Plant Molecular Biology **26**: 1155–1166, 1994.

Molecular characterization of a novel, nuclear-encoded, NAD^+-dependent glyceraldehyde-3-phosphate dehydrogenase in plastids of the gymnosperm *Pinus sylvestris* L.

Gilbert Meyer-Gauen [1], Claus Schnarrenberger [2], Rüdiger Cerff [1] and William Martin [1,*]
[1] *Institut für Genetik, Technische Universität Braunschweig, Spielmannstrasse 7, D-38023 Braunschweig, FRG (*author for correspondence);* [2] *Institut für Pflanzenphysiologie und Mikrobiologie, Freie Universität Berlin, Königin-Luise-Strasse 12–16a, D-14195 Berlin, FRG*

Received 2 June 1994; accepted in revised form 7 September 1994

Key words: sugar phosphate metabolism, chloroplast, cDNA cloning, *in vitro* import, endosymbiotic gene transfer

Abstract

Angiosperms and algae possess two distinct glyceraldehyde-3-phosphate dehydrogenase (GAPDH) enzymes, an NAD^+-dependent tetramer involved in cytosolic glycolysis and an $NADP^+$-dependent enzyme of the Calvin cycle in chloroplasts. We have found that the gymnosperm *Pinus sylvestris* possesses, in addition to these, a nuclear-encoded, plastid-specific, NAD^+-dependent GAPDH, designated GapCp, which has not previously been described from any plant. Several independent full-size cDNAs for this enzyme were isolated which encode a functional transit peptide and mature subunit very similar to that of cytosolic GAPDH of angiosperms and algae. A molecular phylogeny reveals that chloroplast GapCp and cytosolic GapC arose through gene duplication early in chlorophyte evolution. The *GapCp* gene is expressed as highly as that for *GapC* in light-grown pine seedlings. These findings suggest that aspects of compartmentalized sugar phosphate metabolism may differ in angiosperms and gymnosperms and furthermore underscore the contributions of endosymbiotic gene transfer and gene duplication to the nuclear complement of genes for enzymes of plant primary metabolism.

Introduction

The largely parallel pathways of primary carbohydrate metabolism of the cytosol and chloroplast, glycolysis, gluconeogenesis and Calvin cycle, possess a number of enzymatic reactions in common. These are catalysed by separate enzymes unique to each compartment [17, 19, 46] which in turn are encoded by distinct nuclear genes [53]. Of these chloroplast-cytosol isoenzyme pairs, glyceraldehyde-3-phosphate dehydrogenases (GAPDHs) differ in their cosubstrate specificity for NAD^+ and $NADP^+$, respectively, and have been studied in greatest detail at the molecular level [30, 40, 54, 13, 21].

Glyceraldehyde-3-phosphate dehydrogenases are integral to compartmentalized primary carbohydrate metabolism in higher plants. Photosynthate generated by the Calvin cycle or by starch degradation in plastids is transported to the cytosol primarily in the form of divalent dihydroxyacetone phosphate in exchange [48] for inorganic

phosphate via the phosphate translocator, the major protein of the inner chloroplast envelope [15]. The reductive step of the Calvin cycle is catalysed by the $NADP^+$-specific GAPDH of chloroplasts (EC 1.2.1.13). This GAPDH enzyme of the Calvin cycle is a marker enzyme of chloroplasts [6, 39] and exists in angiosperms as either an A_4 homotetramer or an A_2B_2 heterotetramer. Subunit A of chloroplast GAPDH (GapA) is of slightly lower molecular weight than subunit B (GapB) [7] by virtue of a highly charged 30 amino acid carboxy-terminal extension in the latter [4]. The chloroplast enzyme may also show activity with NAD^+ as a substrate *in vitro* [6, 23] and its activity is regulated at the protein level [5, 45] via the thioredoxin system [11, 28]. Expression of the genes for GapA and GapB are regulated at the transcriptional level by light [8, 13].

In the cytosol, photosynthate exported from chloroplasts as dihydroxyacetone phosphate or 3-phosphoglycerate can be channelled into glycolysis for the generation of metabolic energy via the tricarboxylic acid cycle and oxidative phosphorylation in mitochondria. The oxidative step of glycolysis is catalysed by the NAD^+-specific GAPDH of the cytosol (EC 1.2.1.12). Cytosolic GAPDH in angiosperms is a homotetramer of identical [6] or electrophoretically distinguishable [42] subunits (GapC). *GapC* genes of angiosperms can either be constitutively expressed or are inducible by environmental stimuli [29, 42, 54].

The genes for GapA, GapB and GapC are nuclear-encoded. These were transferred from endosymbiont genomes to the nucleus during the course of evolution [31, 21]. Cyanobacteria, from which plastids descend [36], possess in their genomes homologues of both *GapC* and *GapA*. In the cyanobacterium *Anabaena variabilis*, the homologue of glycolytic GAPDH of the plant cytosol, *gap1*, is located in an operon which contains genes for the oxidative pentose phosphate cycle, the major pathway of carbohydrate breakdown in cyanobacteria [47, 38].

Plastids inherited much of their biosynthetic apparatus from free-living antecedants of modern cyanobacteria. Since cyanobacteria possess the homologues of both glycolytic and Calvin cycle GAPDH enzymes of plants, we reasoned that some plastids may have retained a gene product homologous to cytosolic GapC, although such enzymes have not been described for angiosperms or algae to date. We have recently begun study of GAPDH genes in gymnosperms [35] and have previously characterized cDNAs encoding GapC [34] and GapA [20] from *Pinus sylvestris*. We continued investigation of this gymnosperm in search of plastid-specific GapC homologues. Here we report the cDNA cloning and uptake by chloroplasts of a novel, nuclear-encoded NAD^+-dependent GAPDH enzyme, GapCp, from *Pinus*.

Materials and methods

Plant material

$Poly(A)^+$ mRNA was isolated from total RNA (kindly provided by Stefan Jansson, Umeå) isolated from one-week old light-grown *Pinus sylvestris* seedlings. Pea seedlings (*Pisum sativum* var. Rosa Krone) for chloroplast isolation were grown at 25 °C under a 14/10 h light/dark regime.

cDNA cloning

The cDNA library of *Pinus sylvestris* has been previously described [34]. 40000 recombinants were screened by plaque hybridization using a random-labelled cDNA insert for GapC from *Magnolia liliiflora* [32]. Hybridization in the first screen was performed at 55 °C in 3× SSPE, washings were performed at 55 °C in 2× SSPE. Positively hybridizing clones were subcloned into pSK plasmids (Stratagene) and terminal regions sequenced by the dideoxy method. Rescreenings of the original, unamplified library with the 165 bp *Bam* HI fragment from the 3′ non-coding region of pPSD14.6 and the 230 bp *Pst* I fragment from the transit peptide region of pPSD14 were performed as above but at 68 °C. Nested deletion series of clones pPSD14 and pPSD17 were

sequenced on both strands using either radioactive or laser flourescence [1] techniques.

In vitro *transcription and translation*

Plasmid substrates for *in vitro* transcription were purified on CsCl gradients, linearized with the appropriate restriction enzyme, incubated in the same buffer for 30 min at 37 °C with 50 μg/ml Proteinase K (Merck), phenolized, precipitated, and dissolved in 10 mM Tris-HCl and 1 mM EDTA pH 8.0. Transcription was performed for 30 min at 37 °C in 50 μl of 40 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$, 5 mM NaCl, 2 mM spermidine, 0.5 mM ATP, 0.5 mM CTP, 0.5 mM TTP, 0.05 mM GTP, 10 mM dithiothreitol, 1 mM ^{7}mG(5′)ppp(5′)G (New England Biolabs), 1 U/μl RNAsin (Pharmacia), 100 μg/ml BSA, 100 mg/ml DNA containing 1 U/μl of either T_7 or T_3 RNA polymerase (Pharmacia). Transcripts were phenolized, precipitated and dissolved in 10 μl of water for translation. Proteins for import were synthesized using rabbit reticulocyte lysate (Amersham) according to the manufacturer's protocol. Translation of capped mRNAs was more efficient than that of uncapped; reticulocyte lysate yielded more efficient incorporation and fewer low-molecular-weight radioactive peptides than wheat germ extracts (data not shown). Translation for transport experiments was performed for 90 min at 30 °C in 70 μl with 1.2 μg of *in vitro* transcript in the presence of 0.8 μCi/μl ^{35}S-methionine (1000 Ci/mmol, Amersham). Translation mixes were frozen in liquid nitrogen and stored at −80 °C prior to transport.

Import of in vitro *synthesized proteins*

Intact chloroplasts were prepared in linear Percoll gradients from 80 g of 7–10-day old pea leaves by the method described [2]. Import was performed according to [55] with a protocol kindly provided by R.-B. Klösgen (Institut für Botanik, Ludwig-Maximillian Universität München). All procedures were at 4 °C unless otherwise indicated. Isolated intact chloroplasts (100 μg chlorophyll) were incubated in Corex tubes at 25 °C for 30 min under a 200 W photo lamp with 30 μl of *in vitro* translation mix in a total volume of 240 μl transport buffer (50 mM Hepes/KOH, 330 mM sorbitol, 12 mM methionine, 1 mM ATP pH 8.0). Chloroplasts were diluted with 240 μl of transport buffer and centrifuged for one minute at 1000 × ***g***. The pellet was resuspended in 500 μl of transport buffer containing 4 mM $CaCl_2$ and 0.1 mg/ml thermolysin (Serva) and incubated for 30 min on ice. EDTA was added to a final concentration of 20 mM. The suspension was loaded onto 1.5 ml of 45% Percoll in transport buffer in 2 ml Eppendorf tubes and centrifuged for 8 min at 4000 × ***g*** in a swinging bucket rotor. The supernatant was discarded, intact chloroplasts (pellet) were resuspended in 100 μl of 10 mM Tricin and 1 mM EDTA, pH 8.0, and lysed for 15 min on ice. Stroma and thylakoid fractions were separated as described [2]. Thylakoid pellets were resuspended in 100 μl of 2 × Laemmli buffer, stroma proteins were precipitated 15 min on ice by addition of 4 volumes of acetone, pelleted for 2 min in an Eppendorf centrifuge and resuspended in 100 μl of 2 × Laemmli buffer. Proteins were electrophoresed on 12% denaturing acrylamide gels as described [44], impregnated with scintillator subsequent to electrophoresis (Amplify, Amersham), dried and subjected to autoradiography at −80 °C. Other molecular methods were performed as described [44].

Computer analysis

Sequence handling and alignment were perfomed with the GCG package [12]. Numbers of non-synonymous substitutions per non-synonymous (K_a) site between sequences were measured with the weighted pathway method [26]. A tree from the distance matrix was inferred with the neighbour-joining method [43]. Reliability of the tree was estimated by parsimony bootstrap analysis performed on a nucleotide data set of first and second codon positions with DNABOOT of the PHYLIP package [14], this approximates bootstrap parsimony at non-synonymous sites [36].

Results

GapCp of Pinus *is a novel NAD$^+$-dependent GAPDH enzyme*

Searching for a plastid-specific, NAD$^+$-dependent GAPDH species from *Pinus*, 20 positively hybridizing cDNA clones were picked at random from the first low-stringency screen of the *P. sylvestris* cDNA library using a previously described *GapC* probe (see Materials and methods). Of these, six contained hybridizing cDNA inserts greater than 1 kb. They were subcloned into pSK plasmids and their terminal sequences determined. One of these, pPSD14.6, encoded the partial sequence of a GapC-like protein which was highly divergent from that encoded by previously characterized plant *GapC* genes. To obtain full-size clones for this mRNA, the library was rescreened under stringent conditions using the 165 bp *Bam* HI fragment from the 3′ non-coding region of pPSD14.6. Positive clones were detected at a frequency of roughly one per 1000 recombinants, similar to the frequency of cytosolic *GapC* cDNAs in this tissue. λPSD14 contained the largest cDNA insert (1.6 kb) among the hybridizing clones and was subcloned into the *Not* I site of pSK+ and sequenced. The sequence was identical in the 3′ non-coding region to that determined for pPSD14.6. The open reading frame contained a sequence very similar to the GapC subunit, but this was surprisingly preceded by a proline- and serine-rich amino-terminal extension of 98 amino acids showing similarities to chloroplast transit peptides (Fig. 1). This suggested that pPSD14 encodes a plastid-specific GAPDH subunit.

To exclude the possibility that cDNA cloning artefacts might have given rise to the transit peptide-like region of pPSDC14, we prepared the 230 bp *Pst* I fragment from this amino-terminal extension of pPSD14 and rescreened 70000 recombinants from the library. 22 positives were identified which could be classified into two groups on the basis of an additional *Hin*d III site not present in the insert of pPSD14. Four of these clones (two from each class) which contained roughly 1.6 kb inserts were subcloned into the *Not* I site of pSK+ and sequenced from both ends. All were independent cDNAs, since they varied in length in the 3′ and 5′ non-coding regions. One of those which contained the *Hin*d III restriction polymorphism relative to pPSD14, pPSD17, was subcloned and sequenced. pPSD17 differs from pPSD14 by eight base substitutions (seven in the coding region, two of which are non-synonymous) and two single base indels in the 3′ non-coding region (Fig. 1). Thus, the aminoterminal extension of pPSD14 is not a cloning artefact, rather, pPSD14 and pPSD17 are independent cDNAs from allelic variants or very recently duplicated genes encoding a GapC precursor with a transit peptide.

We designate the protein product encoded by these cDNAs as GapCp and its gene as *GapCp* in accordance with the GAPDH gene nomenclature previously set forth [31]. Similar to its cytosolic homologues from higher eukaryotes, *GapCp* of *Pinus* encodes an NAD$^+$-dependent GAPDH enzyme, as evidenced by the presence of Pro-188 in the polypeptide chain (Fig. 2), a residue on the basis of its three-dimensional structure and *in vitro* mutagenesis to confer NAD$^+$ specificity to GAPDH holoenzymes [10, 9]. The amino acid sequence of GapCp is completely collinear with that of angiosperm GapC with the exception of a single amino acid deletion at position 64 of the alignment in Fig. 2.

Fig. 1. Sequences of pPSD14 (accession number L32560) and pPSD17 (L32561). The transit peptide is underlined. In the sequence of pPSD14, only differences to the pPSD17 sequence are shown. | indicates the adenosine commencing the pPSD14 sequence and the cytosine terminating the pPSD17 sequence. Amino acid exchanges between the sequences are indicated. Deletions are shown as dashes. Dots indicate residues identical in the two sequences. The stop codon is marked by an asterisk. Non-coding regions are shown in lower-case letters.

```
14          |.....
17 ttttatatcagttttcttagatttcgttcctttggttttcggatcgcaATGGCAATGGCGGCTTCCACCATGATGAGATCGTCT
                                                  M  A  M  A  A  S  T  M  M  R  S  S

14 .....
17 GTTACGGCGGGATCGTTTCGTGAAAATCATCGTCCTTCTGTGGCGGTTGACGCACAGATCCGATCAAAGATAACCAGCACCGAT
   V  T  A  G  S  F  R  E  N  H  R  P  S  V  A  V  D  A  Q  I  R  S  K  I  T  S  T  D

14 .....
17 GGCACAGTCAGATTGCCTTTCGGAAGCTTTGGAAGCAATTTCTTCGGATCCACTGCAGAGTTTGGCCAGGGTTCAGCCTTACAA
   G  T  V  R  L  P  F  G  S  F  G  S  N  F  F  G  S  T  A  E  F  G  Q  G  S  A  L  Q

14 .....
17 TGTCAGAGTGCTAGGACTTTGCAGCCAGTTAGAGCTACTGCAACAGAGGCCCCTCCTAGAGTTCAAAGATCCTCGGGAAAGGAG
   C  Q  S  A  R  T  L  Q  P  V  R  A  T  A  T  E  A  P  P  R  V  Q  R  S  S  G  K  E

14 .....G......
17 AGGACCAAGATAGGAATAAATGGATTTGGTCGTATTGGGAGGCTGGTTTTGCGTGTTGCATTGAGCAAGGAAGATATCCAAGTA
   R  T  K  I  G  I  N  G  F  G  R  I  G  R  L  V  L  R  V  A  L  S  K  E  D  I  Q/E  V

14 .....
17 GTTGCAATCAATGACCCTTTCATTGATGCCAAATACATGGCTTACATGTTTAAGTATGATTCAACTCATGGGGTATACAAAGGG
   V  A  I  N  D  P  F  I  D  A  K  Y  M  A  Y  M  F  K  Y  D  S  T  H  G  V  Y  K  G

14 .....
17 AGCTTGAAGATTGTTGATGATACGACTCTTGAAATTGATGGTCACAGAATTACTGTCAATTCTAAAAGGGATCCTTCAGAGATT
   S  L  K  I  V  D  D  T  T  L  E  I  D  G  H  R  I  T  V  N  S  K  R  D  P  S  E  I

14 .....T.....
17 CCCTGGGGTAACTATGGAGCTGAGTATGTTGTAGAATCATCTGGAGCTTTTACAACAACCGAGAAGGCTTCCGCACATCTTAAG
   P  W  G  N  Y  G  A  E  Y  V  V  E  S  S  G  A  F  T  T  T  E  K  A  S  A  H  L  K

14 ........G.....
17 GGTGGAGCTAAGAAGGTTGTAATTTCTGCTCCATCAGCGGATGCACCCATGTTTGTGGTTGGAGTAAACGAAGGATCATATAAG
   G  G  A  K  K  V  V  I  S  A  P  S  A  D  A  P  M  F  V  V  G  V  N  E  G  S  Y  K

14 .....
17 CCTGAAATGTCTATTGTTTCAAATGCAAGCTGCACAACTAATTGCCTTGCTCCTCTTGCAAAGGTGGTGAATGAAGAATTTGGT
   P  E  M  S  I  V  S  N  A  S  C  T  T  N  C  L  A  P  L  A  K  V  V  N  E  E  F  G

14 .....
17 ATTGCTGAGGCCCTCATGACCACTGTTCATGCAACAACAGCTACACAGAAGACAGTAGATGGTCCATCTATGAAGGATTGGCGT
   I  A  E  A  L  M  T  T  V  H  A  T  T  A  T  Q  K  T  V  D  G  P  S  M  K  D  W  R

14 .....
17 GGAGGTCGTGGTGCTGGACAAAATATCATTCCAAGCTCAACTGGTGCTGCAAAGGCAGTTGGGAAGGTCCTTCCAGAGTTGAAT
   G  G  R  G  A  G  Q  N  I  I  P  S  S  T  G  A  A  K  A  V  G  K  V  L  P  E  L  N

14 ........T.....
17 GGAAAGCTAACTGGAATGGCTTTCCGTGTACCAACACCCAATGTCTCAGTTGTGGATCTGACATGTCGCCTTGAGAAACCAGCA
   G  K  L  T  G  M  A  F  R  V  P  T  P  N  V  S  V  V  D  L  T  C  R  L  E  K  P  A

14 .....C
17 TCTTACGATGATATAAAAGCAGCAATGAAGGCCGCATCTGAAGGGTCACTAAAAGGCATCCTTGGATACACTGATGAAGATGTT
   S  Y  D  D  I  K  A  A  M  K  A  A  S  E  G  S  L  K  G  I  L  G  Y  T  D  E  D  V

14 ..............................A.....
17 GTTTCGAATGATTTTGTAGGCGATGCAAGGTCGAGTATCTTTGATGCTAAGGCTGGTATAGCCTTGAGTTCTACATTTGTGAAA
   V  S  N  D  F  V  G  D  A  R  S  S  I  F  D  A  K  A  G  I  A  L  S  S  T  F  V  K

14 .....C.....
17 CTTGTTTCTTGGTATGACAATGAGTGGGGATACAGCAACCGAGTGGTGGACTTGATCTCACACATGGCTTTAGTTGGTTCACGC
   L  V  S  W  Y  D  N  E  W  G  Y  S  N  R  V  V  D  L  I  S  H  M  A  L  V  G/A  S  R

14 .....-.....-....
17 AAATAGagattcttctattttggagtatttgatgggtcttttgtggacctctagcctccttacgtagtacactctagggaatga
   K  *

14 ..........c.....
17 gttaaattttaggatccttggtgtggtaattgaagaaagcaggaggtaacccagtgtgtttaggttgcctgagtattttatttt

14 .....
17 gacaaaatatgctatccatcatctctggatggtaaatttttgatgaagtactattatgccgatattggatctttttggaataa


14 ..........cagtagatttcatcttttcaaaaaaaaaaaaaaaaaaaaaaa
17 aatatattcg|
```

GAPDH Transit peptides

```
PeaA                                                          MASATFSVA
ZeaA                                                          MASSALSA
PinA                                                      MASATLSMAAVANCSST
PinCp                                          MAMAASTMMRSSVTAGSFRENHRPSVAVDAQIRS

PeaA  KPAIKANGKGFSEFSGLRNSSRHLPFSRK-SSDDFHSLVTFQTNAVGSSGGHKKSLVVEA>
ZeaA  TTVPLQQ.G.L.........AS..MR.NAT....M.A.S.R.HK..T...PRR-APT..>
PinA  HASLQVS......Y...K...SGF.CL..YNEELRFQ..A...NA.....QKR--GV...>
PinCp .ITSTDGTVRLPFG.FGS.FFGSTAEFGQG.ALQCQ.AR.L.PVRATATEAPPRVQRSSG>
```

GAPDH Mature subunits

```
                bbbbbb aaaaaaaaaaaaaaa       bbbbbb   aaaaaaaaaaa
PeaA     |<KQLKVAINGFGRIGRNFLRCWHGRKD-SPLDVIAINDT-GGVKQASHLLKYDSTLG  60
ZeaA      |<K.....................G.A..........-....E.............
PinA      |<KI.....................-.....VV....-..............I..
PinCp   |<KERT.IG.........LV..VALSKE.---IE.V....PFIDA.YMAYMF.....H.
PinC  MGSTGKI.IG.........LVA.VALT.D.---IELVGV..PFISTDYM.YMF....VH.
PeaC    MGAKI.IG.........LVA.VALK.D.---VELV.V..PFITTDYMTYMF....VH.
ZeaC     MGKI.IG.........LVA.VALQSE.---VELV.V..PFITTDYMTYMF...TVH.

                                bbbb    aaaaaaaaa   bbbbb      aaaaaaaaa
PeaA  IFD-ADVKPVGTDGISVDGKVIKVVSDRNPANLPWKELGIDLVIEGTGVFVDREGAGRHI 120
ZeaA  ...-VG.....DNA...........C....S....G.....................K..
PinA  T..-P..QVA.N.................V.....DYE.EY.V.S....T.K.K.SA.L
PinCp VYK-GSL.I.DDTTLEI..HR.T.N.K.D.SEI..GNY.AEY.V.SS.A.TTT.K.SA.L
PinC  KWKHHE.NVKDSKTLLFGE.SVA.FGC...EEI..G.V.A.....P..........K.L
PeaC  QWKNDELTVKDSNTLLFGQ.PVT.FAH...EEI..AST.A.IIV.S....T.KDK.AA.L
ZeaC  HWKHS.ITLKDSKTLLFGD.PVT.FGI...EEI..G.A.AEY.V.S....T.KDK.AA.L

      aaa  bbbb         aaaa              bbbb  ‡aaaaaaaaaaaaaaaa
PeaA  TAGAKKVLITAPRKGDIPTYVVGVNADAYTHADDIISNASCTTNCLAPFVKVLDQKFGII 180
ZeaA  Q........P..G.............Q.NPDEP.........................
PinA  Q...........G........C...E.K.YP.EE.........................
PinCp KG.....V.S..SA-.A.MF.....EGS.KPEMS.V.............LA..VNEE...A
PinC  K......V.S..S.-.A.MF.....EHQ.KSDVN.V.............LA..IND....V
PeaC  KG.....I.S..S.-.A.MF.....ENE.KPEF................LA..INDR...V
ZeaC  KG.....V.S..S.-.A.MF.....E.K..SDVN.V.............LA..IHDN...V

      bbbbbbbbb~~~~~~~~~~◊~~~~~~~~    bbbbbbb      aaaa        bbbb
PeaA  KGTMTTTHSYTGDQRLLDA-SHRDLRRARAAALNIVPTSTGAAKAVALVLPTLKGKLNGI 240
ZeaA  ....................-...........................S....N........
PinA  ....................-...........................S........
PinCp EAL...V.AT.AT.KTV.GP.MK.W.GG.G.GQ..I.S........GK...E.N...T.M
PinC  E.L...V..I.AT.KTV.GP.NK.W.GG.G.GF..I.S........GK...A.N...T.M
PeaC  E.L...V..I.AT.KTV.GP.SK.W.GG...SF..I.S........GK...A.N...T.M
ZeaC  E.L...V.AI.AT.KTV.GP.AK.W.GG...SF..I.S........GK...D.N...T.M

      bbbbbb  bbbbbbbbbbb   aaaaaaaaaaaaaaaaa   bbbbb
PeaA  ALRVPTPNVSVVDLVVQVSKKTFAEEVNEAFRESAAKELTGILSVCDEPLVSVDFRCTDV 300
ZeaA  ......................L.........DA..N......E...V.........S..
PinA  ....................E.........A....T.D...K...A.........I....S..
PinCp .F............TCRLE.PASYDDIKA.MKAASEGS.K...GYT..DV..N..VGDAR
PinC  .F.....D......T.RLE.SATYD.IKA.IKAESEGN.K...GYTEDAV..T..IGDSR
PeaC  SF....VD......T.RLE.AATYD.IKA.IK.ESEGK.K...GYTEDDV..T..IGDTR
ZeaC  SF....VD......T.RIE.GASY.DIKK.IKAASEGP.K..MGYVE.D...T..LGDSR

       bbbb   bbbb       bbbbbaaaaaaaaaaaaaaaaaaaaa
PeaA  SSTVDSSLTMVMGDDLVKVIAWYDNEWGYSQRVVDLADIVANNWK*
ZeaA  ...I.A.........M....S.....................C..Q..*
PinA  ..............M...V.........................S*
PinCp ..IF.AKAGIALSSTF..LVS.........N.....ISHM.LVASRK*
PinC  ..IF.AQAGIALS.NF..LVS.........S.....IVHM.ATQ*
PeaC  ..IF.AKAGIALN.KF..LVS.....L...T.....IVHI.KQL*
ZeaC  ..IF.AKAGIALN.HF..LVS.........N.....IRHMFKSQ*
```

Import of GapCp into pea chloroplasts

In order to demonstrate that the precursor encoded by pPSD14 possesses a functional chloroplast transit peptide, we synthesized capped mRNA from *Xba* I-linearized pPSD14 using the T_7 promoter. As a negative control for import, mRNA was synthesized from the T_3 promoter of *Xba* I-linearized pPSC15 [32] which contains a full-size cDNA insert for cytosolic GapC from *Pinus sylvestris* in the *Not* I site of pSK+. As a positive control, mRNA was synthesized from the T_3 promoter of an *Xba* I-linearized plasmid containing a full-size cDNA insert for GapB from pea [25] in the *Not* I site of pSK+. About 5 μg of capped mRNA from 1 μg of linearized plasmid substrate as estimated by ethidium bromide staining were typically obtained. These mRNAs were translated *in vitro* (see Materials and methods) to yield radioactive proteins for import. The size of the major *in vitro* synthesis products in SDS gels (45, 39, 50 kDa) correspond well to the expected molecular mass calculated from the cDNA sequences for coding regions of GapCp (46 kDa), GapC (37 kDa) and GapB (48 kDa) (lanes 1, 4 and 7, Fig. 3). The GapC and GapB translation products contained some spurious bands less than 30 kDa in size (lanes 1 and 4). The synthetic precursors were imported into freshly isolated pea chloroplasts. The stroma fraction of the imported pea GapB precursor (positive control) contained a processed radioactive polypeptide of 40 kDa, close to expected size of 39 kDa estimated for the *in vivo* GapB subunit [8] and 39.5 kDa calculated from the sequence.

The stroma fraction of the imported GapCp precursor from *Pinus* (lane 2, Fig. 3) contained a 36 kDa processed radioactive polypeptide,

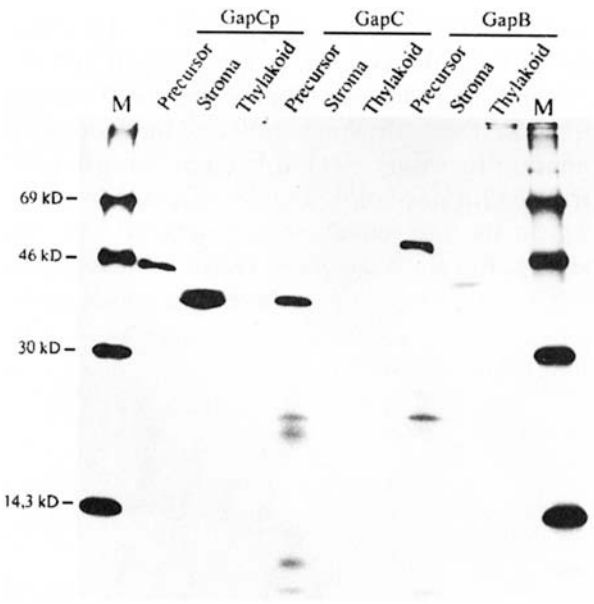

Fig. 3. In vitro synthesized GapCp, GapB, and GapC polypeptides imported into pea chloroplasts. Sizes of molecular weight standards are indicated. Lanes 1, 4 and 7: precursors synthesized *in vitro* from *in vitro* transcribed mRNA before transport, 0.5% of a translation mix was loaded in each slot. Lanes 2, 5, and 8: stroma proteins after transport corresponding to 20% of the total stroma fraction per lane. Lanes 3, 6, and 9: thylakoid proteins after transport corresponding to 20% of the total thylakoid fraction per lane. The gel was exposed for seven days at −80 °C.

slightly smaller than the *Pinus* cytosolic GapC monomer (lane 4, calculated size 37 kDa). Since the processed product has a slightly lower M_r than pine GapC, we estimate the size of the processed GapCp subunit to be about 36 kDa. This is in good agreement with the MW calculated for the mature subunit inferred from the alignment of GAPDH enzymes in Fig. 2 (35.7 kDa), indicating that the *in vivo* processing site for the *Pinus* GapCp is very close to the site suggested in the

Fig. 2. Alignment of several GAPDH sequences from plants. Gaps are indicated as dashes, residues identical to the pea GapA sequence are shown as dots. Residues which are involved in the formation of α-helices and β-pleated sheets in folding of the *Bacillus stearothermophilus* GAPDH enzyme [18] are indicated above the alignment by 'aaa' and 'bbb', respectively. >|< indicates the putative processing sites for transit peptides of *GapA* and *GapCp*. The S-loop region which confers cooperativity between monomers is underscored with a wavy line, Pro-188 which confers NAD-specificity to GAPDH tetramers [10] is indicated (◇), as is the thioester-forming cysteine residue of the catalytic center (‡). Sequence abbreviations are: PeaA, *Pisum sativum GapA*; ZeaA, *Zea mays GapA*; PinA, *Pinus sylvestris GapA*, PinCp, *Pinus sylvestris GapCp*; PinC, *Pinus sylvestris GapC*; PeaC, *Pisum sativum GapC*; ZeaC, *Zea mays GapC*.

sequence alignment. A slightly higher MW band is visible in the GapCp stroma fraction and on original autoradiograms for the pea GapB stroma fraction, these are apparently due to incomplete import processing. A faint band in the thylakoid fraction for pea GapB and pine GapCp (roughly 3% of the respective stroma fractions) was observed. But since the *in vivo* GapB subunit of pea is a stromal enzyme [6], these faint bands probably represent contaminating stromal protein in the thylakoid fraction.

Origin of the GapCp gene

We analysed the gene phylogeny of GapCp in the context of several plant and bacterial GAPDH sequences in order to reconstruct the series of events which gave rise to this coding sequence (Fig. 4). The *GapA* and *GapC* sequences from *Pinus sylvestris* were included in the analysis to underscore the diversity of the GAPDH gene family in this gymnosperm. *GapA*, *GapB*, *GapC* and *GapCp* are all descendants of the bacterial

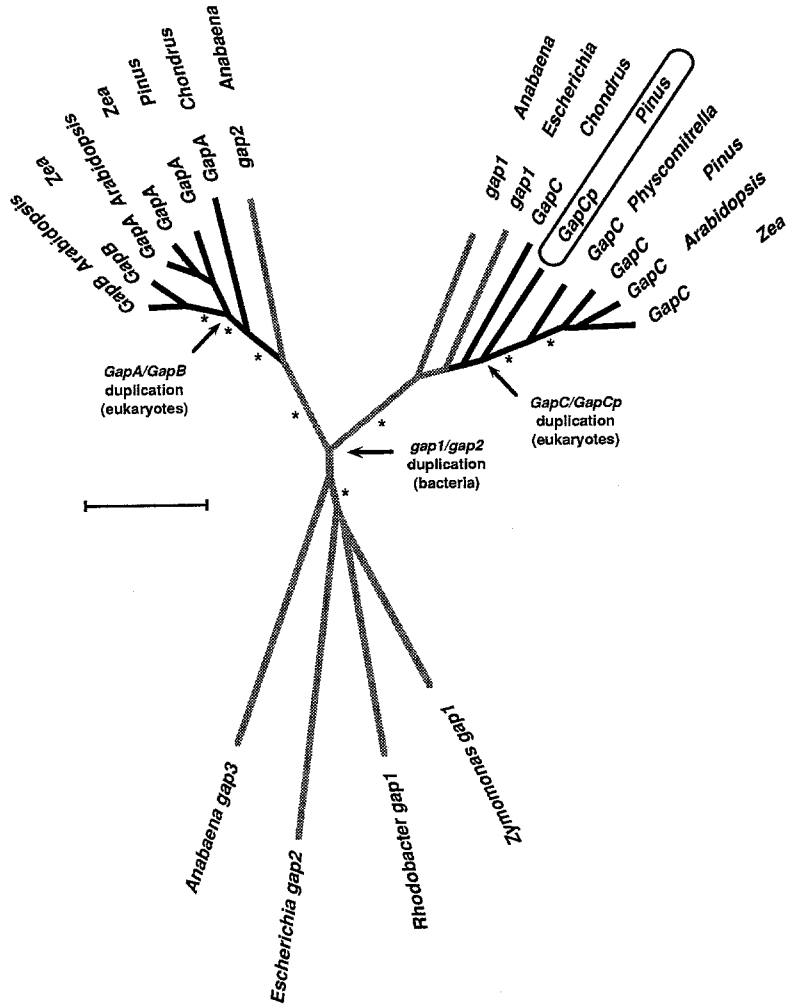

Fig. 4. Phylogenetic tree of GAPDH sequences. Gray branches indicate genes contained within prokaryotic chromosomes, solid branches indicate eukaryotic nuclear genes. The scale bar represents 0.1 non-synonymous substitution per site. Asterisks indicate branches which were detected in >90/100 bootstrap parsimony replicates. Gene duplication events are indicated. Note that the cyanobacterium *Anabaena variabilis* possesses three GAPDH genes. *Escherichia coli* also possesses a third GAPDH gene, the partial sequence of which was recently characterized (accession number L09067). Sources of sequences included in the tree are as given in references 21, 25 and 34.

GAPDH gene family as indicated by their terminal positions on the cyanobacterial *gap1* and *gap2* branches, respectively. The nuclear location of eukaryotic GAPDH genes is the result of endosymbiotic gene transfer from organelles [31, 21]. *GapCp* is only marignally more closely related to *GapC* genes from chlorophytes than it is to the *GapC* gene of the rhodophyte *Chondrus crispus*. This indicates that the gene duplication which gave rise to *GapCp* and *GapC* took place early in chlorophyte evolution, yet subsequent to establishment of ancestral *GapC/GapCp* genes in eukaryotic DNA.

Discussion

Primary metabolism in plastids is similar in many respects to that of modern cyanobacteria, yet most of the genes for enzymes of plastid pathways are located in the nucleus. With the notable exceptions of *rbcL* – and in the case of rhodoplasts also phosphoglycerate mutase and *rbcS* [41] – plastids have relinquished all genes for enzymes of primary carbohydrate metabolism. During the course of endosymbiosis, these genes underwent several different fates [53]. Some were transferred to the nucleus where they acquired a transit peptide for reimport of the protein product into the organelle of their genetic origin as in the case of $NADP^+$-dependent GAPDH [3]. Others were lost and functionally substituted through the import of preexisting cytosolic gene products of the nucleus as in the case of glutamine synthase [49, 22]. Still others apparently were lost and functionally substituted through the import of gene products contributed by the mitochondrial genome as in the case of *rp121* [33], whereas some were lost altogether along with the corresponding enzyme activity. Thus, whereas nuclear genomes are chimaeras of genes acquired from different sources, plastids are biochemical chimaeras since they contain enzymes of diverse evolutionary history. Enzymes of primary carbohydrate metabolism are among the oldest and most conservatively evolving proteins known [16]. For only three of the 17 cytosol-plastid isozyme pairs reviewed by Gottlieb [17] have molecular sequence become available for both isoforms to date: aldolase [37, 50], phosphoglycerate kinase [27] and glyceraldehyde-3-phosphate dehydrogenase [30].

Pinus *possesses an* NAD^+*-dependent chloroplast GAPDH enzyme*

The majority of studies on plant GAPDH enzymes and genes have been conducted on angiosperms. In this study of a gymnosperm, we isolated and characterized several cDNA clones from the gymnosperm *Pinus sylvestris* (Scots pine) which encode a novel NAD^+-dependent, chloroplast isoform of glyceraldehyde-3-phosphate dehydrogenase. The new GAPDH enzyme has been termed GapCp, its gene *GapCp*. GapCp from *Pinus* possesses a 98 amino acid N-terminal extension (Figs. 1 and 2) which shows features common to transit peptides of cytosolic precursors of proteins imported across the chloroplast envelope [52]. Of several different full-size cDNA clones isolated encoding GapCp from *Pinus*, two were sequenced. These differ by only eight base substitutions, two of which result in conservative replacements in the derived amino acid sequences and two insertions/deletions in the 3′ non-coding region, indicating that the corresponding mRNAs are transcribed from either very recently duplicated genes or allelic variants of the same gene. Southern blots with specific probes directed against the transit peptide region indicate that GapCp is encoded in the *Pinus* genome as a small multigene family consisting of 4–6 members (data not shown).

GapCp is synthesized as a cytosolic precursor and is imported into plastids with the aid of a transit peptide (Fig. 3). The mature subunit possesses the critical residue Pro-188, which confers NAD^+ specificity to GAPDH enzymes [10]. Both *GapC* and *GapCp* constitute about 0.1% of the mRNA in young pine seedlings each. RNA extraction from mature tissues of *Pinus* is fraught with difficulties due to the presence of resins, we are currently engaged in expression of GapCp in

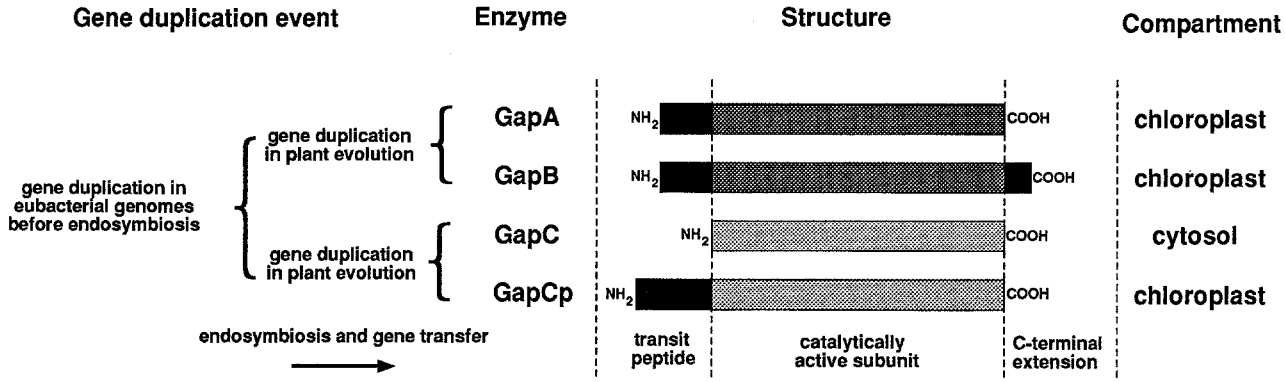

Fig. 5. Schematic structure of plant GAPDH proteins. Transit peptide regions and the carboxyterminal extension of GapB are shown as solid boxes. Catalytic subunits are shaded. Different shading for GapC-GapCp subunits and GapA-GapB subunits, respectively, emphasizes their common ancestry through gene duplication during eukaryotic evolution.

Escherichia coli in order to obtain specific antibodies for studies of gene expression kinetics and specificity in developing and mature plants.

GapCp *arose through gene duplication early in plant evolution*

A molecular phylogeny for the coding region of *GapCp* in the context of several homologues from plants and bacteria reveals that *GapCp* and *GapC* of *Pinus* are the descendants of a gene duplication event which took place early in plant evolution, in striking parallel to the *GapA-GapB* gene duplication [31]. The *GapC-GapCp* gene duplication clearly predated the emergence of land plants, since *GapCp* robustly branches below the point of divergence between *Physcomitrella* (bryophyte) and spermatophyte *GapC* sequences. It therefore seems likely that other land plants, in addition to *Pinus*, may have retained the *GapCp* gene and its product. Our preliminary studies on a cDNA library constructed from the heterosporous fern *Marsilea quadrifolia* show that this plant also expresses a *GapCp* gene which encodes a transit peptide and that the derived amino acid sequence of its encoded mature subunit is completely collinear and 92% identical to GapCp of *Pinus* in contrast to an average of 85% amino acid identity and one deletion (shared by *Pinus* GapCp at position 64 in Fig. 2) relative to higher plant *GapC* sequences (data not shown).

GapA, *GapB* and *GapC* genes and their products have been studied extensively in many angiosperms and algae [24]. As summarized in Fig. 5, the first gymnosyperm studied revealed the presence of a novel and highly expressed member of the GADPH gene family in plants, *GapCp*. The physiological role of GapCp in pine chloroplasts is not yet clear. The enzyme may be involved in carbohydrate catabolism [51], as its homologues in the cytosol of plants (GapC) and cyanobacteria (gap1) are [31]. Previous studies indicate that all genes for phosphorylating GAPDH of plants are of eubacterial origin, the finding that gymnosperm plastids have retained an NAD^+-dependent GAPDH enzyme as a relict of their cyanobacterial origin is easily reconciled with this view.

Acknowledgements

We thank Klaus Kloppstech, Hannover, and Ralf-Bernd Klösgen, Munich, for valuable advice on chloroplast import, the Gesellschaft für Biotechnologische Forschung, Braunschweig for the generous use of their computer facilities, Ole Schüsseler for porting programs to a UNIX operating system and Christiane Köhler for excellent technical assistance. This work was funded

by grants Ma 1426/1-3 and Schn 28/19-1 from the Deutsche Forschungsgemeinschaft to W.M. and C.S., respectively.

References

1. 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).
2. Bartlett SG, Grossmann AR, Chua N-H: *In vitro* synthesis and uptake of cytoplasmically-synthesized chloroplast proteins. In: Edelmann M, Hallick RB, Chua N-H (eds) Methods in Chloroplast Molecular Biology, pp. 1081–1092. Elsevier Biomedical Press, Amsterdam (1982).
3. Brinkmann H, Martinez P, Quigley F, Martin W, Cerff R: Endosymbiotic origin and codon bias of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. J Mol Evol 26: 24–33 (1987).
4. 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).
5. Buchanan BB: Role of light in the regulation of chloroplast enzymes. Annu Rev Plant Physiol 31: 341–374 (1980).
6. Cerff R: Separation and purification of NAD- and NADP-linked glyceraldehyde-3-phosphate dehydrogenases from higher plants. In: Edelmann M, Hallick RB, Chua N-H (eds) Methods in Chloroplast Molecular Biology, pp. 683–694. Elsevier Biomedical Press, Amsterdam (1982).
7. Cerff R, Chambers S: Subunit structure of higher plant glyceraldehyde-3-phosphate dehydrogenases (EC 1.2.1.12 and 1.2.1.13). J Biol Chem 254: 6094–6098 (1979).
8. Cerff R, Kloppstech K: Structural diversity and differential light control of mRNAs coding for angiosperm glyceraldehyde-3-phosphate dehydrogenases. Proc Natl Acad Sci USA 79: 7624–7628 (1982).
9. Clermont S, Corbier C, Mely Y, Gerard D, Wonacott A, Branlant G: Determinants of coenzyme specificity in glyceraldehyde-3-phosphate dehydrogenase: role of the acidic residue in the fingerprint region of the nucleotide binding fold. Biochemistry 32: 10178–10184 (1993).
10. Corbier C, Clermont S, Billard P, Skarzynski T, Branlant C, Wonacott A, Branlant G: Probing the coenzyme specificity of glyceraldehyde-3-phosphate dehydrogenases by site directed mutagenisis. Biochemistry 29: 7101–7106 (1990).
11. Cséke C, Buchanan BB: Regulation of the formation and utilisation of photosynthate in leaves. Biochim Biophys Acta 853: 43–64 (1986).
12. Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programms for the VAX. Nucl Acids Res 12: 387–395 (1984).
13. Dewdney J, Conley TR, Shih M-C, Goodmann H: Effects of blue and red light on expression of nuclear genes encoding chloroplast glyceraldehyde-3-phosphate dehydrogenase of *Arabidopsis thaliana*. Plant Physiol 103: 1115–1121 (1993).
14. Felsenstein J: Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791 (1985).
15. Flügge U-I, Weber A, Fischer K, Lottspeich F, Eckerskorn C, Waegemann K, Soll J: The major chloroplast envelope polypeptide is the phosphate translocator and not the protein import receptor. Nature 353: 364–367 (1991).
16. Fothergill-Gilmore LA, Michels PAM: Evolution of glycolysis. Progr Biophys Mol Biol 59: 105–238 (1993).
17. Gottlieb LD: Conservation and duplication of isozymes in plants. Science 216: 373–380 (1982).
18. Harris JI, Waters M: Glyceraldehyde-3-phosphate dehydrogenase. In: Boyer PD (ed) The Enzymes, Vol. 13, pp. 1–50. Academic Press, New York (1976).
19. Heber U, Pon NG, Heber M: Localization of carboxydismutase and trisoephosphate dehydrogenases in the chloroplast. Plant Physiol 38: 355–360 (1963).
20. Jansson S, Meyer-Gauen G, Cerff R, Martin W: Nucleotide distribution in gymnosperm nuclear sequences suggest a model for GC-content change in land plant nuclear genomes. J Mol Evol 39: 34–46 (1994).
21. 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).
22. Kumada Y, Benson DR, Hillemann D, Hosted TJ, Rochefort DA, Thompson CJ, Wohlleben W, Tateno Y: Evolution of the glutamin synthase gene, one of the oldest existing and functioning genes. Proc Natl Acad Sci USA 90: 3009–3013 (1993).
23. Levy LM, Betts GF: The physiological significance of aggregation phenomena of NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from spinach chloroplasts. Biochim Biophys Acta 832: 186–191 (1985).
24. Liaud M-F, Valentine C, Martin W, Bouget F-Y, Kloareg B, Cerff R: The evolutionary origin of red algae as deduced from the nuclear genes encoding cytosolic and chloroplast glyceraldehyde-3-phosphate dehydrogenases from Chondrus crispus. J Mol Evol 38: 319–327 (1994).
25. Liaud M-F, Zhang D-X, Cerff R: Differential intron loss and endosymbiotic transfer of chloroplast glyceraldehyde-3-phosphate dehydrogenase genes to the nucleus. Proc Natl Acad Sci USA 87: 8918–8922 (1990).
26. Li W-H, Wu C-I, Luo C-C: A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol 2: 150–174 (1985).
27. Longstaff M, Raines CA, McMorrow E, Bradbeer JW,

Dyer TA: Wheat phosphoglycerate kinase: evidence for recombination between the chloroplastic and cytosolic enzymes. Nucl Acids Res 17: 6569–6580 (1989).
28. Macdonald FD, Buchanan BB: The reductive pentose phosphate pathway and its regulation. In: Dennis DT, Turpin DH (eds) Plant Physiology, Biochemistry and Molecular Biology, pp. 249–262. Longman, Singapore (1990).
29. Martinez P, Martin W, Cerff R: Structure, evolution and anaerobic regulation of a nuclear gene encoding cytosolic glyceraldehyde-3-phosphate dehydrogenase from maize. J Mol Biol 208: 551–565 (1989).
30. Martin W, Cerff R: Prokaryotic features of a nucleus encoded enzyme: cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*). Eur J Biochem 159: 323–331 (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, Gierl A, Saedler H: Molecular evidence for pre-Cretaceous angiosperm origins. Nature 339: 46–48 (1989).
33. 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).
34. Martin W, Lydiate D, Brinkmann H, Forkmann G, Saedler H, Cerff R: Molecular phylogenies in angiosperm evolution. Mol Biol Evol 10: 140–162 (1993).
35. Martin W, Nock S, Meyer-Gauen G, Häger K-P, Jensen U, Cerff R: A method for isolation of cDNA-quality mRNA from immature seeds of a gymnosperm rich in polyphenolics. Plant Mol Biol 22: 555–556 (1993).
36. Martin W, Somerville CC, Loiseaux-de Goër S: Molecular phylogenies of plastid origins and algal evolution. J Mol Evol 35: 385–403 (1992).
37. 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).
38. Peschek GA: Respiratory electron transport. In: Fay P, van Baalen C (eds) The Cyanobacteria, pp. 119–162. Elsevier, Amsterdam (1987).
39. Pupillo P, Faggiani R: Subunit structure of three glyceraldehyde-3-phosphate dehydrogenases of some flowering plants. Arch Biochem Biophys 194: 581–592 (1979).
40. Quigley F, Martin W, Cerff R: Intron conservation across the prokaryote-eukaryote boundry: Structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. Proc Natl Acad Sci USA 85: 2672–2676 (1988).
41. 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).
42. Russell DA, Sachs MM: The maize cytosolic glyceraldehyde-3-phosphate dehydrogenase gene family: organ-specific expression and genetic analysis. Mol Gen Genet 229: 219–228 (1991).
43. Saitou N, Nei M: The neighbor-joining method: a new method for the reconstruction of phylogenetic trees. Mol Biol Evol 4: 406–425 (1987).
44. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
45. Scagliarini S, Trost P, Pupillo P, Valenti V: Light activation and molecular mass changes of NAD(P)-glyceraldehyde 3-phosphate dehydrogenase of spinach and maize leaves. Planta 190: 313–319 (1993).
46. Schnarrenberger C: Regulation and structure of isozymes of sugar phosphate metabolism in plants isozymes. In: Current Topics in Biological and Medical Research, vol. 16: Agriculture, Physiology, and Medicine, pp. 223–240. A.R. Liss, New York (1987).
47. Smith AJ: Modes of cyanobacterial carbon metabolism. In: Carr NG, Whitton BA (eds) The Biology of Cyanobacteria, pp. 47–85. Blackwell Scientific Publications, Oxford (1982).
48. Stitt M: The flux of carbon between the chloroplast and the cytosol. In: Dennis DT, Turpin DH (eds) Plant Physiology, Biochemistry and Molecular Biology, pp. 319–338. Longman, Singapore (1990).
49. 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).
50. Tsutsumi K, Kagaya Y, Hidaka S, Suzuki J, Tokairin J, Harai T, Hu D, Ishikawa K, Ejiri S: Structural analysis of the genes encoding chloroplastic and cytoplasmic aldolases in rice. Gene, in press.
51. Vermeglio A, Ravenel J, Peltier G: Chlororespiration: a respiratory activity in the thylakiod membrane of microalgae and higher plants. In: Weissner W, Robinson DG, Starr RC (eds) Experimental Phycology. 1. Cell Walls and Surfaces, Reproduction, Photosynthesis, pp. 188–205. Springer-Verlag, Berlin (1990).
52. von Heijne G, Stepphuhn J, Herrmann RG: Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem 180: 535–545 (1989).
53. Weeden NF: Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. J Mol Evol 17: 133–139 (1981).
54. Yang Y, Kwon H-B, Peng H-P, Shih M-C: Stress responses and metabolic regulation of glyceraldehyde-3-phosphate dehydrogenase genes in *Arabidopsis*. Plant Physiol 101: 209–216 (1993).
55. Clausmeyer S, Klösgen R-B, Herrmann RG: Protein import into chloroplasts: The hydrophilic lumenal proteins exhibit unexpected import and sorting specificities in spite of structurally conserved transit peptides. J Biol Chem 268: 13869–13876 (1993).