J Mol Evol (2004) 58:367–375
DOI: 10.1007/s00239-003-2558-7

JOURNAL OF MOLECULAR EVOLUTION

Lateral Transfer and Recompartmentalization of Calvin Cycle Enzymes of Plants and Algae

Matthew Rogers, Patrick J. Keeling

Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC V6T 1Z4 Canada

Received: 28 June 2003 / Accepted: 24 October 2003

Abstract. Certain Calvin cycle enzymes also function in glycolysis or gluconeogenisis, thus photosynthetic eukaryotes would be predicted to have ancestrally possessed cytosolic homologues of these enzymes derived from the eukaryotic host and plastid homologues from the cyanobacterial endosymbiont. In practice, the evolutionary histories of these enzymes are often more complex. Focusing on eukaryotes with secondary plastids, we have examined the evolution of four such genes: class I and II fructose bisphosphate aldolase (FBA), sedoheptulose bisphosphatase (SBPase), and fructose bisphosphatase (FBPase). We show that previously observed distributions of plastid and cytosolic homologues are not always found in algae with secondary plastids: there is evidence for multiple events of both lateral gene transfer and retargeting to a new cellular compartment for both cytosolic and plastid enzymes of plants and algae. In particular, we show that a clade of class II FBAs spans a greater diversity of eukaryotes that previously recognized and contains both plastid-targeted (*Phaeodactylum*, *Odontella*) and cytosolic (ascomycetes, oomycetes, *Euglena*, and *Bigelowiella*) forms. Lateral transfer events also gave rise to a subset of plant cytosolic FBA, as well as cytosolic FBPase in *Toxoplasma* and other coccidian apicomplexa. In contrast, it has recently been suggested that the *Trypanosoma* FBA and SBPase are derived from a plastid, however, greater taxonomic sampling shows that these enzymes provide no evidence for a plastid-containing ancestor of *Trypanosoma*. Altogether, the evolutionary histories of the FBA and SBPase/FBPase gene families are complex, including extensive paralogy, lateral transfer, and retargeting between cellular compartments.

Key words: Lateral transfer — Calvin cycle enzymes — Recompartmentalization — Plants — Algae

Introduction

Plastids, the photosynthetic organelles of plants and algae, arose through the endosymbiotic uptake of a cyanobacterium. Primary plastids, found in glaucophytes, red algae, green algae, and plants, are the direct descendants of the original symbiosis between a cyanobacterium and a nonphotosynthetic eukaryote. All other plastids have arisen through an endosymbiotic partnership between one of these groups of algae and a nonphotosynthetic eukaryote. These plastids are referred to as secondary plastids. Ultimately, all plastids trace back to the cyanobacterial endosymbiont, so plastid biochemistry is typically performed by cyanobacterial enzymes, even though the genes for most of these enzymes have been relocated to the nuclear genome of the host alga. Almost all Calvin cycle reactions (or reverse reactions) are also used in glycolysis, gluconeogenesis, or the pentose phosphate pathway, and many of these

Correspondence to: Patrick J. Keeling; *email:* pkeeling@interchange.ubc.ca

are carried out by homologous enzymes in the host cytosol. In these cases, the ancestor of plants and algae would have contained two genes with distinct evolutionary histories, one cyanobacterial gene for a plastid-targeted protein and a second eukaryotic gene for the cytosolic enzyme. In a handful of cases, this simple prediction is apparently not met, as gene duplications, losses, and transfers all seem to have taken place (Archibald et al. 2003; Brinkmann and Martin 1996; Martin et al. 1996, 2002). If these processes are common in host–plastid molecular relationships or in the general composition of their respective proteomes, then any hard line between host and endosymbiont would be blurred at the molecular level. In addition, it is presently not clear whether these processes operate in a similar manner in algae with primary and secondary plastids, which may be an important consideration since secondary endosymbiosis has been a major force in the generation of plastid and algal diversity. Here we have examined two Calvin cycle enzyme families that appear to be particularly complex: fructose bisphosphate aldolase (FBA) and sedoheptulose bisphosphatase/fructose bisphosphatase (SBPase/FBase).

FBA is divided into two classes of phylogenetically and structurally unrelated enzymes characterized by several divergent properties (Marsh and Lebherz 1992). Class I FBAs are homotetramers that form a Schiff base with their substrate and can be inhibited with borohydride reagents (Lebherz and Rutter 1969; Rutter 1964). Class II FBAs, on the other hand, are found as homodimers and require divalent cations as cofactors and are inhibited by EDTA (Rutter 1964; Zgiby et al. 2000). Both classes of FBA catalyze the cleavage of fructose-1,6-bisphosphate to glyceraldehyde-3-phosphate and dihydroxytacetone phosphate in glycolysis or the reverse aldol condensation in the Calvin cycle. In addition to the aldol condensation of triose sugars, the plastid FBA I of *Spinacia* can condense DHAP and erythrose-4-phosphate to form sedoheptulose-1,7-bisphosphate. The taxonomic distribution of both classes is complex and probably not fully realized. Class I FBA is found primarily in eukaryotes, where it is widespread (Gross et al. 1999; Jacobshagen and Schnarrenberger 1990; Lebherz et al. 1984; Marchand et al. 1988; Pelzer-Reith et al. 1993, 1994; Plaumann et al. 1997; Rutter 1964; Schnarrenberger et al. 1994). A small group of proteobacteria also possesses class I FBA, the role of which is unclear, and archaebacteria and various Gram-positive bacteria possess distinct FBAs that are distantly related to the FBA I of eukaryotes and proteobacteria (Siebers et al. 2001). Red algae, green algae, and plants possess two distinct class I FBAs, one glycolytic and gluconeogenic enzyme that functions in the cytosol and a second, plastid-targeted Calvin cycle enzyme, which is thought to have arisen by gene duplication (Gross et al. 1999; Kruger and Schnarrenberger 1983). Class II FBA is chiefly found in eubacteria, where there is extensive paralogy, and is divided into two distinct subgroups, type A and type B. The two types are distantly related and are distinguished by several large insertions and deletions (Henze et al. 1998; Nickol et al. 2000; Plaumann et al. 1997). Type A FBA II has been characterized from a variety of eubacteria, as well as the cytosol of ascomycete fungi and *Euglena* (Marsh and Lebherz 1992; Pelzer-Reith et al. 1994; Plaumann et al. 1997; Rutter 1964). Type B FBA II is also present in diverse eubacteria, several ''amitochondriate'' protists, and the plastid of the glaucophyte, *Cyanophora paradoxa* (Gross et al. 1994; Henze et al. 1998; Nickol et al. 2000; Sanchez et al. 2002).

FBPase and SBPase are related enzymes that catalyze similar reactions in the Calvin cycle: SBPase catalyzes the substrate level dephosphorylation of sedoheptulose-1,7-bisphosphate to sedoheptulose-7-phosphate (Cadet and Meunier 1988), while FBPase catalyzes the dephosphorylation of fructose-1,6-bisphosphate to fructose-6-phosphate (Zimmermann et al. 1976). FBPase is widely distributed, and the plastid FBPase of red algae, green algae, and plants are not cyanobacterial but instead are related to cytosolic forms, suggesting that they originated through gene duplication (Martin et al. 1996). In contrast, SBPase is found only in plastids of green algae and plants and the kinetoplastid, *Trypanosoma*. Like FBA, the presence of an SBPase in *Trypanosoma* is a key element in the argument for an ancestral plastid in kinetoplastid parasites (Hannaert et al. 2003).

We have investigated the evolution of plastid and cytosolic FBA, FBPase, and SBPase genes by reconstructing the phylogenies of all three enzymes including plastid and cytosolic sequences from algae with secondary plastids derived from both green and red algae. The evolution of plant and algal FBA, FBPase, and SBPase is more complex than previously considered: genes from all three families investigated here have been involved in similar kinds of events including recompartmentalization, duplicaton, and lateral gene transfer from a variety of sources. Altogether, this suggests that these processes are generally active factors in contributing to the proteome of both host and plastid components of primary and secondary algae.

Methods

Identification, Sequencing, and Primary Analysis of FBA and FBPase/SBPase Genes. Messenger RNAs encoding class I and class II FBA as well as SBPase and FBPase were identified from an ongoing *Bigelowiella natans* EST sequencing project (www.botany.ubc.ca/keeling/ChlorEST/ChlorEST.html) based on sequence similarity to known eukaryotic and eubacterial

FBA genes. All EST clones matching any known FBA, FBPase, or SBPase were isolated and fully sequenced as described (Archibald et al. 2003). Inferred amino acid sequences from all algal genes were analyzed for probable cellular location by identifying and characterizing any putative N-terminal targeting information. Secondary plastids such as those in *Bigelowiella*, diatoms, apicomplexa, and *Euglena* use a bipartite leader consisting of a signal peptide followed by a transit peptide, both of which have several predictable chemical characteristics (McFadden 1999). All genes from these organisms were examined for the presence of a peptide leader, which was analyzed using SignalP and iPsort to identify putative signal and transit peptides, respectively (Bannai et al. 2002; Nielsen et al. 1997).

Phylogenetic Analysis. Available FBA, FBPase, and SBPase were aligned using Clustal X (Jeanmougin et al. 1998). ESTs encoding an SBPase-like protein from *Magnaporthe grisea* were retrieved from the Whitehead Institute *Magnaporthe* EST databases (www-genome.wi.mit.edu and www.fungalgenomics.ncsu.edu). Trees were inferred from amino acids using distance and protein maximum likelihood methods. Distances were calculated by TREE-PUZZLE 5.0 (Strimmer and von Haeseler 1996) using the WAG substitution matrix (Goldman and Whelan 2000), and correcting for rates-across-sites variation according to a discrete gamma distribution with eight categories of variable sites and an invariable sites category. The shape parameter alpha and the proportion of invariable sites were estimated from the data. Trees were inferred from gamma-corrected distances by weighted neighbor joining using WEIGHBOR 1.0.1a (Bruno et al. 2000). Bootstrapped distances were calculated by PUZZLEBOOT (A. Roger and M. Holder, http:/www.tree-puzzle.de) under described conditions with the alpha and proportion of invariable sites parameters estimated from the original data. Protein maximum likelihood trees were also inferred for class I FBA, class II type A FBA, and class II type B FBA, FBPase, and SBPase using ProML 3.6a (Felsenstein 1993) with global rearrangements and 10 randomized sequence additions. Rates-across-sites heterogeneity was modeled using the R-option with seven rate categories and frequencies estimated by TREE-PUZZLE (six variable and one invariable categories). Bootstrap resampling was performed using protein maximum likelihood in the same way, except that four variable rate categories were used. As both type A and type B class II FBA contain many insertions and deletions with respect to one another, otherwise informative characters are unalignable between the two subclasses. To optimize the character sets of both types of class II FBA, and to more precisely resolve the phylogenies of type A and type B class II FBA individually, class II FBA was analyzed in three steps. Initially, the phylogeny of both type A and type B together (using 270 alignable characters) was inferred and bootstrapped using distance methods. Then the phylogenies of type A and type B were inferred independently using distance and maximum likelihood methods (using 324 and 267 characters, respectively). The phylogeny shown is based on the protein maximum likelihood trees of the two individual types, joined at the nodes found in the phylogeny of the entire class.

Results and Discussion

Class I FBA. The phylogeny of class I FBA (Fig. 1) shows a number of well-supported groups, including apicomplexans, animals, proteobacteria, and both plastid-targeted and cytosolic forms from plants. The fact that both plastid and cytosolic FBAs in plants are class I enzymes, and that the plastid enzyme shows no resemblance to known cyanobacterial enzymes, has been interpreted as evidence for an endosymbiotic gene replacement event in red algae, green algae, and plants. It is thought that the gene for the cytosolic enzyme duplicated and the protein product of one paralogue was targeted to the plastid, leading to the deletion of the original cyanobacterial enzyme (Gross et al. 1999). Plastid-targeted proteins from three green algae are also strongly related to the plant plastid clade, while the plastid-targeted protein from the red alga *Galdieria* is also weakly related to the plant clade. Oddly, the *Euglena* plastid-targeted FBA shows no phylogenetic affinity to other plastid-targeted enzymes, green algal enzymes, or the glycosome-targeted FBA of its kinetoplastid relative, *Trypanosoma*. The origin of the *Euglena* plastid enzyme is accordingly uncertain (Plaumann et al. 1997) but seems unlikely to be derived from green algal plastid genes, as would be expected. The cytosolic and plastid-targeted class I FBAs of *Bigelowiella* branch weakly with their respective cytosolic and plastid-targeted counterparts in algae and land plants, as would be expected. Interestingly, the *Bigelowiella* plastid-targeted FBA branches with poor support (45% ProML bootstrap support) with a divergent FBA from *Chlamydomonas* that encodes a leader predicted to be a plastid transit peptide. It appears that *Bigelowiella* has retained this divergent type of green algal FBA in its plastid.

A recent analysis of the class I FBA of *Trypanosoma* has been used to suggest that kinetoplastids have acquired their glycosomal FBA enzyme through an ancient endosymbiosis with an alga (Hannaert et al. 2003). However, we do not find this relationship when a broad sampling of eukaryotes and proteobacterial FBAs is included in the analysis (several key taxa such as *Euglena*, *Dictyostelium*, *Cryptosporidium*, and proteobacteria were not included in the original analysis). Rather, the position of *Trypanosoma* in the tree is equivocal, but it never branches with plastid-targeted FBAs. Consequently, there is no evidence that it is derived from a plastid-targeted FBA. Likewise, the FBA of alveolates shares no close affinity with that of algae and plants, as has also been suggested (Hannaert et al. 2003).

Class II FBA. The composite phylogeny of type A and B subgroups of class II FBA is shown in Fig. 2. Both types are predominantly composed of eubacterial sequences, where paralogy is apparent in several cases, and the overall relationships are not well supported. Eukaryotic and plastid-targeted class II FBA genes appear in various places in the tree. The cytosolic class II type A FBA of *Bigelowiella* falls within a clade of eukaryotic genes and shares a robustly supported branch with the cytosolic FBA of *Euglena*. This relationship is intriguing, as both organisms contain secondary plastids of green algal origin but

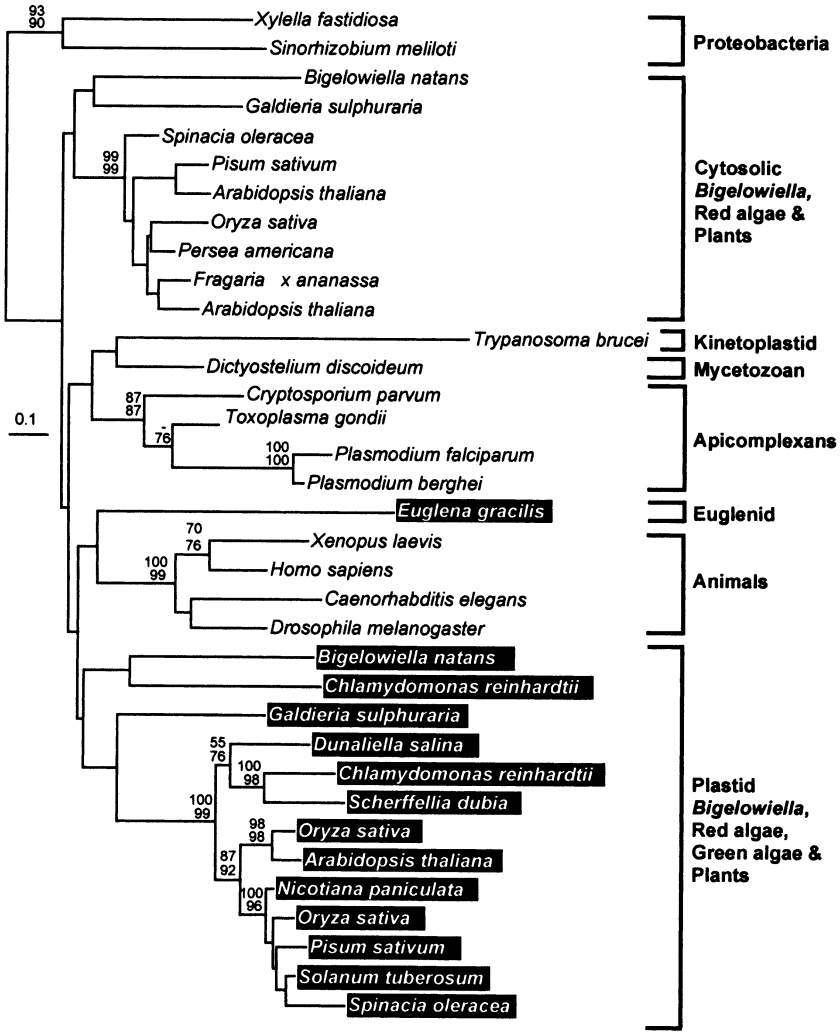

Fig. 1. Protein maximum likelihood tree of class I FBA. Numbers at nodes correspond to bootstrap values greater than 50% for major nodes obtained from weighted neighbor-joining (top) and protein maximum likelihood (bottom). Plastid-targeted genes are in black boxes. Major groups are indicated by brackets and labeled to the right.

share few other features that would support a close common line of nucleocytoplasmic descent. Nevertheless, a relationship between euglenids and chlorarachniophytes (and many other eukaryotes) has been proposed based on the proposition that their plastids originated through a single common secondary endosymbiotic event (Cavalier-Smith 1998, 2000). However, the kinetoplastid *Trypanosoma* has a class I FBA, but no class II FBA is present in the unfinished genomes of *Trypanosoma brucei* or *Leishmania major*. If the *Euglena* and *Bigelowiella* class II FBAs were both inherited from a common ancestor, then the trypanosomatids should also have inherited this enzyme. Alternatively, both may have acquired the enzyme (in parallel or in common) from their endosymbiont, but once again, no class II FBA sequences have been identified from chlorophytes including the complete genome of *Chlamydomonas*. The very strong relationship between the cytosolic FBAs of *Euglena* and *Bigelowiella* lacks an obvious explanation, but it is nonetheless an intriguing link between these two lineages. The position of the *Euglena* and *Bigelowiella* enzymes within the larger picture of class II type A FBAs is also of interest. Formerly only Euglena and fungi were known to have genes of this class, but now this strongly supported clade also includes *Bigelowiella* and a cytosolic homologue (based on the absence of an N-terminal extension and a methionine start codon at approximately the same position as the confirmed cytosolic FBA of fungi) assembled from ESTs from the oomycete heterokont *Phytophthora sojae*.

A second, strongly supported, grouping of FBAs from the diatoms *Pnaeodactylum tricornutum* and *Odontella sinensis* consistently branch at the base of this clade with modest support. Like chlorarachniophytes and euglenids, these algae possess secondary plastids but, in this instance, of red algal origin. These proteins have bipartite leaders with signal and transit peptide moieties (one of the *Phaeodactylum* FBAs is truncated within a partial leader sequence), indicating that they are plastid-targeted. Interestingly, however, these enzymes are apparently not derived from the red algal endosymbiont plastid FBA, since this enzyme would

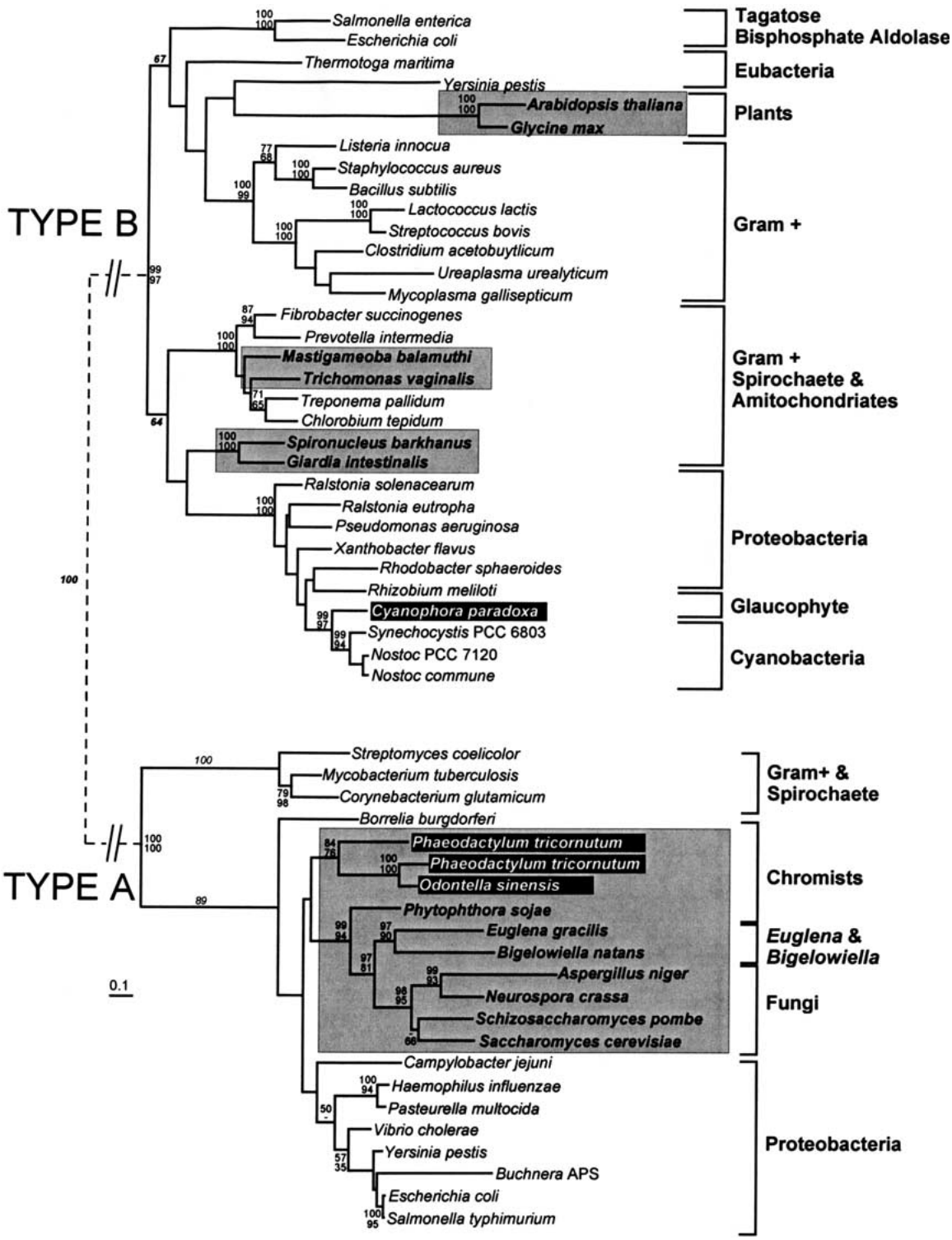

Fig. 2. Composite protein maximum-likelihood tree of type A and type B class II FBA. Type A and B trees were individually constructed and rooted (indicated by a dashed line) as described under Methods. Numbers at nodes of subtrees correspond to bootstrap values greater than 50% for major nodes obtained from weighted neighbor-joining (top) and protein maximum likelihood (bottom). Bootstrap values in italics correspond to weighted neighbor-joining values for the position of the root in global analyses. Eukaryotic genes are in a shaded box, while plastid-targeted genes are in black boxes.

be predicted to have been a class I FBA, like that of the red alga *Galdieria*. The two *Phaeodactylum* paralogues are not specifically related, indicating a gene duplication sometime in the ancestry of diatoms.

Several eukaryotes possess class II type B FBAs, but unlike type A, these are scattered about the type B subtree. These are the plastid-targeted FBA of the glaucophyte *Cyanophora paradoxa* (discussed below),

the amitochondriates, *Giardia*, *Spironucleus*, *Trichomonas*, and *Mastigamoeba*, and poorly characterised land plant genes that were identified in genome sequencing data or assembled from ESTs. Neither of the two type B sequences from *Arabidopsis* or *Glycine* is predicted to encode an N-terminal transit peptide, indicating that these enzymes are localized in the cytosol. In some analyses they show a weak affinity for the gamma proteobacterial genes for tagatose bisphosphate aldolase and an uncharacterized gene from *Yersinia pestis*. This probably reflects a bacterial origin and perhaps a nonfructose substrate for these enzymes. *Cyanophora paradoxa* has a class II type B FBA that is unrelated to any of the other eukaryotic homologues but occupies a position of special interest. This gene encodes a plastid-targeted FBA, which branches with cyanobacterial FBAs with very strong support (Nickol et al. 2000). Normally, a plastid-targeted enzyme of cyanobacterial affinity would not be particularly interesting, but in this case the plastid enzymes of red and green algae are class I enzymes likely derived from endosymbiotic gene replacement (Gross et al. 1999). This key difference is potentially important evidence for the early divergence of glaucophytes, since it alone of all algae possesses the ancestral plastid-targeted FBA. A deep position of glaucophytes has been debated extensively (Bhattacharya et al. 1995; Cavalier-Smith 1982; Helmchen et al. 1995; Martin et al. 2002), but there is not yet conclusive evidence. The FBA gene replacement may be quite helpful in supporting the glaucophytes as the earliest lineage of primary algae because it is possible to infer if the replacement took place once or twice independently based on the phylogeny of class I FBA. If an ancestral class II FBA has been replaced by a class I FBA once in a common ancestor of red algae and green algae, then class I FBA phylogeny should show the plastid and cytosolic genes of red and green algae forming distinct clades. Conversely, if red and green algae independently replaced their plastid class II enzymes with class I enzymes, then the plastid and cytosolic enzymes from each algal group should be most closely related to each other. Unfortunately the overall level of support seen in class I FBA phylogeny is low, but it does consistently support the former alternative in all analyses (Fig. 1). It therefore appears most likely that the gene plastid FBA replacement took place once in the common ancestor of green and red algae and by extension, that glaucophytes diverged prior to this event.

FBPase/SBPase. The evolution of plastid-targeted FBPase in plants and algae is similar to that of FBA I: The plastid-targeted FBPase likely arose through duplication of an ancestral cytosolic FBPase (Martin et al. 1996). The plastid-targeted FBPase of *Bigelowiella* branches with other plastid-targeted FBPase genes and it also shares a 12-amino acid insertion with them. This insertion is absent in cytosolic FBPase and has been shown to contain amino acids involved in thioredoxin binding (Hermoso et al. 1999). The second *Bigelowiella* FBPase is equivocal in its position in the tree; it branches with low support at the base of the clade consisting of eukaryotic FBPase in a manner analogous to the cytosolic FBA I of *Bigelowiella*. In contrast, the cytosolic FBPase of the apicomplexan *Toxoplasma gondii* branches within a larger clade of eubacteria, consisting of a cyanobacterium, *Chlorobium*, and several gamma-proteobacteria. This relationship is well supported and consistent with a eubacterial origin of this enzyme, although it is unclear which group of eubacteria may have served as a donor. Additional ESTs from two other coccidians, *Eimeria tenella* and *Neospora caninum* (GenBank accession Nos.: BI895768 and BF248534), are highly similar to the *Toxoplasma* sequence, but no indications of an FBPase are present in genomic sequences from *Plasmodium*, *Cryptosporidium*, or *Theileria*. Altogether, this suggests that an eubacterial FBPase has most likely been transferred to a common ancestor of coccidians following their divergence from other apicomplexans. This transfer is of functional interest, as it suggests that an important differential core carbon metabolic capability exists between different apicomplexans that was enhanced by lateral transfer. The evolutionary history of other gluconeogenic enzymes for which little information from coccidians is available (e.g., glucose-6-phosphatase) would be interesting to examine.

In plants and algae that have been examined, SBPase is exclusively found in the plastid and has no cytosolic role. Accordingly, the characterization of a cytosolic SBPase in the nonphotosynthetic kinetoplastid *Trypanosoma brucei* led to the suggestion that this enzyme (like FBA) was derived from a cryptic plastid endosymbiont (Hannaert et al. 2003). However, Fig. 3 shows that the fungi *Magnaporthe grisea* and *Neurospora crassa* also possess an apparently cytosolic SBPase, and this fungal gene forms a clade with the kinetoplastid SBPase that receives good bootstrap support from both maximum likelihood and distance methods (we also analyzed these data using Fitch–Margoliash, where the position of *Trypanosoma* was equivocal: not shown). On the whole, it appears that SBPase may have been a late addition to plastid metabolism, having persisted in the cytosol of nonphotosynthetic eukaryotes prior to the initial acquisition of the chloroplast. Indeed, cyanobacteria lack SBPase and use a dual-specificity FBPase (Tamoi et al. 1996; Yoo and Bowien 1995), so the plastid SBPase most likely came from some source other than cyanobacteria. Accordingly, there is no reason to believe that SBPase in nonphotosynthetic

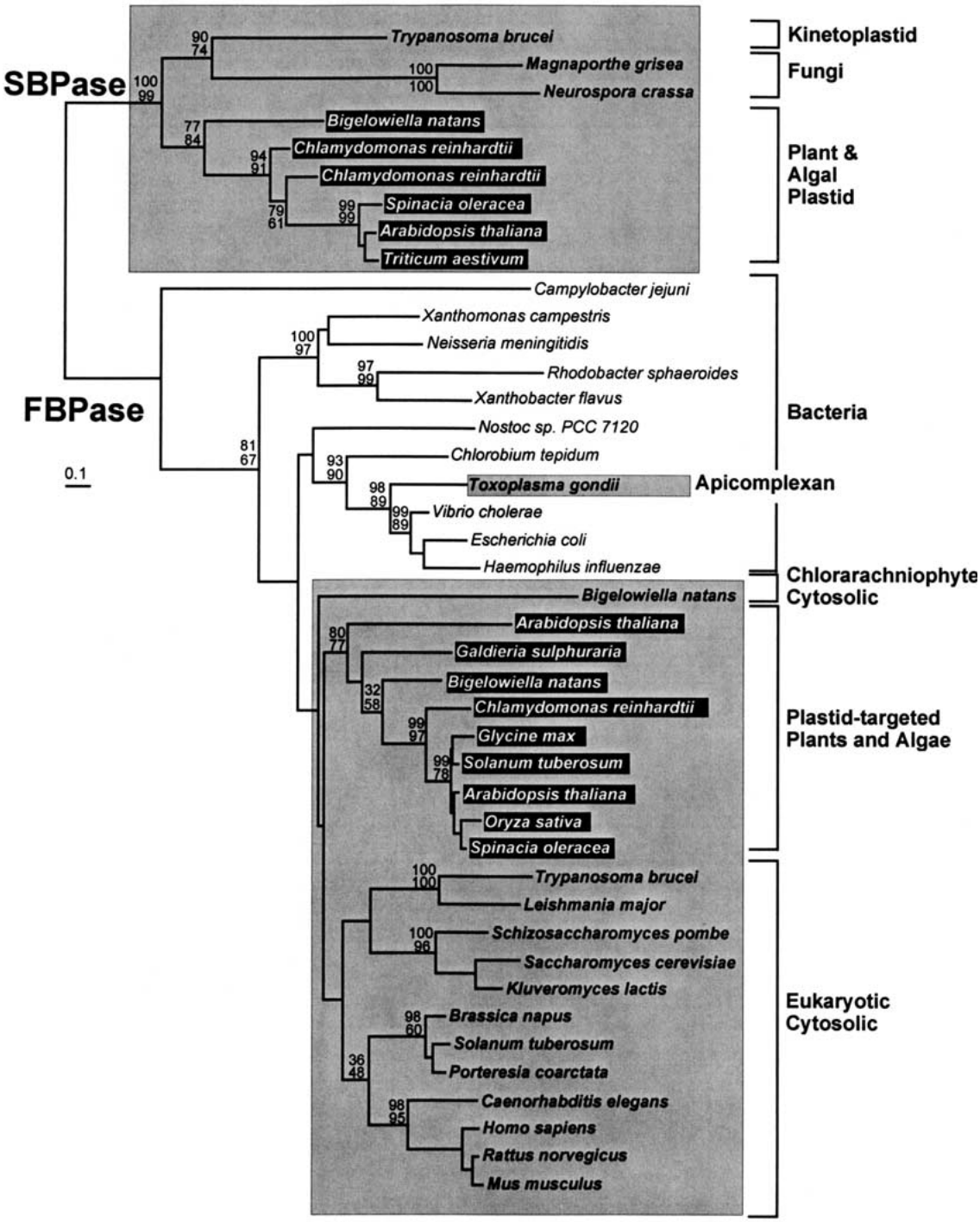

Fig. 3. Protein maximum-likelihood phylogeny of FBPase and SBPase. Numbers at nodes, shading, and brackets are as described in the legend to Fig. 2.

protests are indicative of a plastid. Still, the function of SBPase in *Trypanosoma* and fungi is puzzling. Hannaert et al. (2003) suggest that the SBPase of *Trypanosoma* may function in a modified pentose phosphate pathway. This would require the action of an FBA I specific for erythrose to generate sedoheptulose-1,7-bisphosphate, as is known to occur in the plastids of plants and alpha-proteobacteria.

Concluding Remarks: The Evolution of FBA, FBPase, and SBPase. The evolutionary history of enzymes in eukaryotic core carbon metabolism has frequently been quite colorful, and the redundancy of enzymes resulting from both primary and secondary endosymbiotic origins of plastids has added to this in several interesting ways. In all three gene families examined here, we find similar instances of lateral

gene transfer, retargeting of enzymes between cellular compartments, and cases where one or both of these processes support important relationships in the tree of eukaryotes. Two strong cases of lateral transfer are found in the plant class II type B FBA and the coccidian FBPase, both of which have interesting functional implications for the metabolism of the recipient. The diverse group of eukaryotic class II type B FBA genes might also have arisen by one or more lateral transfer events, but this is not so clear. The retargeting of an enzyme to a new cellular compartment has been observed previously in both FBA (plants and primary plastid-containing algae) and FBPase (plants), and we now also show that it most likely accounts for plastid SBPase and perhaps also the heterokont plastid class II type B FBA. While characterization of such events is becoming more common (Brinkmann and Martin 1996; Fast et al. 2001), they may still provide important evidence for major events or lineages in the tree of eukaryotes. In this case, the retention of an ancestral class II type A FBA in the plastid of *Cyanophora* is a strong indication that glaucophytes diverged prior to green and red algae. The facts that several such events have now been found to affect both primary and secondary plastid-containing algae and that such events are evident in all three genes examined here taken together suggest that these processes probably play a general role in the evolution of the plastid and host proteome in plants and algae. This is not to say that most plastid proteins are not still directly descended from the cyanobacterial endosymbiont genome, but exceptions due to these processes apparently form a significant minority with important implications.

Acknowledgments. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (227301-00). P.J.K. is a Scholar of the Canadian Institute for Advanced Research and a new investigator of the Michael Smith Foundation for Health Research and the Canadian Institutes for Health Research. We thank J.M. Archibald, J.T. Harper, and B.S. Leander for critical reading of the manuscript.

References

Archibald JM, Rogers MB, Toop M, Ishida K, Keeling PJ (2003) Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. Proc Natl Acad Sci USA 100:7678–7683

Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S (2002) Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18:298–305

Bhattacharya D, Helmchen T, Bibeau C, Melkonian M (1995) Comparisons of nuclear-encoded small-subunit ribosomal RNAs reveal the evolutionary position of the Glaucocystophyta. Mol Biol Evol 12:415–420

Brinkmann H, Martin W (1996) Higher-plant chloroplast and cytosolic 3-phosphoglycerate kinases: A case of endosymbiotic gene replacement. Plant Mol Biol 30:65–75

Bruno WJ, Socci ND, Halpern AL (2000) Weighted neighbor joining: A likelihood-based approach to distance-based phylogeny reconstruction. Mol Biol Evol 17:189–197

Cadet F, Meunier JC (1988) pH and kinetic studies of chloroplast sedoheptulose-1,7-bisphosphatase from spinach (Spinacia oleracea). Biochem J 253:249–254

Cavalier-Smith T (1982) The evolutionary origin and phylogeny of eukaryote flagella. Symp Soc Exp Biol 35:465–493

Cavalier-Smith T (1998) A revised six-kingdom system of life. Biol Rev Cambr Philos Soc 73:203–266

Cavalier-Smith T (2000) Membrane heredity and early chloroplast evolution. Trends Plant Sci 5:174–182

Fast NM, Kissinger JC, Roos DS, Keeling PJ (2001) Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. Mol Biol Evol 18:418–426

Felsenstein J (1993) PHYLIP (phylogeny inference package). University of Washington, Seattle

Goldman N, Whelan S (2000) Statistical tests of gamma-distributed rate heterogeneity in models of sequence evolution in phylogenetics. Mol Biol Evol 17:975–978

Gross W, Bayer MG, Schnarrenberger C, Gebhart UB, Maier TL, Schenk H (1994) Two distinct aldolases of class II type in the cyanoplasts and in the cytosol of the alga *Cyanophora paradoxa*. Plant Physiol 105:1393–1398

Gross W, Lenze D, Nowitzki U, Weiske J, Schnarrenberger C (1999) Characterization, cloning, and evolutionary history of the chloroplast and cytosolic class I aldolases of the red alga *Galdieria sulphuraria*. Gene 230:7–14

Hannaert V, Saavedra E, Duffieux F, Szikora JP, Rigden DJ, Michels PA, Opperdoes FR (2003) Plant-like traits associated with metabolism of *Trypanosoma* parasites. Proc Natl Acad Sci USA 100:1067–1071

Helmchen TA, Bhattacharya D, Melkonian M (1995) Analyses of ribosomal RNA sequences from glaucocystophyte cyanelles provide new insights into the evolutionary relationships of plastids. J Mol Evol 41:203–210

Henze K, Morrison HG, Sogin ML, Muller M (1998) Sequence and phylogenetic position of a class II aldolase gene in the amitochondriate protist, *Giardia lamblia*. Gene 222:163–168

Hermoso R, Lazaro JJ, Chueca A, Sahrawy M, Lopez-Gorge J (1999) Purification and binding features of a pea fructose-1,6-bisphosphatase domain involved in the interaction with thioredoxin f. Physiol Plantarum 105:756–762

Jacobshagen S, Schnarrenberger C (1990) 2 Class-I aldolases in *Klebsormidium–Flaccidum* (Charophyceae)—An evolutionary link from chlorophytes to higher-plants. J Phycol 26:312–317

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci 23:403–405

Kruger I, Schnarrenberger C (1983) Purification, subunit structure and immunological comparison of fructose-bisphosphate aldolases from spinach and corn leaves. Eur J Biochem 136:101–106

Lebherz HG, Rutter WJ (1969) Distribution of fructose diphosphate aldolase variants in biological systems. Biochemistry 8:109–121

Lebherz HG, Leadbetter MM, Bradshaw RA (1984) Isolation and characterization of the cytosolic and chloroplast forms of spinach leaf fructose diphosphate aldolase. J Biol Chem 259:1011–1017

Marchand M, Poliszczak A, Gibson WC, Wierenga RK, Opperdoes FR, Michels PA (1988) Characterization of the genes for fructose-bisphosphate aldolase in Trypanosoma brucei. Mol Biochem Parasitol 29:65–75

Marsh JJ, Lebherz HG (1992) Fructose-bisphosphate aldolases: An evolutionary history. Trends Biochem Sci 17:110–113

Martin W, Mustafa AZ, Henze K, Schnarrenberger C (1996) Higher-plant chloroplast and cytosolic fructose-1,6-bisphosphatase isoenzymes: Origins via duplication rather than prokaryote-eukaryote divergence. Plant Mol Biol 32:485–491

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246–12251

McFadden GI (1999) Plastids and protein targeting. J Eukaryot Microbiol 46:339–346

Nickol AA, Muller NE, Bausenwein U, Bayer MG, Maier TL, Schenk HE (2000) *Cyanophora paradoxa*: Nucleotide sequence and phylogeny of the nucleus encoded muroplast fructose-1,6-bisphosphate aldolase. Z Naturforsch 55:991–1003

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6

Pelzer-Reith B, Penger A, Schnarrenberger C (1993) Plant aldolase: cDNA and deduced amino-acid sequences of the chloroplast and cytosol enzyme from spinach. Plant Mol Biol 21:331–340

Pelzer-Reith B, Wiegand S, Schnarrenberger C (1994) Plastid class I and cytosol class II aldolase of *Euglena gracilis* (purification and characterization). Plant Physiol 106:1137–1144

Plaumann M, Pelzer-Reith B, Martin WF, Schnarrenberger C (1997) Multiple recruitment of class-I aldolase to chloroplasts and eubacterial origin of eukaryotic class-II aldolases revealed by cDNAs from *Euglena gracilis*. Curr Genet 31:430–438

Rutter WJ (1964) Evolution of aldolase. Fed Proc 23:1248–1257

Sanchez L, Horner D, Moore D, Henze K, Embley T, Müller M (2002) Fructose-1,6-bisphosphate aldolases in amitochondriate protists constitute a single protein subfamily with eubacterial relationships. Gene 295:51–59

Schnarrenberger C, Pelzer-Reith B, Yatsuki H, Freund S, Jacobshagen S, Hori K (1994) Expression and sequence of the only detectable aldolase in *Chlamydomonas reinhardtii*. Arch Biochem Biophys 313:173–178

Siebers B, Brinkmann H, Dorr C, Tjaden B, Lilie H, van der Oost J, Verhees CH (2001) Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. J Biol Chem 276:28710–28718

Strimmer K, von Haeseler A (1996) Quartet puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969

Tamoi M, Ishikawa T, Takeda T, Shigeoka S (1996) Molecular characterization and resistance to hydrogen peroxide of two fructose-1,6-bisphosphatases from Synechococcus PCC 7942. Arch Biochem Biophys 334:27–36

Yoo JG, Bowien B (1995) Analysis of the cbbF genes from Alcaligenes eutrophus that encode fructose-1,6-sedoheptulose-1,7-bisphosphatase. Curr Microbiol 31:55–61

Zgiby SM, Thomson GJ, Qamar S, Berry A (2000) Exploring substrate binding and discrimination in fructose 1,6-bisphosphate and tagatose 1,6-bisphosphate aldolases. Eur J Biochem 267:1858–1868

Zimmermann G, Kelly GJ, Latzko E (1976) Efficient purification and molecular properties of spinach chloroplast fructose 1,6-bisphosphatase. Eur J Biochem 70:361–367