Available online at www.sciencedirect.com
SCIENCE DIRECT®
Gene 313 (2003) 139–148
GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES
www.elsevier.com/locate/gene

Plastid-derived Type II fatty acid biosynthetic enzymes in chromists

Kim Ryall[1], James T. Harper, 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 3 January 2003; received in revised form 12 March 2003; accepted 29 April 2003
Received by A. Roger

Abstract

Fatty acid biosynthesis is a critical process for living organisms, but the evolution of the enzymes involved in this pathway is poorly understood. Animals and fungi use a Type I fatty acid synthase (FAS), a large multifunctional protein found in the cytosol. Bacteria use a Type II complex, where each enzymatic domain is a discrete polypeptide. In plants, fatty acid biosynthesis takes place in the plastid, and utilises a Type II enzyme complex. Recently, the apicomplexan parasites *Plasmodium* and *Toxoplasma* have been shown to contain the plastid-targeted Type II FAS. To investigate the distribution of this pathway, we have characterised two Type II enzymes, FabD and FabI, in three other eukaryotes with plastids derived from red algal endosymbionts: cryptomonads, heterokonts, and haptophytes. Collectively, these are referred to as chromists, and are thought to be related to apicomplexa and their relatives. Phylogenies of these enzymes show that the plastid Type II FAS enzymes are found in all groups studied, which most likely means that they originated from the red algal endosymbiont at the outset of the secondary endosymbiosis of their plastids. In addition, although plastid *fab* D genes are clearly related to one another, they are not related to cyanobacterial homologues, as would be expected. On the other hand, the strongly supported plastid *fab* I clade is related to cyanobacteria, and contains genes from chlamydiales.

Keywords: Plastid; Evolution; Phylogeny; *Plasmodium*

1. Introduction

As a central component of most lipids, the biosynthesis of fatty acids is a critical anabolic pathway in most organisms. In addition to being the major component of membranes (except in archaebacteria, where isoprenyl ethers are used instead of fatty acid esters), fatty acid lipids are an important energy storage molecule, and lipid derivatives have been found to possess a variety of other specific physiological and biological functions (Harwood, 1996).

The fundamental process of fatty acid biosynthesis is highly conserved, starting with carboxylation of acetyl-CoA to yield malonyl-CoA. The malonate group of malonyl-CoA is transferred to the phosphopantetheine prosthetic group of a small, acidic protein or protein domain called the acyl carrier protein (ACP). Malonyl-ACP is then condensed with acetyl-CoA, reduced, dehydrated, and reduced once again yielding an acyl-ACP (originally with four carbon units). The chain is elongated two carbon units at a time by condensing another malonyl-ACP with the acyl-ACP and repeating the reaction cycle. While this basic process may be conserved, many important aspects of the pathway are extremely variable when compared between different groups of organisms and, overall, the pathway appears to have had a very colourful evolutionary history. In animals and fungi, fatty acids are synthesised using a single, large, multifunctional polypeptide localised in the cytosol called the Type I Fatty Acid Synthase (FAS) (Harwood, 1996). Some fungi and bacteria possess similar multifunctional proteins that synthesise polyketides, polyketide synthases (PKS). These enzymes generally have several repeating enzymatic units, each encoding all or some of the FAS-like domains, and it is thought that the growing polyketide is passed from unit to unit as it grows (Keating and Walsh, 1999). In contrast to animals and fungi, bacteria possess Type II, or dissociable FAS, in which each enzymatic

Abbreviations: FAS, fatty acid synthase; PKS, polyketide synthase; ACP, acyl carrier protein.

* Corresponding author. Tel.: +1-604-822-4906; fax: +1-604-822-6089.
E-mail address: pkeeling@interchange.ubc.ca (P.J. Keeling).
[1] Current address: Biology Department, Duke University, P.O. Box 90338, Durham, NC 27708, USA.

0378-1119/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0378-1119(03)00671-1

activity is encoded on a separate polypeptide (Harwood, 1996). The Type I FAS is thought to have evolved by the fusion of a Type II complex into a single protein. In plants, fatty acid biosynthesis takes place in the plastid (chloroplast), and the FAS is accordingly a Type II complex since the plastid is derived from a cyanobacterial endosymbiont. The genes for these enzymes are all encoded in the nuclear genome, and the proteins are post-traslationally targeted to the plastid as is common with plastid enzymes in plants and algae (McFadden, 1999).

Recently, the obligate intracellular parasite, *Plasmodium* (the causative agent of malaria) and other members of the Phylum Apicomplexa were found to contain a plastid (McFadden et al., 1996; Wilson et al., 1996). The organelle is now known to be the product of secondary endosymbiosis, or the uptake and retention of a photosynthetic eukaryote by a second eukaryote (Archibald and Keeling, 2002). There has been some debate over whether the apicomplexan plastid originated from a red or a green alga, but the current data strongly favour the conclusion that the plastid is derived from a red alga (reviewed in (Archibald and Keeling, 2002; Williams and Keeling, 2003)). The function of this organelle has been studied extensively, and it appears to be responsible for the biosynthesis of haem, isopentenyl diphosphate (the core unit of isoprenoid compounds), and fatty acids (Wilson, 2002). Nucleus-encoded genes for plastid-localised FAS enzymes have been characterised from *Toxoplasma* and *Plasmodium*, and are Type II (Waller et al., 1998). Moreover, neither cytosolic FAS genes nor Type I FAS complex is encoded in the complete genome of *Plasmodium* (Gardner et al., 2002), suggesting that fatty acid biosynthesis in apicomplexans is restricted to the plastid. However, a FAS enzyme from the apicomplexan parasite *Cryptosporidium parvum* has recently been characterised (Zhu et al., 2000b), and it is found to be an unusually large Type I enzyme, that has three imperfect, tandomly-repeated enzymatic units. Interestingly, *Cryptosporidium* also possesses a PKS that resembles its FAS, and it has been suggested that in *Cryptosporidium* the two evolved by gene duplication (Zhu et al., 2002). This discrepancy between the situation in *Cryptosporidium* on one hand, and in *Plasmodium* and *Toxoplasma* on the other, leaves the origin of fatty acid biosynthesis in apicomplexans unclear.

To examine the distribution of the plastid Type II FAS pathway, we have characterised genes encoding Type II FAS enzymes from several other groups of plastid-containing eukaryotes that are thought to be related to apicomplexa. The apicomplexa, together with dinoflagellates and ciliates, make up the alveolates. The alveolates, in turn are now considered to be related to heterokonts, haptophytes, and cryptomonads, other eukaryotic lineages that contain plastids derived from red algae that are collectively called chromists (Cavalier-Smith, 1998; Fast et al., 2001; Yoon et al., 2002). We have characterised two enzymes from chromists, FabD and FabI. FabD (EC 2.3.1.39) is the malonyl-CoA:ACP transacylase that forms malonyl-ACP, the source of all incoming C2 units in fatty acid biosynthesis (Harwood, 1996). Interestingly, FabD has also been shown to be a link between FAS and PKS complexes of bacteria, providing the malonyl substrate for both complexes in some species (Florova et al., 2002). FabI (EC 1.3.1.9), or enoyl-ACP reductase, catalyses the final reduction to yield acyl-ACP. FabI is also a major point of regulation for bacterial and plastid fatty acid biosynthesis, and is a common drug target, including the potential antimalarial agent triclosan. Phylogenies of both enzymes indicate that plastid-derived Type II FAS is common to all three chromist groups, based on the growing evidence for a relationship between chromists and alveolates, suggests that the plastid FAS was transferred from the red algal endosymbiont to its new host very early in the evolution of the endosymbiosis that gave rise to their plastids.

2. Methods

2.1. Strains and culture conditions

Unialgal cultures of the haptophyte *Prymnesium parvum* (strain CCMP 1926), the synurophyte heterokont (stramenopile) *Mallomonas rasilis* (strain CCMP 478), the bacillariophyte heterokont *Phaeodactylum tricornutum* (strain CCMP 630), and the cryptomonads *Guillardia theta* (strain CCMP 327) and *Hemiselmis virescens* (strain CCMP 442) were obtained from the Provasoli-Guillard National Centre for Culture of Marine Phytoplankton (West Boothbay Harbor, ME, USA). All cultures were maintained in f/2-Si medium at 16 °C (12:12 light:dark cycle). DNA from the oomycete heterokont *Thraustotheca clavata* (strain CBS 343.33) was generously provided by A.W.A.M. de Cock (Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands).

2.2. Amplification of *fab D* and *fab I* genes

Depending on cell density, 20–50 ml of culture was harvested by centrifugation and mechanically disrupted under liquid nitrogen using a mortar and pestle. DNA was extracted using the DNeasy Plant Mini Kit (Qiagen). Genomic DNA was used in PCR amplification reactions using the following primers: *fab* D genes from *Guillardia*, *Phaeodactylum, Prymnesium,* and *Thraustotheca* were amplified using primers CCGCGTTCCTGTTYCCNGGNCARGG and ACCTTGCCNGGNCCNAGYTCRTA; *fab* D genes from *Mallomonas* and *Hemiselmis* were amplified using CCGCGTTCCTGTTYCCNGGNCARGG and GTCGGCGTCGACRTTNGADATNACNGG; *fab* I from *Phaeodactylum* was amplified using TTTATCGCAGGNATHGGNGAY and CGGACCTGCNGADATNGTRTT. Primer sets for *fabD* amplify approximately 81 and 70% of the mature coding regions, respectively, and the primer sets for *fabI* amplify approximately 78% of the mature coding

region. All PCR products were cloned into pCR2.1 using TOPO TA cloning (Invitrogen) and multiple individual isolates sequenced. In all cases a single product was identified. New sequences were submitted to GenBank under accession numbers AY254311 to AY254317.

2.3. Phylogenetic analyses

New sequences were aligned with conceptual translations of homologues from public databases using Clustal X. In the case of FabD, genes from a variety of bacterial Type II FAS complexes were included, as well as plastid genes from plants and *Plasmodium*. In addition, the homologous malonyl-CoA:ACP transacylase domain of several animal and fungal Type I FAS and numerous polyketide synthase genes from bacteria and fungi were also included in the alignment, as well as the FAS and polyketide synthase from the apicomplexan *Cryptosporidium parvum*. In the case of FabI, genes from a variety of bacteria as well as plant and *Plasmodium* plastid genes were aligned, but the homologous enoyl-ACP reductase domain from Type I FAS and polyketide complexes was too divergent to be aligned and was not considered. Genes from the green alga *Chlamydomonas reinhardtii* were derived by assembling a large number of expressed sequence tags. For both FabD and FabI a single gene was found and a large number of expressed sequence tags (at least 15) provided a single unambiguous sequence that encoded a plastid transit-peptide. Alignments can be provided upon request.

Amino acid alignments were used to infer phylogenies using distance and maximum likelihood methods. For distance methods, maximum likelihood distances were calculated with site-to-site rate variation modelled on a discrete gamma distribution with eight equally probable rate categories, invariable sites and the WAG substitution frequency matrix using TREE-PUZZLE 5.0 (Strimmer and von Haeseler, 1996). The proportion of invariable sites and the shape parameter alpha for the gamma distribution were estimated from the data by TREE-PUZZLE. Trees were constructed from distances using weighted neighbour-joining [using WEIGHBOR 1.0.1a: (Bruno et al., 2000)] and Fitch–Margoliash [using FITCH 3.6a: (Felsenstein, 1993)] methods. For FabD, data sets of 227 amino acid positions and 101 sequences (Type II + Type I) and 59 taxa (Type II) were analysed. For these data the alpha shape parameters were estimated to be 1.91 and 1.66, respectively, and the proportion of invariable sites were estimated to be 0.01 and 0.03, respectively. For FabI, a 232 amino acid position, 53 sequence data set was analysed and estimated to have a shape parameter alpha of 1.25 and a proportion of invariable sites of 0.02. Bootstrap support was estimated using the above method, except that the proportion of invariable sites and alpha parameter from the original data were used with each replicate, using the shell script puzzleboot (http://www.tree-puzzle.de). With the Type II FabD and FabI data sets, protein maximum likelihood trees were also inferred with site-to-site rate variability modelled using the −r option and five categories of rates (corresponding to invariable sites and four variable categories – rates and frequencies estimated by TREE-PUZZLE as described above) using ProML 3.6a (Felsenstein, 1993). Trees inferred using different methods did not differ in any significantly supported node, so only the Fitch-Margoliash trees are shown.

3. Results and discussion

3.1. FabD and FabI from heterokonts, haptophytes, and cryptomonads

Genes encoding proteins with a high similarity to FabD were amplified and sequenced from the haptophyte *Prymnesium*, the heterokonts *Mallomonas*, *Phaeodactylum,* and *Thraustotheca*, and the cryptomonds *Guillardia* and *Hemiselmis*. Genes from *Prymnesium*, *Mallomonas*, and *Thraustotheca* each contained introns, but none at positions shared with one another or *Plasmodium*. The active site of FabD has been studied in some detail and a crystal structure has been solved (Kremer et al., 2001), so the active residues and their role in catalysis are all known. Comparing the chromist and *Plasmodium* sequences with other homologues in these regions (Fig. 1A) reveals that all four active site residues encoded in the region amplified are conserved in chromist and *Plasmodium* sequences. In particular, S193 and H305 (numbered according to *P. falciparum*) are known to be located at the active site, they are involved in catalysis, and their mutation abolishes activity in bacterial enzymes (Kremer et al., 2001). S193, the nucleophile in catalysis, is part of a highly conserved GHSLG motif, which corresponds to the GXSXG motif of serine-dependent acylhydrolases (Brenner, 1988).

A single FabI was characterised from the heterokont *Phaeodactylum*. The *Phaeodactylum* gene contained three introns, but none corresponded to introns in *Plasmodium*. As with FabD, the *Phaeodactylum* FabI is highly conserved, and maintains most, but not all residues implicated in catalysis. The active site of FabI is shown in Fig. 1B. FabI is part of the short-chain alcohol dehydrogenase/reductase family, which is characterised by a $S(X)_{10}$-Y-$(X)_n$-K motif (Persson et al., 1991). In FabI, the S is replaced by a Y, and $n = 6$. This conformation is highly conserved, but curiously in all plastid targeted FabI proteins and those of chlamydiales, $n = 7$ (Fig. 1B). In addition, the *Phaeodactylum* FabI has a F residue at position 264 rather than the canonical Y. This residue is not thought to play as critical a role in FabI as in other members of the short-chain alcohol dehydrogenase/reductase family (Heath et al., 2001), so it is not clear whether this substitution is important. In addition, both *P. falciparum* and *P. yoelii* FabI sequences share a long, low complexity insertion in the otherwise highly conserved substrate-binding fold (positions 325–367

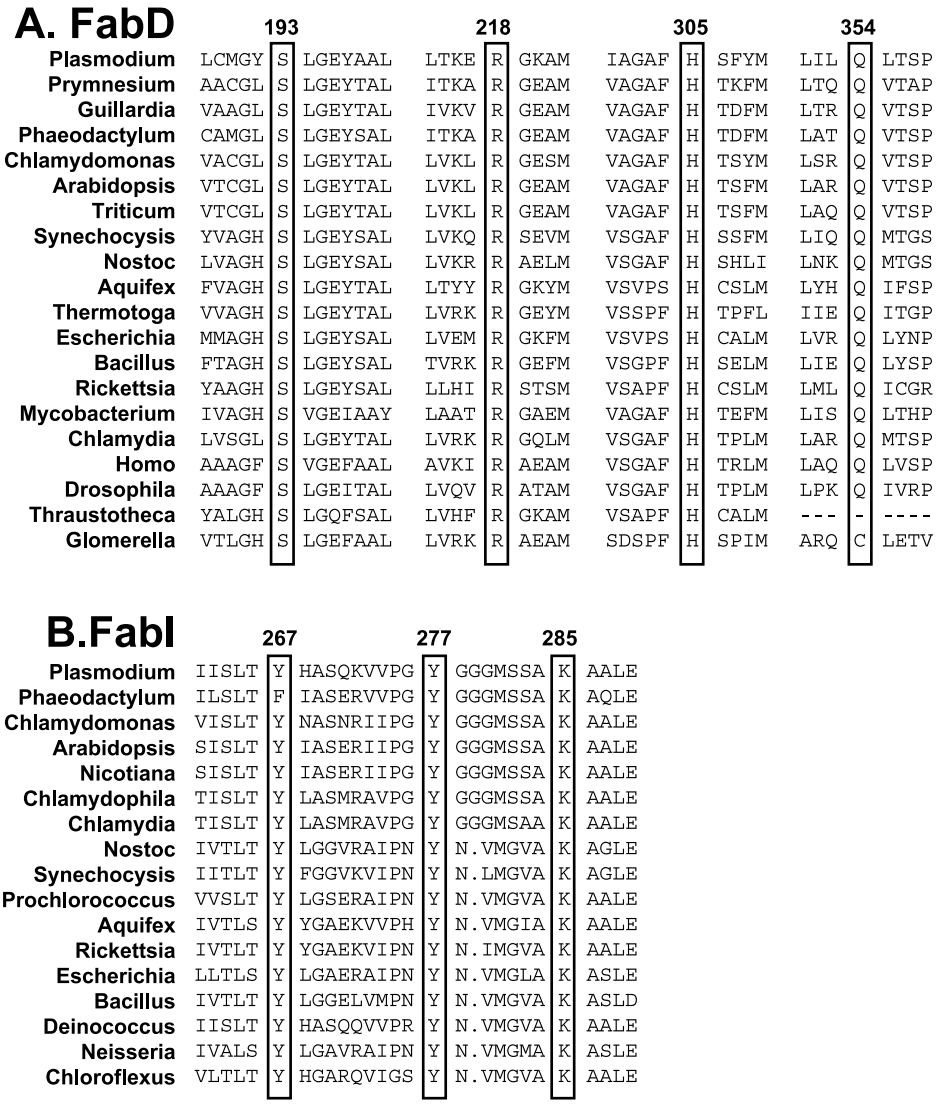


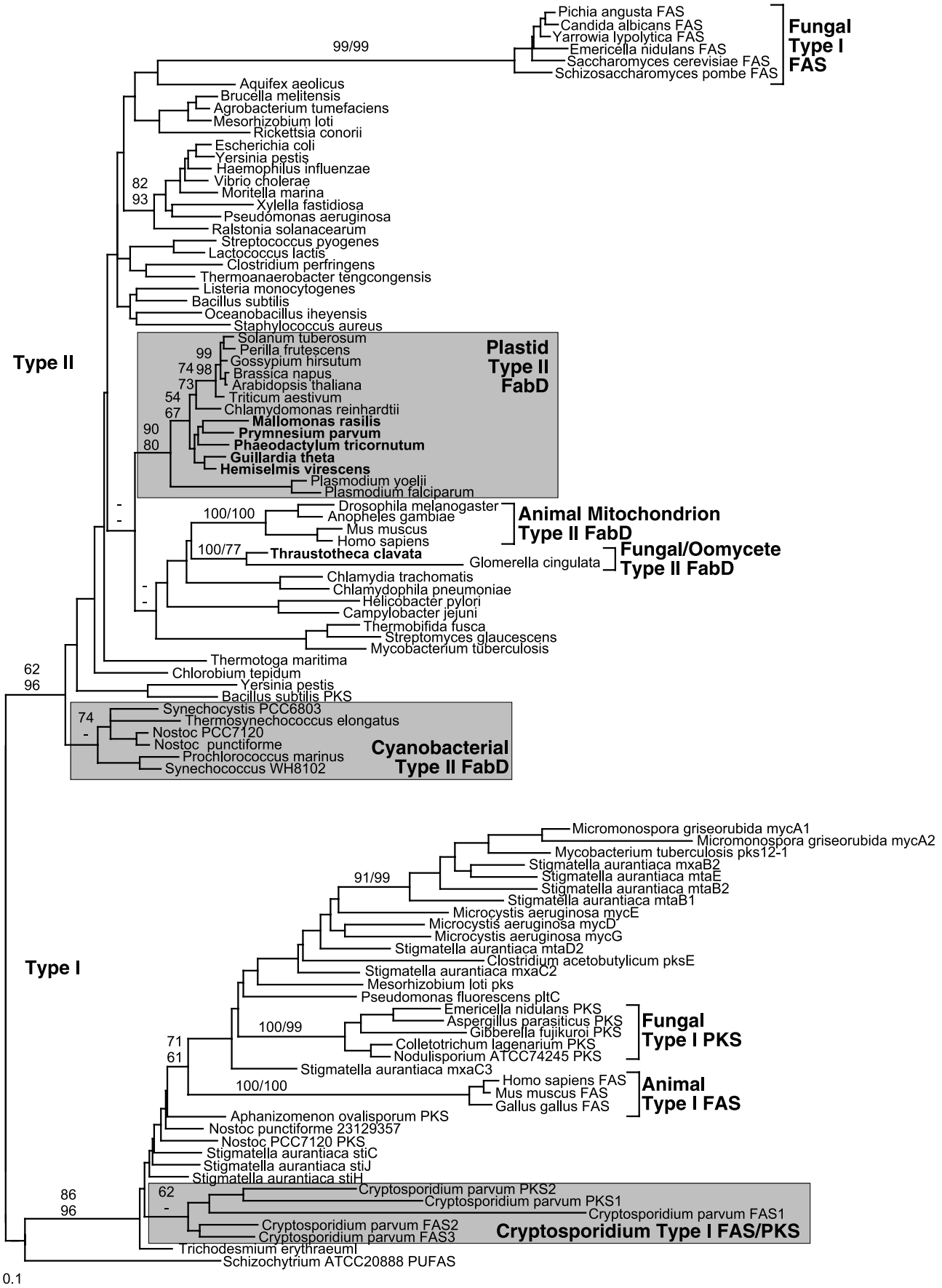
Fig. 1. Conservation of active sites of FabD (A) and FabI (B). Blocks of aligned amino acid sequences centred around residues of known function (boxed). Numbers above residues are according to the *P. falciparum* sequence. Dots indicate gaps, dashes indicate missing data.

in *P. falciparum*). *Plasmodium* proteins commonly contain such insertions (Gardner et al., 2002), and this one is typical in being notably asparagine-rich, probably due to the high AT bias of the genome.

3.2. Phylogeny of FabD and FabI

The origins of the chromist FabD and FabI sequences, and the evolution of the plastid FAS genes in general were examined by inferring phylogenies of the two proteins. FabD from plastids and bacteria were aligned with the malonyl-CoA:ACP transacylase domain of a number of animal, fungal and *Cryptosporidium* Type I FAS as well as PKS complexes from bacteria, fungi and *Cryptosporidium*. In addition, several other proteins were obviously related to FabD, in particular a subfamily of stand-alone *fab* D-like genes identified from unannotated ORFs from several animal genome projects. These sequences were all found to encode N-terminal leaders that were predicted to be mitochondrial transit peptides based on iPSORT.

Fig. 2. Phylogeny of Type II FabD, Type I FAS FabD domains, and Type I PKS FabD domains. Phylogeny shows two distinct subgroups, labelled Type I and Type II, although the fungal Type I domains branch within the Type II clade. Plastid, cyanobacterial, and *Cryptosporidium* cytosolic sequences are boxed, and other groups of relevance to the discussion are bracketed. Type I domains are labelled by their genus and species name, followed by the gene name (e.g. *myc*G), and if multiple FabD (acyl) domains exist in the same PKS, they are numbered (e.g. *mxa*C2). Numbers at nodes are bootstrap support from Fitch-Margoliash (top) and weighted neighbor-joining (bottom).


Pichia angusta FAS
Candida albicans FAS
Yarrowia lypolytica FAS
Emericella nidulans FAS
Saccharomyces cerevisiae FAS
Schizosaccharomyces pombe FAS
Fungal Type I FAS
99/99
Aquifex aeolicus
Brucella melitensis
Agrobacterium tumefaciens
Mesorhizobium loti
Rickettsia conorii
Escherichia coli
Yersinia pestis
Haemophilus influenzae
Vibrio cholerae
Moritella marina
Xylella fastidiosa
Pseudomonas aeruginosa
Ralstonia solanacearum
Streptococcus pyogenes
Lactococcus lactis
Clostridium perfringens
Thermoanaerobacter tengcongensis
Listeria monocytogenes
Bacillus subtilis
Oceanobacillus iheyensis
Staphylococcus aureus
Type II
Solanum tuberosum
Perilla frutescens
Gossypium hirsutum
Brassica napus
Arabidopsis thaliana
Triticum aestivum
Chlamydomonas reinhardtii
Mallomonas rasilis
Prymnesium parvum
Phaeodactylum tricornutum
Guillardia theta
Hemiselmis virescens
Plasmodium yoelii
Plasmodium falciparum
Plastid Type II FabD
Drosophila melanogaster
Anopheles gambiae
Mus muscus
Homo sapiens
Animal Mitochondrion Type II FabD
Thraustotheca clavata
Glomerella cingulata
Fungal/Oomycete Type II FabD
Chlamydia trachomatis
Chlamydophila pneumoniae
Helicobacter pylori
Campylobacter jejuni
Thermobifida fusca
Streptomyces glaucescens
Mycobacterium tuberculosis
Thermotoga maritima
Chlorobium tepidum
Yersinia pestis
Bacillus subtilis PKS
Synechocystis PCC6803
Thermosynechococcus elongatus
Nostoc PCC7120
Nostoc punctiforme
Prochlorococcus marinus
Synechococcus WH8102
Cyanobacterial Type II FabD
Micromonospora griseorubida mycA1
Micromonospora griseorubida mycA2
Mycobacterium tuberculosis pks12-1
Stigmatella aurantiaca mxaB2
Stigmatella aurantiaca mtaE
Stigmatella aurantiaca mtaB2
Stigmatella aurantiaca mtaB1
Microcystis aeruginosa mycE
Microcystis aeruginosa mycD
Microcystis aeruginosa mycG
Stigmatella aurantiaca mtaD2
Clostridium acetobutylicum pksE
Stigmatella aurantiaca mxaC2
Mesorhizobium loti pks
Pseudomonas fluorescens pltC
Type I
Emericella nidulans PKS
Aspergillus parasiticus PKS
Gibberella fujikuroi PKS
Colletotrichum lagenarium PKS
Nodulisporium ATCC74245 PKS
Fungal Type I PKS
Stigmatella aurantiaca mxaC3
Homo sapiens FAS
Mus muscus FAS
Gallus gallus FAS
Animal Type I FAS
Aphanizomenon ovalisporum PKS
Nostoc punctiforme 23129357
Nostoc PCC7120 PKS
Stigmatella aurantiaca stiC
Stigmatella aurantiaca stiJ
Stigmatella aurantiaca stiH
Cryptosporidium parvum PKS2
Cryptosporidium parvum PKS1
Cryptosporidium parvum FAS1
Cryptosporidium parvum FAS2
Cryptosporidium parvum FAS3
Cryptosporidium Type I FAS/PKS
Trichodesmium erythraeum!
Schizochytrium ATCC20888 PUFAS
0.1

Phylogenetic analysis of 101 FabD and FabD domains (Fig. 2) revealed that the family can be roughly divided into two subgroups. One contained the PKS FabD domains and most of the Type I FAS domains, including all the *Cryptosporidium* sequences (which form a single cluster), and the FabD domain of a previously characterised Type I polyunsaturated fatty acid synthase from the thraustochytrid (heterokont) *Schizochytrium*. The second subgroup included the stand-alone FabDs: Type II FabD (including those of plastids) together with the animal mitochondrial proteins, and a *Bacillus subtilis* gene identified as part of a Type II PKS complex (*pks* C). The fungal Type I FabD domains also branched in this group is most analyses, but in some trees branched between the two subgroups, leaving their position unclear. With the exception of the oomycete *Thraustotheca*, all the chromist sequences branched within the plastid targeted FabD clade, as expected. The *Thraustotheca* gene is not derived from the plastid, but instead branches in the heterogeneous clade consisting of bacterial FabD sequences, animal mitochondrial proteins, and the stand alone FabD from the fungus *Glomerella* (which has a short N-terminal leader that is not predicted to be a transit peptide). Interestingly, the plastid FabD sequences did not show any affinity for cyanobacterial homologues, but consistently branched with this heterogeneous clade (with no statistical support).

To focus on the relationship between plastid FabD and those of other bacteria, including cyanobacteria, the PKS domains and animal FAS domains were excluded and the smaller data set analysed. Here (Fig. 3), an unrooted tree shows many of the same relationships observed in the larger data set. The plastids still form a strongly supported clade including the new sequences from heterokonts, haptophytes, and cryptomonads, but the plastid clade still shows no affinity for the cyanobacteria. The cyanobacteria are consistently found to branch with the *B. subtilis* Type I PKS FabD and an unannotated protein from *Yersinia* (since *Yersinia* also has a FabD that is closely related to that of *E. coli*, it is possible that this second gene is also part of a PKS pathway). This relationship is relatively well supported (at 86–88% bootstrap), suggesting that the lack of affinity between plastid and cyanobacterial FabDs might indeed indicate that either the plastid or cyanobacterial homologues have an unusual evolutionary history. The *Thraustotheca* FabD branches with the *Glomerella* gene with high support, and together these are related to the animal mitochondrial genes and the FabDs from the chlamydiales. The animal/fungal/heterokont affinity may indicate an unrecognised and widespread eukaryotic FabD clade, but it is impossible to be certain of this without more thorough sampling.

The phylogeny of FabI (Fig. 4) is more restricted in scope, since the homologous domains of PKS and Type I FAS complexes are too poorly conserved to be alignable. In the phylogeny of bacterial and plastid FabI sequences, the *Phaeodactylum* FabI falls with strong statistical support within the plastid-derived FabI clade. In contrast to FabD, the FabI plastid clade is consistently found to be related to the cyanobacterial clade with modest support (59–74%). Most unusually, however, the plastid clade also consistently contains the chlamydiales (with 99–100% support). A relationship between the FabI of chlamydiales and plastids (or in some cases 'plants') has been noted previously (Stephens et al., 1998; Brinkman et al., 2002), however, the chlamydial sequences appear to have a weak but consistent affinity for the *Phaeodactylum* FabI, suggesting that these genes may be derived from a chromist or red algal plastid gene. While such a transfer has never been observed previously and would require further sampling to clarify, the phylogeny is extremely well supported, and the plastid and chlamydial sequences all contain a number of large and extremely highly conserved insertions that are found in no other FabI known to date (one small insertion at the active site is shown in Fig. 1, other larger and highly conserved insertions are not shown). Altogether there is no doubt that these sequences share a close common ancestor.

3.3. Evolution of type II fatty acid biosynthesis in chromalveolates

Following the discovery of plastids in apicomplexan parasites (McFadden et al., 1996; Wilson et al., 1996), a remarkable body of information about the biology of this organelle has quickly been assembled. This is in part due to the potential for anti-malarial targets residing in the prokaryotic biochemistry of plastids, and also due to the evolutionary novelty of plastids in intracellular parasites. Ironically, the apicomplexan plastids are arguably the best understood secondary plastids in many respects, despite being the most recently discovered and certainly the most unusual.

The first function of the apicomplexan plastids to be discovered was fatty acid biosynthesis (Waller et al., 1998). The *Toxoplasma* and *Plasmodium* FAS genes were clearly Type II enzymes and also clearly related to the plastid clade. In the simplest view, these enzymes would have ultimately originated in the cyanobacterial endosymbiont, and were first transferred to the nuclear genome of the protist that first ingested the plastid endosymbiont and established photosynthesis in eukaryotes: the ancestor of red, green, and glaucophyte algae. Subsequently, the genes encoding the enzymes of this pathway were transferred once again from the red alga that gave rise to the apicomplexan plastid to the nuclear genome of its new, secondary host. However, not all plastid biochemistry is necessarily transferred to the host during secondary endosymbiosis; indeed, the plastid shikimate pathway was not retained in the plastids of apicomplexa (Keeling et al., 1999). Moreover, the characterisation of the Type I FAS in *Cryptosporidium* (Zhu et al., 2000b) and the continuing lack of evidence for a plastid in that organism (Zhu et al., 2000a) add further uncertainty to the origin and evolution of this biochemical pathway in apicomplexa and their relatives.

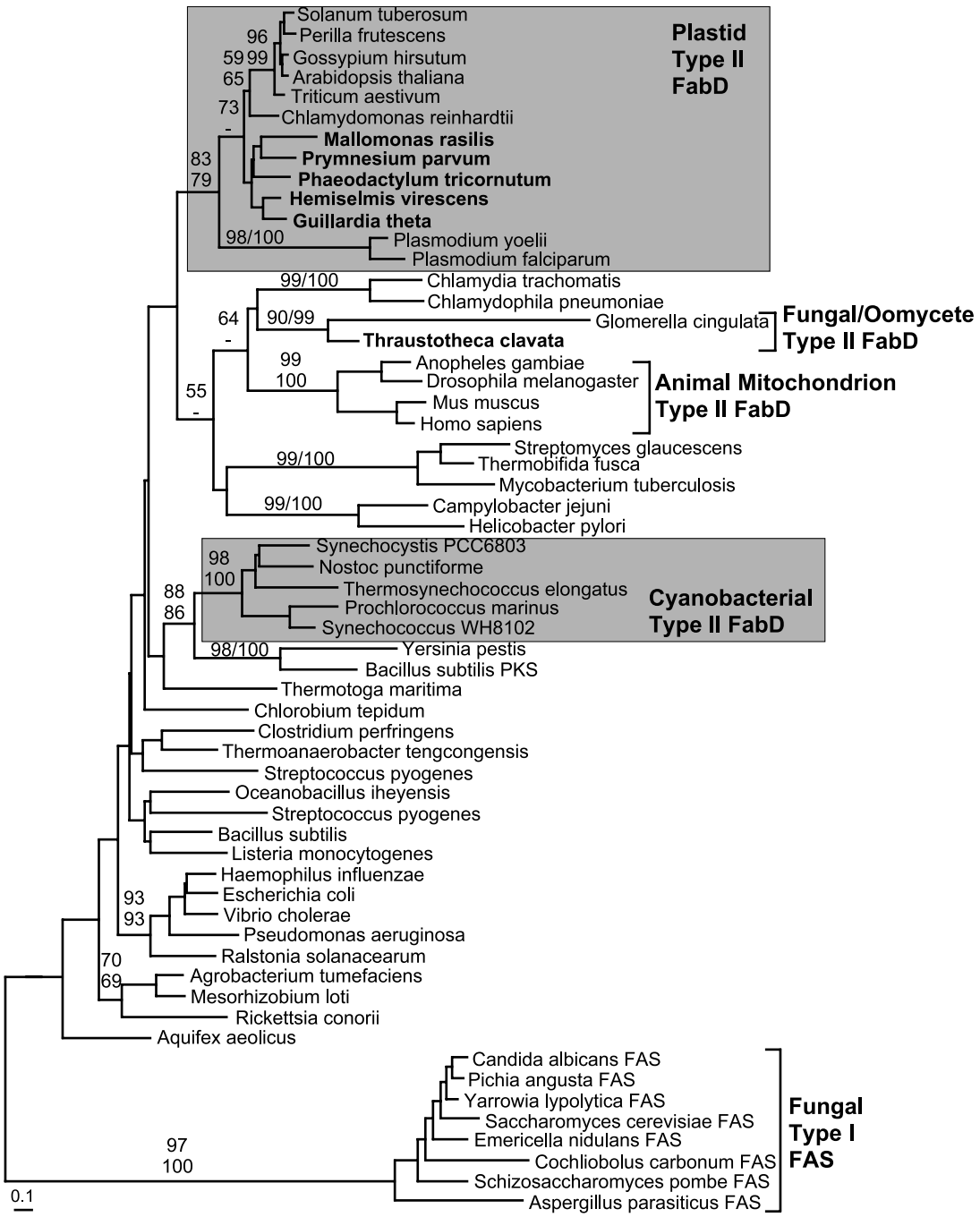


Fig. 3. Phylogeny of Type II FabD clade (including fungal Type I FAS domain). Taxa labelled as in Fig. 2. Numbers at nodes are bootstrap support from Fitch-Margoliash (top) and weighted neighbor-joining (bottom).

Recent work has tied the plastids of apicomplexa (and their close relatives, dinoflagellates) to those of chromists, altogether forming a large and diverse protist groups dubbed the chromalveolates (Cavalier-Smith, 1998). A characteristic gene replacement event suggests quite strongly that the plastids of all these organisms evolved from a single endosymbiosis (Fast et al., 2001), indicating that the plastid found in apicomplexa was acquired long before the group evolved. Here, we have demonstration that all three chromist groups, heterokonts, haptophytes, and cryptomonads, contain a plastid-derived FabD and that plastid-derived FabI is also present in heterokonts. In the context of the growing evidence for single origin of all plastids in chromists and alveolates (Cavalier-Smith, 1998; Fast et al., 2001; Yoon et al., 2002), this would mean that the Type II FAS genes were transferred from the red algal endosymbiont to the secondary host nucleus at the origin of chromalveolates, and that they are ancestral to all members

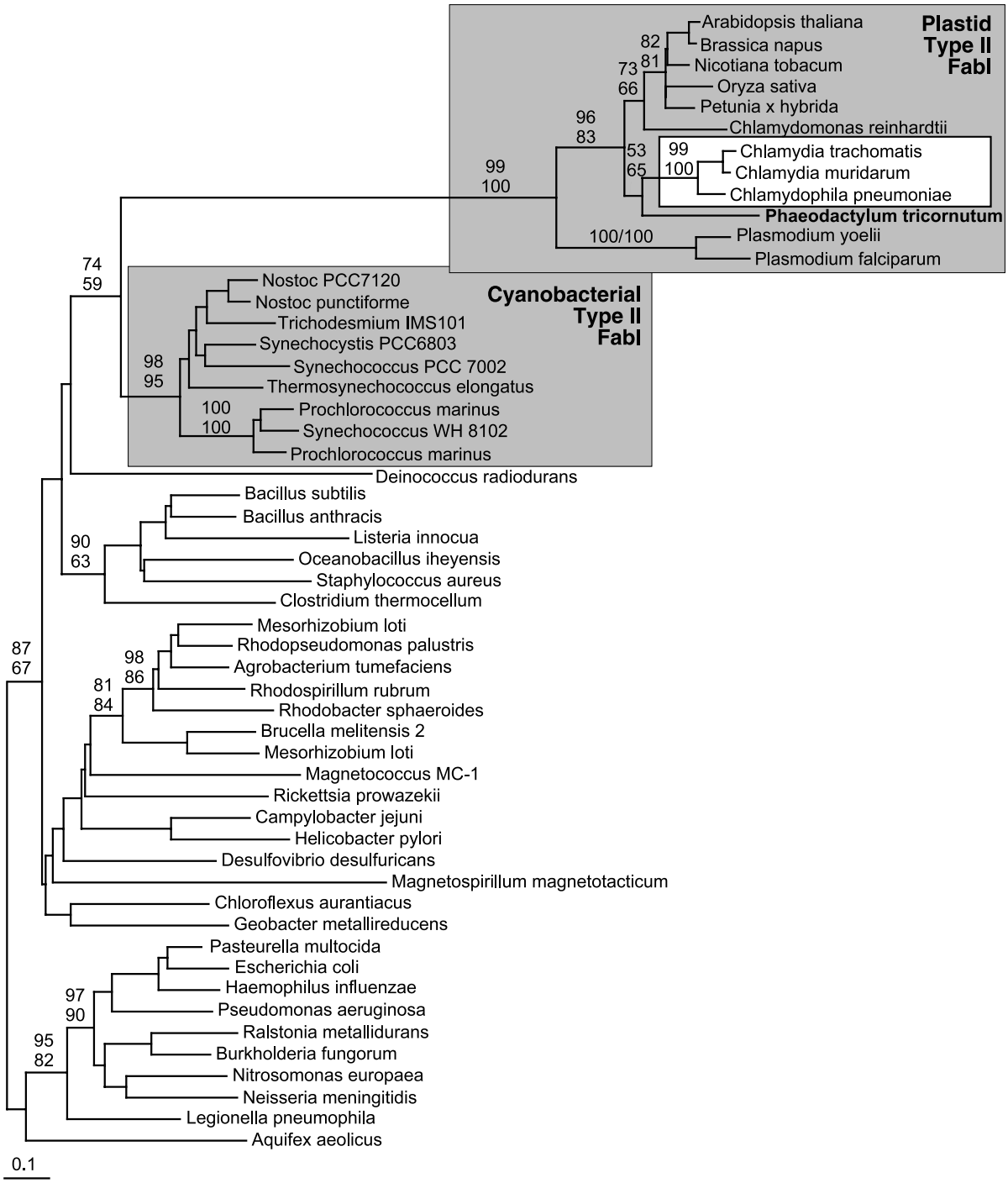


Fig. 4. Phylogeny of Type II FabI. Plastid and cyanobacterial sequences are boxed, and all other sequences are bacterial Type II FAS enzymes. Numbers at nodes are bootstrap support from Fitch-Margoliash (top) and weighted neighbor-joining (bottom). The strongly supported plastid/chlamydial group is also supported by numerous insertions, deletions, and signature sequences (see text).

of the group (although the phylogenies of FabD and FabI fail to resolve this relationship themselves). This is particularly interesting in the case of chromalveolates that are traditionally thought to lack plastids, such as ciliates (alveolates) and oomycetes (heterokonts): an early fixation of the plastid Type II FAS makes it very likely that the plastid was essential even without photosynthesis from a very early stage in its evolution. This in turn may mean that the plastid still exists in these organisms, and if so, that Type II FAS enzymes might yet be discovered. The finding of an oomycete Type II FabD was tantalising, but phylogenetic analysis shows that it is not of the plastid type, but instead more closely related to the enigmatic stand-alone Type II FabD proteins that seem to be targeted to the

mitochondrion of animals and fungi. The role of these enzymes is not clear, but they may be involved in remodelling mitochondrial lipids.

If the plastid Type II FAS did originate in an early ancestor of chromalveolates, then fatty acid biosynthesis in *Cryptosporidium* must have undergone a massive alteration. *Cryptosporidium* is now believed to be a very early branch among apicomplexans, closely related to the gregarines (Carreno et al., 1999; Leander et al., 2003). *Cryptosporidium* contains a novel Type I FAS derived from a PKS complex and there is evidence of neither Type II FAS nor plastids (Zhu et al., 2000a,b, 2002). An early origin of the plastid would mean that the Type II FAS has been lost and replaced with this unusual Type I enzyme, and the close relationship between the *Cryptosporidium* FAS and PKS suggests that this probably took place relatively recently. The loss of formerly essential biochemical pathways in the plastid could lead to the complete loss of the organelle, which may have taken place in *Cryptosporidium*.

One unexpected feature of the FabD phylogeny is the lack of affinity between plastids and cyanobacteria. Although not strongly supported, the plastid genes show a consistent relationship with a heterogeneous grouping of mitochondrial and bacterial enzymes, while the cyanobacteria tend to be related to certain other bacterial lineages. Whether this indicates that the plastid FabD is derived from a source other than the cyanobacterial endosymbiont (i.e. the original gene was replaced), will be difficult to resolve.

Lastly, the origin of FabI in chlamydiales presents an intriguing case of apparent lateral gene transfer. The similarity between chlamydial FabI and plastid FabI has been noted previously and indeed, chlamydiales have been found to contain a number of proteins with affinities to plant homologues (Brinkman et al., 2002). In particular, it has been suggested that an ancestral *Chlamydia* infected an *Acanthamoeba*-like protist, and acquired plant-like genes from it *via* lateral gene transfer (Stephens et al., 1998). This scenario assumes that *Acanthamoeba* is somehow related to plants, but this is not the case: *Acanthamoeba* is actually closely related to animals and fungi (Bapteste et al., 2002), so this fails to explain plastid-like genes in chlamydiales. Another suggestion was that chlamydiales are closely related to cyanobacteria, and the 'plant-like' genes in *Chlamydia* are really plastid-derived genes. With a poor sampling of cyanobacteria, such genes could appear more closely related to plants than the cyanobacteria (Brinkman et al., 2002). However, there are now complete genomes from ten distantly related cyanobacteria, and the FabI in chlamydiales is clearly far more closely related to plastid homologues than to any of the cyanobacterial genes. Altogether it seems most likely that the chlamydial FabI is derived from a plastid-targeted gene, however unlikely such a transfer may seem. The chamydial FabI shows a weak but consistent affinity for the *Phaeodactylum* FabI, suggesting that the source of this gene may be a red algal or red alga-derived plastid. What possible ecological or evolutionary context may have allowed this transfer to take place is not obvious, but further sampling of plastid and bacterial *fab* I genes may also shed additional light on this very unusual case of lateral gene transfer.

Acknowledgements

We thank A.W.A.W. deCock for providing *Thraustotheca clavata* DNA, and S. Ralph and J.M. Archibald for critical reading of the manuscript. This work was supported by a grant from the Canadian Institute for Health Research (MOP-42517). PJK is a scholar of the CIHR, the CIAR, and the MSFHR.

References

Archibald, J.M., Keeling, P.J., 2002. Recycled Plastids: a green movement in eukaryotic evolution. Trends Genet. 18, 577–584.

Bapteste, E., Brinkmann, H., Lee, J., et al., 2002. The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and, *Mastigamoeba*. Proc. Natl. Acad. Sci. USA 99, 1414–1419.

Brenner, S., 1988. The molecular evolution of genes and proteins: a tale of two serines. Nature 334, 528–530.

Brinkman, F.S., Blanchard, J.L., Cherkasov, A., et al., 2002. Evidence that plant-like genes in *Chlamydia* species reflect an ancestral relationship between Chlamydiaceae, cyanobacteria, and the chloroplast. Genome Res. 12, 1159–1167.

Bruno, W.J., Socci, N.D., Halpern, A.L., 2000. Weighted neighbor joining: a likelihood-based approach to distance-based phylogeny reconstruction. Mol. Biol. Evol. 17, 189–197.

Carreno, R.A., Martin, D.S., Barta, J.R., 1999. *Cryptosporidium* is more closely related to the gregarines than to coccidia as shown by phylogenetic analysis of apicomplexan parasites inferred using small-subunit ribosomal RNA gene sequences. Parasitol. Res. 85, 899–904.

Cavalier-Smith, T., 1998. A revised six-kingdom system of life. Biol. Rev. 73, 203–266.

Fast, N.M., Kissinger, J.C., Roos, D.S., Keeling, P.J., 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). J. Felsenstein, University of Washington, Seattle

Florova, G., Kazanina, G., Reynolds, K.A., 2002. Enzymes involved in fatty acid and polyketide biosynthesis in *Streptomyces glaucescens*: role of FabH and FabD and their acyl carrier protein specificity. Biochemistry 41, 10462–10471.

Gardner, M.J., Hall, N., Fung, E., et al., 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419, 498–511.

Harwood, J.L., 1996. Recent advances in the biosynthesis of plant fatty acids. Biochim. Biophys. Acta 1301, 7–56.

Heath, R.J., White, S.W., Rock, C.O., 2001. Lipid biosynthesis as a target for antibacterial agents. Prog. Lipid Res. 40, 467–497.

Keating, T.A., Walsh, C.T., 1999. Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. Curr. Opin. Chem. Biol. 3, 598–606.

Keeling, P.J., Palmer, J.D., Donald, R.G., Roos, D.S., Waller, R.F., McFadden, G.I., 1999. Shikimate pathway in apicomplexan parasites. Nature 397, 219–220.

Kremer, L., Nampoothiri, K.M., Lesjean, S., et al., 2001. Biochemical characterization of acyl carrier protein (AcpM) and

malonyl-CoA:AcpM transacylase (mtFabD), two major components of *Mycobacterium tuberculosis* fatty acid synthase II. J. Biol. Chem. 276, 27967–27974.

Leander, B.S., Clopton, R.E., Keeling, P.J., 2003. Phylogeny of gregarines (Apicomplexa) as inferred from SSU rDNA and beta-tubulin. Int. J. System. Evol. Microbiol. 53, 345–354.

McFadden, G.I., 1999. Plastids and protein targeting. J. Eukaryot. Microbiol. 46, 339–346.

McFadden, G.I., Reith, M.E., Munholland, J., Lang-Unnasch, N., 1996. Plastid in human parasites. Nature 381, 482.

Persson, B., Krook, M., Jornvall, H., 1991. Characteristics of short-chain alcohol dehydrogenases and related enzymes. Eur. J. Biochem. 200, 537–543.

Stephens, R.S., Kalman, S., Lammel, C., et al., 1998. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science 282, 754–759.

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

Waller, R.F., Keeling, P.J., Donald, R.G., et al., 1998. Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. Proc. Natl. Acad. Sci. USA 95, 12352–12357.

Williams, B.A.P., Keeling, P.J., 2003. Cryptic organelles in parasitic protists and fungi. Adv. Parasitol. in press.

Wilson, R.J., 2002. Progress with parasite plastids. J. Mol. Biol. 319, 257–274.

Wilson, R.J., Denny, P.W., Preiser, P.R., et al., 1996. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. J. Mol. Biol. 261, 155–172.

Yoon, H.S., Hackett, J.D., Pinto, G., Bhattacharya, D., 2002. The single, ancient origin of chromist plastids. Proc. Natl. Acad. Sci. USA 99, 15507–15512.

Zhu, G., Marchewka, M.J., Keithly, J.S., 2000a. *Cryptosporidium parvum* appears to lack a plastid genome. Microbiology 146, 315–321.

Zhu, G., Marchewka, M.J., Woods, K.M., Upton, S.J., Keithly, J.S., 2000b. Molecular analysis of a Type I fatty acid synthase in *Cryptosporidium parvum*. Mol. Biochem. Parasitol. 105, 253–260.

Zhu, G., LaGier, M.J., Stejskal, F., Millership, J.J., Cai, X., Keithly, J.S., 2002. *Cryptosporidium parvum*: the first protist known to encode a putative polyketide synthase. Gene 298, 79–89.