 250 column (Bio-Rad) equilibrated with buffer B. Elution was monitored by UV absorption. Fractions of 250 µl were collected and those containing TPI activity were pooled, concentrated by dialysis against PEG 20,000 and dialysed overnight against 2 l of buffer B.

For the isolation of cytosolic triosephosphate isomerase (cTPI), the enzyme after the first AX 300 column was rechromatographed as above on the AX 300 column using a 0.05–0.1 M KCl gradient in buffer B. Fractions with cTPI activity were pooled, concentrated by dialysis against PEG 20000, dialysed overnight and chromatographed on Biosil TSK 250 as described for cpTPI (PGK coeluted with cTPI from Biosil TSK 250). Fractions with TPI activity were pooled, concentrated, and dialysed against 2 l of buffer C (10 mM potassium phosphate, 20 mM 2-mercaptoethanol pH 6.5). In a final step, the solution was re-chromatographed as above on the AX 300 column using a 0.05–0.15 M KCl gradient in buffer C. Fractions containing cTPI activity, but devoid of PGK activity, were pooled, concentrated, and dialysed against buffer C.

Enzyme assays

Enzyme assays were performed at 20 °C and monitored spectrophotometrically at 334 nm in a total volume of 1 ml. One unit of activity corresponds to 1 µmol of NADH oxidized per minute. TPI assays contained 100 mM triethanolamine pH 7.6, 5 mM glyceraldehyde-3-phosphate, 200 µM NADH, and 2.6 units glyceraldehyde-3-phosphate dehydrogenase [3]. PGK assays contained 80 mM Tricin pH 7.5, 4.5 mM $MgCl_2$, 5 mM dithiothreitol, 5 mM adenosine triphosphate, 200 µM NADH, 10 mM 3-phosphoglycerate and 0.5 units glyceraldehyde-3-phosphate dehydrogenase [4]. KCl concentrations in eluted fractions were measured by conductivity.

Protein sequencing

Purified proteins were digested with endoproteinase LysC as described by Eckerskorn and Lottspeich [10]. Peptides were separated on a 2 mm × 125 mm Supersher 60 RP select B column (Merck) at a flow rate of 200 µl/min in a 1%/min gradient of 0.1% (v/v) trifluoroacetic acid in water to 0.1% (v/v) trifluoroacetic acid in acetonitrile. Peptides were sequenced by amino-terminal degradation [11] in an automatic Porton 3600 Sequencer (Beckman) and amino acids were identified in a Mikrobore HPLC System Gold (Beckman).

Nucleic acid isolation

One-week-old spinach seedlings (80 g) were ground in liquid nitrogen and vigorously shaken for 10 min in 100 ml of 50 mM Tris-HCl, 100 mM NaCl, 10 mM EDTA, 2% (w/v) sodium dodecyl sulfate (SDS), 200 µg/ml proteinase K (Merck) pH 9.0, followed by two extractions of one volume each with 80% (w/v) phenol/chloroform/isoamyl alcohol (25:24:1, v/v) and two further extractions of one volume each with chloroform/isoamyl alcohol (24:1, v/v). The aqueous phase of the final extraction was precipitated with 2 volumes of ethanol, nucleic acids were precipitated for 1 h at −20 °C and collected by 30 min centrifugation at 10000 rpm. The pellet was dissolved in 10 ml of water containing 10 µg/ml protein-

ase K, 2.5 ml of 10 M LiCl were added, RNA was precipitated for 3 h at 0 °C and collected by centrifugation for 30 min at 15000 rpm and 4 °C.

DNA in the supernatant of the LiCl precipitation was collected by isopropanol precipitation and was further purified by two rounds of CsCl centrifugation. RNA in the pellet of the LiCl precipitation was dissolved in 50 ml of binding buffer (400 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.2% (w/v) SDS, 10 μg/ml proteinase K) and bound for 30 min to 0.5g oligo(dT) cellulose (Type 7, Pharmacia). Cellulose was washed in batch four times with 25 ml of binding buffer and twice with 25 ml of column buffer (100 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.2% (w/v) SDS, 10 μg/ml proteinase K), transferred to a 1 cm × 10 cm column and washed with 100 ml of column buffer. Poly(A)$^+$ mRNA was eluted at 50 °C with 10 ml of 10 mM Tris-HCl pH 7.5, 10 μg/ml proteinase K and purified again by chromatography on oligo(dT) cellulose. The final eluate was phenolized and precipitated with ethanol.

cDNA synthesis and cloning

cDNA was synthesized with the Pharmacia kit with modifications as described [32]. From 5 μg of mRNA, 2 μg of adaptor-ligated dscDNA was obtained which was used for cloning and for PCR reactions. For cloning, 10 ng of cDNA was ligated into 200 ng of *Eco* RI-digested λnm1149 [38] and packaged with extracts prepared as described [28]. Recombinants were plated on *Escherichia coli* POP13 and screened by plaque hybridization on 82 mm nitrocellulose filters (Gelman Sciences) at 68 °C in 3 × SSPE, 0.1% (w/v) SDS, 0.02% (w/v) PVP, 0.02% (w/v) Ficoll-400 and 50 μg/ml calf thymus DNA containing 10 ng/ml of hybridization probe. The hybridization probe was the 0.45 kb *Bam* HI-*Hin*d III fragment of pPCRcpT1 random-labelled to 1.6×10^7 cpm/μg [13] with α-^{32}P dCTP. Filters were washed for 60 min at 68 °C in 2 × SSPE, 0.1% (w/v) SDS and autoradiographed overnight on XAR films. *Not* I inserts of four positively hybridizing cDNAs were subcloned into pBluescriptSK plasmids (Stratagene). Their terminal sequences were determined by the dideoxy method [49].

Isolation of hybridization probe for cpTPI

The degenerate PCR primers 5′-GTNGGNG-GNAAYTGGAA-3′ and 5′-GCCCANACNG-GYTCRTA-3′ were constructed against the peptide motifs VGGNWK and YEPVWA, respectively, obtained from oligopeptide sequencing. These were used to amplify cpTPI-specific fragments from uncloned cDNA. The PCR reaction contained 52 mM Tris-HCl pH 8.8, 5.2 mM $MgCl_2$, 5 μM of each dNTP, 1.7 μM of each primer, 0.003 U/μl *Taq* polymerase (USB) and 3.2 ng/μl double-stranded spinach cDNA in a total volume of 25 μl. 35 cycles of 93 °C/50 °C/72 °C were performed. The 0.45 kb amplification product was used as a substrate for re-amplification under the same conditions, the product of which was subcloned into *Eco* RV-cut pBluescriptSK + (Stratagene) to yield the plasmid pPCRcpT1. The sequence of pPCRcpT1 was determined by the dideoxy method [49] to confirm the identity of the amplified product by comparison to the peptide sequences from the isolated enzyme. The 0.45 kb *Bam* HI-*Hin*d III fragment of pPCRcpT1 was isolated by gel electrophoresis and DE-52 ion-exchange chromatography as a hybridization probe for cDNA screening. Positives from this screening were selected on the basis of insert size. The largest inserts were subcloned into pBluescript plasmids and sequenced.

Genomic Southern blots

A 25 μg aliquot of spinach DNA, digested with the appropriate enzyme, was loaded per lane and electrophoresed at 3 V/cm on an 0.7% agarose gel in 40 mM Tris-acetate, 1 mM EDTA pH 8.0. DNA was transferred to Hybond-N (Amersham) by capillary blot and UV-cross-linked according to the manufacturer's protocol. Filters were prehybridized at 68 °C in 6 × SSPE, 0.1% (w/v)

SDS, 0.02% (w/v) PVP, 0.02% (w/v) Ficoll-400 containing 50 μg/ml calf thymus DNA. Hybridization was carried out at 68 °C in 3× SSPE, 0.1% (w/v) SDS, 0.02% (w/v) PVP, 0.02% (w/v) Ficoll-400, 50 μg/ml calf thymus DNA containing 10 ng/ml of hybridization probe. The hybridization probe was the 335 bp *Hind* III fragment from the 5′ end of pSOTA2 random-labelled to 1.5×10^8 cpm/μg [13] with α-^{32}P dCTP. Filters were washed for 60 min at 68 °C in 2× SSPE, 0.1% (w/v) SDS and autoradiographed for seven days on XAR films.

Other molecular methods

Nested deletions for sequencing were prepared with exonuclease III according to Henikoff [17] using mung bean (Pharmacia) instead of S1 nuclease. Exonuclease III-resistant ends were generated by Klenow fill-in of 5′-protruding termini for 30 min at 37 °C in the presence of 20 μM of each αS-dNTPs, unincorporated thiophosphates were removed by gel filtration. DNA sequences were determined from alkaline-denatured plasmids [7] by the dideoxy chain termination method with modified T_7 DNA polymerase [57] using either radioactive [49] or flourescence [2] detection. Other DNA manipulations were performed as described [28]. Oligonucleotides were chemically synthesized with the Pharmacia machine.

Sequence analysis

Standard sequence analysis was performed with the GCG package [9]. Sequences were aligned with CLUSTAL V [18] and by eye with LINEUP. After exclusion of regions of uncertain positional homology, 227 positions, 177 of which were polymorphic, remained in the final alignment which was used for phylogenetic inference. Distance between sequences was measured as numbers of amino acid substitutions per site corrected for multiple substitutions by assuming a gamma distribution for the variability of substitution rate across positions [40]. For this, a neighbour-joining tree [48] was constructed using the proportion of amino acid differences between TPI sequences, from which the gamma parameter was estimated. The value of *a* thus determined (2.295) was used to estimate numbers of substitutions per site between sequences using the gamma correction. The resulting distance matrix yielded the final neighbor joining tree (calculated with a program kindly provided by M. Nei). The reliability of branches was estimated by bootstrapping using the same gamma parameter.

Results

Enzyme purification

The chloroplast and cytosolic isoenzymes of trisoephosphate isomerase from spinach leaves were purified by acid precipitation, $(NH_4)_2SO_4$ precipitation and conventional ion-exchange chromatography on DEAE-Fractogel (see Table 1). The isoenzymes were effectively separated by HPLC using ion-exchange chromatography on Synchropak AX 300 (Fig. 1a). Phosphoglycerate kinase activity tends to copurify with cTPI through most steps, but is easily separated

Table 1. Purification scheme for chloroplast and cytosolic TIM from spinach leaves.

	Total activity (units)	Specific activity (units/mg)	Recovery (%)	Purification (fold)
Total TIM				
Crude extract	4600	1.4	100.0	0
Acid precipitation	3840	5.0	83.4	3.4
DEAE-Fractogel	2560	7.0	55.6	5.1
Chloroplast TIM				
AX 300, pH 7.5	1000	119.0	21.7	86.3
AX 300, pH 7.5	1160	240.2	25.0	174.0
BioSil TSK 250	200	333.3	4.3	241.5
Cytosolic TIM				
AX 300, pH 7.5	1080	120.9	23.6	87.6
AX 300, pH 7.5	560	172.3	12.2	124.9
AX 300, pH 7.5	600	340.9	13.0	247.0
AX 300, pH 6.5	85	400.0	1.8	290.0

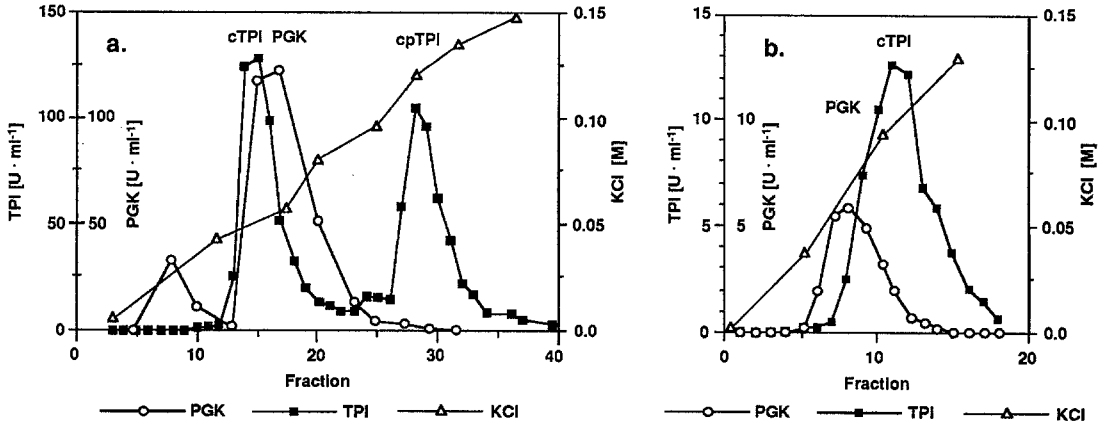


Fig. 1. Elution profiles of TPI. a. Elution profile from Synchropak AX 300 eluted at pH 7.5 showing separation of cytosolic and chloroplast TPI isoenzymes. b. Elution of cTPI and PGK of from AX 300 at pH 6.5. cTPI, cytosolic TPI; cpTPI, chloroplast TPI; PGK, 3-phosphoglycerate kinase.

from the TPI enzyme of isolated chloroplasts [23]. Sufficient separation to obtain a pure cTPI fraction was achieved by lowering the pH to 6.5 in the final HPLC step (Fig. 1b). The chloroplast enzyme was purified 240-fold and the cytosolic enzyme 290-fold. These purification factors are similar to those previously reported for the cTPI and cpTPI from rye [25] and lettuce [43].

The final preparations of both enzymes were electrophoretically homogeneous (Fig. 2). The molecular mass of the cytosolic enzyme under denaturing conditions is 28 kDa, the chloroplast enzyme has a slightly lower mass of 27 kDa. Roughly 300 μg of cpTPI and 100 μg of cTPI were obtained from 500 g of mature leaves. Aliquots of these preparations were subjected to proteolysis with endopeptidase LysC. The resulting fragments were separated by HPLC chromatography and sequenced by the Edman degradation method. For the chloroplast enzyme, five peptide sequences were determined: FFVGGNWK, GGAFTGEISVEQLK, AFADALPSWDNVVVAYEPVWA, VASPDQAQEVHVAVRDWLK and GPEFATIVNSVTA. For the cytosolic enzyme, two peptide sequences were determined: FFVGGNYK and GGAFTGVESAEMLADL.

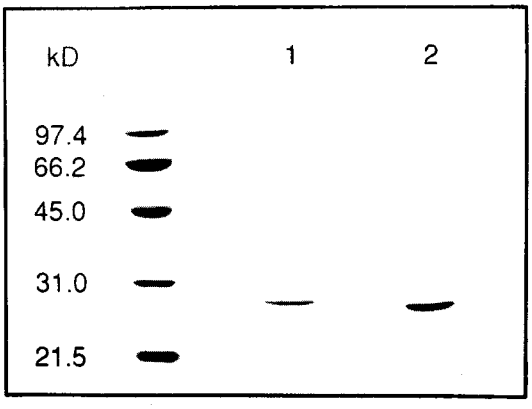
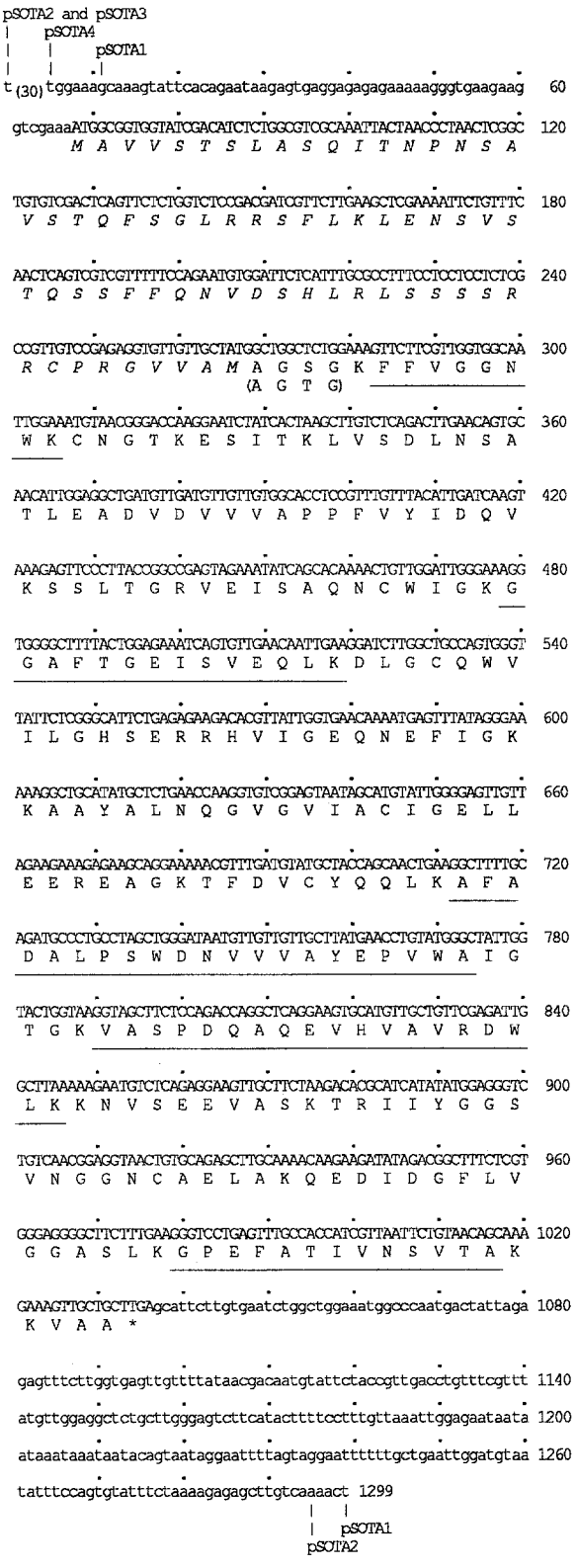


Fig. 2. Coomassie blue-stained SDS-polyacrylamide gel of purified chloroplast and cytosolic triosephosphate isomerases from spinach leaves. Ca. 1 μg of protein was loaded in lanes 1 (cytosolic TPI) and 2 (chloroplast TPI). Sizes of molecular weight standards are indicated.

cDNA clones for chloroplast TPI

Using the protein sequence data, a PCR amplification product specific to cpTIM was obtained and used as a hybridization probe to isolate several full-size cDNA clones for cpTIM from the unamplified library. From 50000 recombinants four positives were obtained, all of which encoded cpTPI. This indicates an abundance for chloroplast TPI transcripts of roughly one in 10000 in Poly(A)$^+$ mRNA, which is about tenfold



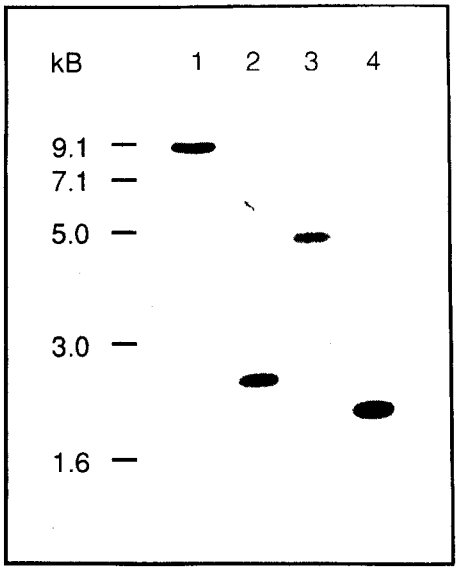


Fig. 4. Southern blot of genomic DNA from spinach probed with the 335 bp *Hin*d III fragment of pSOTA2 which encompasses the region encoding the transit peptide and the first 23 amino acids of the mature subunit. 25 μg of spinach DNA was loaded per lane. Lane 1, *Bam* HI; lane 2, *Nco* I; lane 3, *Nde* I; lane 4 *Hin*d III. Molecular weight standards are indicated. The film was exposed for 7 days at −80 °C.

lower than transcript abundance for chloroplast GAPDH in green tissues of higher plants [19]. The inserts of pSOTA1-4 were all about 1.3 kb long.

pSOTA2 was sequenced on both strands (Fig. 3). pSOTA1, pSOTA3 and pSOTA4 were sequenced from their 5′ ends and shown to have no nucleotide polymorphisms relative to pSOTA2 over about 250 bases, indicating that these are independent cDNAs which probably emanate from the same gene. pSOTA2 is 1295 bases long. It encodes an open reading frame of 255 codons which contains all five oligopeptides sequenced from the protein (underlined in Fig. 3) unambigu-

Fig. 3. Sequence of cDNA clones for chloroplast TPI from spinach. Terminal nucleotides of clones pSOTA1–4 are indicated by '|'. The amino terminus of the chloroplast TPI enzyme from lettuce [43] is shown in parentheses, the transit peptide is shown in italics. Regions sequenced from the isolated protein are underlined. Non-coding regions are given in lower case.

ously identifying the encoded product as chloroplast triosephosphate isomerase. The 5′ and 3′ non-coding regions are 100 bp and 257 bp long, respectively. Interestingly, the 5′ ends of pSOTA2 and pSOTA3 carry a stretch of roughly 30 thymidine residues, corresponding to 3′ poly(A) on the opposite strand. This unusual feature is not a cloning artefact since it is found in two independent cDNA clones but its significance, if any, is obscure. The amino-terminal region of 67 amino acids is serine-rich and appears to encode the transit peptide. The amino acid motif 'AGSG' at codon 68 of pSOTA2 is similar to the amino terminal sequence 'AGTG' which was determined for the chloroplast TPI protein from lettuce [43] and very probably represents the cleavage site for

Fig. 5. Alignment of triosephosphate isomerase amino acid sequences. Gaps are indicated by periods, strictly conserved residues in the alignment are indicated by '•'. The arrow at the amino terminus of spinach chloroplast TPI indicates the processing site. Alpha-helices and β-pleated sheets found in the 0.18 nm crystal structure of the *T. brucei* enzyme [61] are indicated as 'aaaa' and 'bbbb', respectively. Numbering is arbitrary and does not apply to the residues in the crystal structure. Sources of sequences are: *Arabidopsis thaliana* [54]; *Coptis japonica* [39]; maize [30]; rice [64]; human [29]; *Culex tarsalis* [59]; *Saccharomyces cerevisiae* [1]; *Trypanosoma brucei* [56]; *Giardia lamblia* [36]; *Thermotoga maritima* (accession number L27492); *Vibrio* sp. strain ANT-300 (accession number L27493); *Bacillus megaterium* [50]; *Corynebacterium glutamicum* [12]; *Escherichia coli* [42].

the transit peptide. The calculated molecular mass of the mature subunit expected for this cleavage site is 27.2 kDa, which is very close to the value of 27 kDa determined for the purified spinach enzyme by SDS-PAGE.

Chloroplast TPI is encoded by a single gene in spinach

When total spinach DNA was probed with the 335 bp *Hind* III fragment from the 5′ end of pSOTA2, only one hybridizing band was detected in *Bam* HI, *Nco* I, *Nde* I or *Hind* III digests, indicating that only one gene for chloroplast TPI exists in the genome of this inbred line (Fig. 4). This is congruent with our findings in that neither hints for the existence of cpTPI isozymes in enzyme purification nor ambiguities in protein sequencing were detected.

Comparison to other TPI sequences

Molecular sequences for glycolytic TPI have been characterized from plants, animals, fungi, protists and several bacteria. We integrated the oligopeptide sequences of spinach cytosolic TPI and the complete derived sequence from pSOTA2 for chloroplast TPI into an alignment of several TPI sequences from the data base (GenBank Release 84). The partial peptide sequence of cTPI from spinach possesses the unambiguous signature motif 'SAEML' characteristic of other plant cytosolic TPI enzymes (Fig. 5). Cytosolic TPI from spinach is, thus, clearly the product of a nuclear gene which is orthologous to those for other plant cTPI enzymes. Chloroplast TPI from spinach, on the other hand, does not possess this motif but is collinear with and very similar to other TPI proteins. The derived amino acid sequence of cpTPI shares an average of 60% iden-

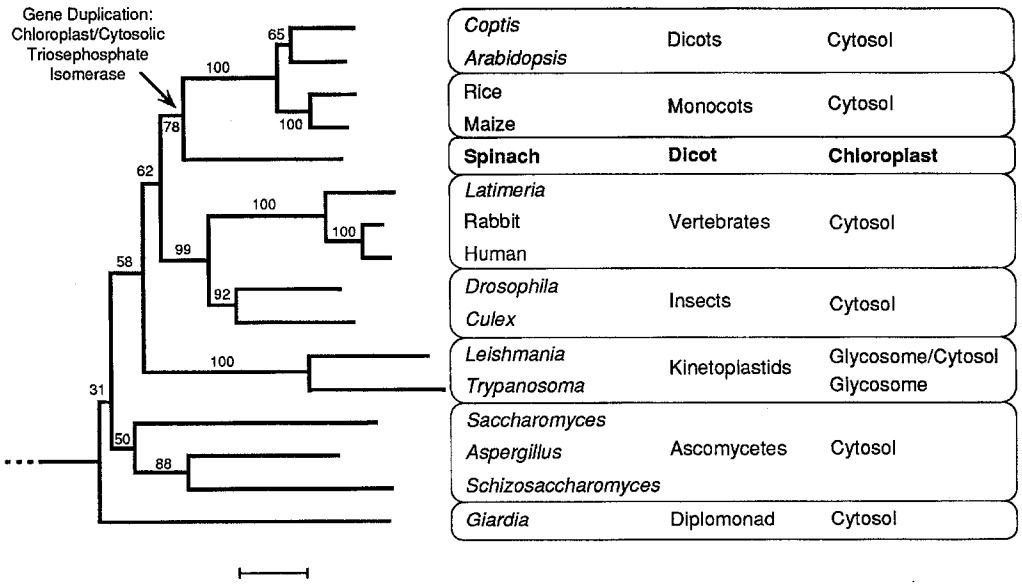


Fig. 6. Gene tree constructed by the neighbour-joining method for triosephosphate isomerase sequences. Numbers above branches indicate the number of times the respective branch was found out of 100 bootstrap replicates. The scale bar indicate 0.1 substitution per site. The gamma parameter [40] was estimated on the basis of a neighbour-joining tree constructed from proportions of amino acid differences, the branching order of which was not different from that shown in the figure. Higher taxa and enzyme compartments are indicated. The tree was rooted to the TPI sequences from *E. coli* and *T. maritima* as outgroups (dotted line). Sources of sequences are as in Fig. 5 in addition to *Latimeria chalumnae* [22], *Leishmania mexicana* [21], *Schizosaccharomyces pombe* [47]; *Aspergillus nidulans* [34]; *Drosophila melanogaster* [53] and rabbit [8].

tity with cytosolic TPI from plants, 58% with animal, 48% with fungal and 42% with bacterial sequences. We note that the *Coptis japonica* sequence shown was originally described as the sequence for (S)-tetrahydroberberine oxidase [39], although its similarity to cTPI sequences suggests that the gene product is triosephosphate isomerase.

A gene phylogeny was inferred for eukaryotic TPI sequences (Fig. 6). The sequences of cytosolic TPI from two monocots branch together as does the *Arabidopsis* TPI sequence with that of *Coptis* (Ranunculaceae). The sequence of chloroplast TPI from spinach shares a common branch with its homologues from the plant cytosol. Although this branch receives only moderate support from neighbour-joining bootstrap analysis (78/100 replicates), the results in Fig. 6 indicate that the gene for cpTPI from spinach arose through duplication of the preexisting nuclear gene for cTPI during eukaryotic evolution. Remaining sequences in the TPI phylogeny correlate well with expectation, since vertebrate, invertebrate, ascomycete and kinetoplastid TPI sequences share common branches, respectively, and since the sequence from the amitochondriate protist *Giardia lamblia* assumes a basal position in eukaryotic phylogeny. The fact that fungi branch below kinetoplastids does not conflict with the results of rRNA [55] and other protein phylogenies [16], since the branching orders for lower eukaryotes in this TPI tree are not robust.

The only non-photosynthetic organisms known to possess differentially compartmentalized TPI isozymes are kinetoplastids. In these protists, TPI may either be localized entirely in glycolytic microbodies (glycosomes), as in the case of *T. brucei* [56], or may occur as glycosomal/cytosolic isoforms, as in the case of *Leishmania*. Interestingly, glycosomal and cytosolic triosephosphate isomerases of *Leishmania* are the product of one and the same gene [21].

Discussion

Enzymes of primary carbohydrate metabolism are ubiquitous among eukaryotes, eubacteria and archaebacteria. They also belong to the most conservatively evolving proteins known and the crystal structures of many have been determined [14]. In plants, carbohydrate metabolism occurs both in the cytosol and in chloroplasts; those reactions common to glycolysis, gluconeogenesis, the oxidative pentose phosphate pathway and the Calvin cycle are catalysed by distinct nuclear-encoded isoenzymes unique to each. Molecular probes exist for a number of nuclear-encoded Calvin-cycle enzymes of higher plants. These include the small subunit of Rubisco [24], phosphoglycerate kinase [5], subunit A and subunit B of chloroplast glyceraldehyde-3-phosphate dehydrogenase [6], fructose-1,6-bisphosphatase [44], fructose-1,6-bisphosphate aldolase [41], seduheptulose-1,7-bisphosphatase [44] and phosphoribulokinase [46]. Chloroplast triosephosphate isomerase is one of a few remaining Calvin cycle enzymes for which molecular probes have been lacking.

Previous enzymatic and immunological studies of plant TPI had shown that the cytosolic and chloroplast enzymes are of slightly different molecular weight, have similar catalytic properties, are both encoded in nuclear DNA [15, 27] but are immunologically distinct [25, 26, 43]. These findings, in addition to the lack of cross-reactivity between bacterial TPI and antiserum raised against the chloroplast enzyme, led Kurzok and Feierabend [26] to postulate '... that the gene for the chloroplast enzyme arose from a duplication of the ancestral nuclear gene ...'. Our sequence results clearly bear out their prediction and furthermore reveal that the gene duplication event for the origin of *TpiP1* took place within plant genomes, since spinach cpTPI is more similar to plant cTPI than it is to other eukaryotic homologues (Fig. 6). Thus, distinct genes for cpTPI and cTPI arose more recently in evolution than those which brough forth other isoenzymes of sugar phosphate metabolism, such as chloroplast and cytosolic aldolases, which arose through duplication in the genomes of primitive eukaryotes, long prior to the separation of plant and animal lineages [52], or chloroplast and cytosolic GAPDH, which arose through gene duplication in eubacterial genomes [31].

As a general model for the origin of nuclear genes for chloroplast/cytosol isoenzymes, the gene transfer corollary to endosymbiotic theory [60] posits a eukaryotic origin of genes for cytosolic, and a eubacterial origin of genes for chloroplast isoenzymes. Yet, none of the nuclear-encoded glycolytic/Calvin cycle enzyme pairs studied to date have provided evidence in support of such a scenario. The isoenzyme genes are related through duplication, rather than through prokaryote/eukaryote divergence. Further studies of cpTPI from other higher plants, algae, photosynthetic protists and its homologues in cyanobacteria are, however, warranted. It is possible that either plastid-bearing eukaryotes which branched off very early in eukaryotic evolution (i.e. Euglenophytes), or those which possess complex plastids (i.e. Cryptophytes or Chlorarachniophytes) [33] might have obtained genes for chloroplast TPI through endosymbiotic gene transfer from organellar genomes.

Acknowledgements

We thank Renate Grübnau and Christiane Köhler for excellent technical assistance and the Gesellschaft für Biotechnologische Forschung (Braunschweig) for the generous use of their computer facilities. This work was supported by a joint grant from the Deutsche Forschungsgemeinschaft to W.M. (Ma 1426/3-1) and C.S. (Schn 28/21-1).

References

1. Alber T, Kawasaki G: Nucleotide sequence of the triose phosphate isomerase gene of *Saccharomyces cerevisiae*. J Mol Appl Genet 1: 419–434 (1982).
2. Ansorge W, Sproat BS, Stegemann J, Schwager C: A non-radioactive automated method for DNA sequence determination. J Biochem Biophys Meth 13: 315–323 (1986).
3. Beisenherz G: Triosephosphate isomerase from calf muscle. Meth Enzymol 1: 387–391 (1955).
4. Bergmeyer HU: Methods of Enzymatic Analysis, vol. 2. Verlag Chemie, Weinheim (1983).
5. Bertsch U, Schlicher TB, Schröder I, Soll J: Sequence of mature phosphoglycerate kinase from spinach chloroplasts. Plant Physiol 103: 1449–1450 (1993).
6. Brinkmann H, Cerff R, Salomon M, Soll J: Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GapA and GapB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach. Plant Mol Biol 13: 81–94 (1989).
7. Chen EY, Seeburg PH: Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. DNA 4: 165–170 (1985).
8. Corran PH, Waley SG: The amino acid sequence of rabbit muscle triose phosphate isomerase. Biochem J 145: 335–344 (1975).
9. Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programms for the VAX. Nucl Acids Res 12: 387–395 (1984).
10. Eckerskorn C, Lottspeich F: Internal amino acid sequence analysis of proteins separated by gel electrophoresis after tryptic digestion in polyacrylamid matrix. Chromatographia 28: 92–94 (1989).
11. Edman P, Begg G: A protein sequenator. Eur J Biochem 1: 80–91 (1967).
12. Eikmanns BJ: Identification, sequence analysis and expression of a *Corynebacterium glutamicum* gene cluster encoding the three glycolytic enzymes glyceraldehydephosphate dehydrogenase, 3-phosphoglycerate kinase and triosephosphate isomerase. J Bact 174: 6076–6086 (1992).
13. Feinberg AP, Vogelstein B: A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 137: 266–267 (1984).
14. Fothergill-Gilmore LA, Michels PAM: Evolution of glycolysis. Progr Biophys Mol Biol 59: 105–238 (1993).
15. Gottlieb LD: Conservation and duplication of isozymes in plants. Science 216: 373–380 (1982).
16. Hasegawa M, Hashimoto T, Adachi J, Iwabe N, Miyata T: Early branchings in the evolution of eukaryotes: ancient divergence of *Entamoeba* that lacks mitochondria revealed by protein sequence data. J Mol Evol 36: 380–388 (1993).
17. Henikoff S: Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28: 351–359 (1984).
18. Higgins DH, Sharp PM: CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73: 237–244 (1988).
19. Jansson S, Meyer-Gauen G, Cerff R, Martin W: Nucleotide distribution in gymnosperm nuclear sequences suggests a model for GC-content change in land plant nuclear genomes. J Mol Evol 39: 34–46 (1994).
20. Kersanach R, Brinkmann H, Liaud M-F, Zhang D-X, Martin W, Cerff R: Five identical intron positions in ancient duplicated genes of eubacterial origin. Nature 367: 387–389 (1994).
21. Kohl L, Callens M, Wierenga R, Opperdoes F, Michels P: Triosephosphate isomerase from *Leishmania mexicana mexicana*: cloning and characterization of the gene, overexpression in *Escherichia coli* and analysis of the protein. Eur J Biochem 220: 331–338 (1994).

22. Kolb E, Harris JI, Bridgen J: Triose phosphate isomerase from the coelecanth. Biochem J 137: 185–197 (1974).
23. Köpke-Secundo E, Molnar I, Schnarrenberger C: Isolation and characterization of the cytosolic and chloroplastic 3-phosphoglycerate kinase from spinach leaves. Plant Physiol 93: 40–47 (1990).
24. Krebbers E, Seurinck J, Herdies L, Cashmore AR, Timko MP: Four genes in two diverged subfamilies encode the ribulose-1,5-bisphosphate carboxylase small subunit polypeptides of *Arabidopsis thaliana*. Plant Mol Biol 11: 745–759 (1988).
25. Kurzok HG, Feierabend J: Comparison of a cytosolic and a chloroplast triosephosphate isomerase isoenzyme from rye leaves. I. Purification and catalytic properties. Biochim Biophys Acta 788: 214–221 (1984).
26. Kurzok HG, Feierabend J: Comparison of a cytosolic and a chloroplast triosephosphate isomerase isoenzyme from rye leaves. II. Molecular properties and phylogenetic relationships. Biochim Biophys Acta 788: 222–233 (1984).
27. Kurzok HG, Feierabend J: Comparison of the development and site of synthesis of a cytosolic and a chloroplast isoenzyme of triosephosphate isomerase in rye leaves. J Plant Physiol 126: 207–212 (1986).
28. Maniatis T, Fritsch E, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989).
29. Maquat LE, Chilcote R, Ryan PM: Human triosephosphate isomerase cDNA and protein structure: studies of triosephosphate isomerase deficiency in man. J Biol Chem 260: 3748–3753 (1985).
30. Marchionni M, Gilbert W: The triosephosphate isomerase gene from maize: Introns antedate the plant-animal divergence. Cell 46: 133–141 (1986).
31. Martin W, Brinkmann H, Savona C, Cerff R: Evidence for a chimaeric nature of nuclear genomes: Eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci USA 90: 8692–8696 (1993).
32. Martin W, Lagrange T, Li Y-F, Bisanz-Seyer C, Mache R: Hypothesis for the evolutionary origin of the chloroplast ribosomal protein L21 of spinach. Curr Genet 18: 553–556 (1990).
33. McFadden GI, Gilson PR, Hofmann CJB, Adcock GJ, Maier U-G: Evidence that an amoeba acquired a chloroplast by retaining part of an engulfed eukaryotic alga. Proc Natl Acad Sci USA, in press (1994).
34. McKnight GL, O'Hara PJ, Parker ML: Nucleotide sequence of the triosephosphate isomerase gene from *Aspergillus nidulans*: implications for a differential loss of introns. Cell 46: 143–147 (1986).
35. Morden CW, Golden SS: Sequence analysis and phylogenetic reconstruction of the genes encoding the large and small subunits of ribulose-1,5-bisphosphate carboxylase/oxygenase from the chlorophyll *b*-containing prokaryote *Prochlorothrix hollandica*. J Mol Evol 32: 379–395 (1991).
36. Mowatt MR, Weinbach EC, Howard TC, Nash TE: Complementation of an *Escherichia coli* glycolysis mutant by *Giardia lamblia* triosephosphate isomerase. Exp Parasitol 78: 85–92 (1994).
37. Mo Y, Harris BG, Gracy RW: Triosephosphate isomerases and aldolases from light- and dark-grown *Euglena gracilis*. Arch Biochem Biophys 57: 580–587 (1973).
38. Murray NE: Phage lambda and molecular cloning. In: Hendrix RW (ed) Lambda II, pp. 395–432. Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY (1983).
39. Okada N, Koizumi N, Tanaka T, Ohkubo H, Nakanishi S, Yamada Y: Isolation, sequence, and bacterial expression of a cDNA for (S)-tetrahydroberberine oxidase from cultured berberine-producing *Coptis japonica* cells. Proc Natl Acad Sci USA 86: 534–538 (1989).
40. Ota T, Nei M: Estimation of the number of amino acid substitutions per site when the substitution rate varies among sites. J Mol Evol 38: 642–643 (1994).
41. Pelzer-Reith B, Penger A, Schnarrenberger C: Plant aldolase: cDNA and deduced amino-acid sequence of the chloroplast and cytosol enzyme from spinach. Plant Mol Biol 21: 331–340 (1993).
42. Pichersky E, Gottlieb LD, Hess JF: Nucleotide sequence of the triose phosphate isomerase gene of *Escherichia coli*. Mol Gen Genet 195: 314–320 (1984).
43. Pichersky E, Gottlieb LD: Plant triose phosphate isomerase isoenzymes: Purification, immunological and structural characterisation and partial amino acid sequences. Plant Physiol 74: 340–347 (1984).
44. Raines CA, Lloyd JC, Willingham NM, Potts S, Dyer TA: cDNA and gene sequences of wheat chloroplast sedulose-1,7-bisphosphatase reveal homology with fructose-1,6-bisphosphatases. Eur J Biochem 205: 1053–1059 (1992).
45. Reith M, Munholland J: A high resolution gene map of the chloroplast genome of the red alga *Porphyra purpurea*. Plant Cell 5: 465–475 (1993).
46. Roesler KR, Ogren WL: Nucleotide sequence of spinach cDNA encoding phosphoribulokinase. Nucl Acids Res 16: 7192 (1988).
47. Russell PR: Transcription of the triose-phosphate-isomerase gene of *Schizosaccharomyces pombe* initiates from a start point different from that in *Saccharomyces cerevisiae*. Gene 40: 125–130 (1985).
48. Saitou N, Nei M: The neighbor-joining method: a new method for the reconstruction of phylogenetic trees. Mol Biol Evol 4: 406–425 (1987).
49. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
50. Schlaepfer BS, Zuber H: Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase (gap operon) from mesophilic *Bacillus megate-*

rium: comparison with corresponding sequences from thermophilic *B.* Gene 122: 53–63 (1992).
51. Schnarrenberger C, Jacobshagen S, Müller B, Krüger I: Evolution of isozymes of sugar phosphate metabolism in green algae. In: Isozymes: Structure, Function and Use in Biology and Medicine, pp. 743–764. Wiley-Liss, New York (1990).
52. Schnarrenberger C, Pelzer-Reith B, Yatsuki H, Freund S, Jacobshagen S, Hori K: Expression and sequence of the only detectable aldolase in *Chlamydomonas reinhardtii*. Arch Biochem Biophys 313: 173–178 (1994).
53. Shaw-Lee RL, Lissemore JL, Sullivan DT: Structure and expression of the triose phosphate isomerase (*Tpi*) gene of *Drosophila melanogaster*. Mol Gen Genet 230: 225–229 (1991).
54. Shih M: Cloning and sequencing of a cDNA clone encoding cytosolic triose phosphate isomerase from *Arabidopsis thaliana*. Plant Physiol 194: 1103–1104 (1994).
55. Sogin M, Gunderson J, Elwood H, Alonso R, Peattie D: Phylogenetic meaning of the kingdom concept: an unusual Ribosomal RNA from *Giardia lamblia*. Science 243: 75–77 (1989).
56. Swinkels BW, Gibson WC, Osinga KA, Kramer R, Veeneman GH, van Boom JH, Borst P: Characterization of the gene for the microbody (glycosomal) triosephosphate isomerase of *Trypanosoma brucei*. EMBO J 5: 1291–1298 (1986).
57. Tabor S, Richardson CC: DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc Natl Acad Sci USA 84: 4767–4771 (1987).
58. Tingey S, Tsai F, Edwards J, Walker E, Coruzzi G: Chloroplast and cytosolic glutamine synthetase are encoded by homologous nuclear genes which are differentially expressed *in vivo*. J Biol Chem 263: 9651–9657 (1988).
59. Tittiger C, Whyard S, Walker VK: A novel intron site in the triosephosphate isomerase gene from the mosquito, *Culex tarsalis*. Nature 361: 470–472 (1993).
60. Weeden NF: Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. J Mol Evol 17 133–139 (1981).
61. Wierenga RK, Noble MEM, Vriend G, Nauche S, Hol-WGJ: Refined 1.83 Å structure of trypanosomal triosephosphate isomerase crystallized in the presence of 2.4 M ammonium sulfate. J Mol Biol 220: 995–1015 (1991).
62. Wolfe KH, Morden CW, Palmer JD: Function and Evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA 89: 10648–10652 (1992).
63. Wolter F, Fritz C, Willmitzer L, Schell J, Schreier P: *rbcS* genes in *Solanum tuberosum*: conservation of transit peptide and exon shuffling during evolution. Proc Natl Acad Sci USA 85: 846–850 (1988).
64. Xu Y, Harris-Haller LW, McCollum JC, Hardin SH, Hall TC:Nuclear gene encoding cytosolic triosephosphate isomerase from rice (*Oryza sativa* L.). Plant Physiol 102: 697 (1993).