Title
Ferredoxin-dependent bilin reductases in eukaryotic algae: Ubiquity and diversity

Authors
Rockwell, Nathan C
Lagarias, J Clark

Publication Date
2017-10-01

DOI
10.1016/j.jplph.2017.05.022

Peer reviewed

HHS Public Access

Author manuscript

J Plant Physiol. Author manuscript; available in PMC 2018 October 01.

Published in final edited form as:

J Plant Physiol. 2017 October ; 217: 57–67. doi:10.1016/j.jplph.2017.05.022.

Ferredoxin-dependent bilin reductases in eukaryotic algae: ubiquity and diversity

Nathan C. Rockwell and **J. Clark Lagarias***

Department of Molecular and Cellular Biology, University of California, Davis, CA 95616

Abstract

Linear tetrapyrroles (bilins) are produced from heme by heme oxygenase, usually forming biliverdin IXα (BV). Fungi and bacteria use BV as chromophore for phytochrome photoreceptors. Oxygenic photosynthetic organisms use BV as a substrate for ferredoxin-dependent bilin reductases (FDBRs), enzymes that produce diverse reduced bilins used as light-harvesting pigments in phycobiliproteins and as photoactive photoreceptor chromophores. Bilin biosynthesis is essential for phototrophic growth in *Chlamydomonas reinhardtii* despite the absence of phytochromes or phycobiliproteins in this organism, raising the possibility that bilins are more generally required for phototrophic growth by algae. We here leverage the recent expansion in available algal transcriptomes, cyanobacterial genomes, and environmental metagenomes to analyze the distribution and diversification of FDBRs. With the possible exception of euglenids, FDBRs are present in all photosynthetic eukaryotic lineages. Phylogenetic analysis demonstrates that algal FDBRs belong to the three previously recognized FDBR lineages. Our studies provide new insights into FDBR evolution and diversification.

Keywords

algal evolution; biliverdin; bioinformatics; cyanobacteriochromes; oxygenic photosynthesis; phycocyanobilin; phycoerythrobilin; phylogenetics; phytochromobilin; secondary endosymbiosis; tetrapyrrole metabolism

INTRODUCTION

Tetrapyrroles are required for diverse metabolic processes essential to terrestrial life. Linear tetrapyrroles derived from heme (bilins) play important roles in the biology of photosynthetic organisms. Bilins are synthesized from heme via heme oxygenase (HO) to yield biliverdin IXα (BV, Fig. 1). In oxygenic photosynthetic organisms, BV is a substrate for ferredoxin-dependent bilin reductases (FDBRs: (Dammeyer and Frankenberg-Dinkel, 2008)) that generate a range of bilins by reducing different double bonds (Fig. 1). Cyanobacteria contain as many as three FDBRs: PcyA, PebA, and PebB. Phylogenetic

*Author for correspondence: J. Clark Lagarias, Tel: 530-752-1865, jclagarias@ucdavis.edu.

analysis places these three enzymes into distinct lineages (Rockwell et al., 2017). Eukaryotic FDBRs characterized to date also belong to these three lineages. For example, HY2 enzymes from streptophytes are part of the PebB lineage, whereas PCYA from chlorophyte algae belongs to the PcyA lineage. PUBS from prasinophyte and streptophyte algae and nonvascular plants belongs to the PebA lineage (Rockwell et al., 2017). Thus, photosynthetic organisms can contain multiple FDBRs producing different reduced bilins, sometimes referred to as phycobilins or phytobilins.

The bilins produced by FDBRs play important roles in the photobiology of oxygenic photosynthetic organisms. Phycobilin chromophores are used by light-harvesting phycobiliproteins in cyanobacteria and eukaryotic glaucophyte, rhodophyte, and cryptophyte algae, frequently as components of large phycobilisome antennae. Phycobilisome-deficient mutant strains of *Synechococcus* sp. PCC 7002 are viable (Alvey et al., 2011). Land plants use phytochromobilin (PΦB, Fig. 1) as the chromophore for phytochrome photoreceptors. Plant phytochromes are critical master regulators for many processes, but plants lacking phytochrome are also viable (Hu et al., 2013). FDBRs would thus not be expected to be essential for oxygenic photosynthesis.

Surprisingly, there is evidence that FDBRs are essential for phototrophic growth. Mutant strains of *Synechococcus* sp. PCC 7002 lacking PcyA could only be obtained in the presence of heterologously expressed HY2 from *Arabidopsis thaliana* (Alvey et al., 2011). PcyA and HY2 produce structurally related phycocyanobilin (PCB) and PΦB, respectively (Fig. 1). There is also evidence that bilin biosynthesis is essential in the chlorophyte alga *Chlamydomonas reinhardtii* (Duanmu et al., 2013), which lacks phytochromes and phycobiliproteins but retains HO and PCYA. A mutant strain lacking *HMOX1,* encoding plastid HO, was unable to grow phototrophically. Addition of exogenous BV afforded partial rescue of *hmox1Δ* phenotypes (Duanmu et al., 2013). These results thus implicate essential functions for bilins in cyanobacteria and green algae.

A requirement for bilins in oxygenic photosynthesis would seem surprising given the intense research into this process for over a century. However, a requirement for bilins in oxygenic photosynthesis implies the presence of HO and one or more FDBRs in all oxygenic photosynthetic organisms, including diverse algae (Fig. 2). Preliminary BLAST searches of algal transcriptomes (Keeling et al., 2014; Matasci et al., 2014) confirmed the presence of FDBRs in many algal lineages. In the current work, we use phylogenetic analysis to examine the distribution and diversification of FDBRs in photosynthetic organisms. The results confirm the presence of FDBRs in all algal lineages, with the possible exception of photosynthetic euglenids, and confirm that eukaryotic FDBRs belong to the three known lineages. Our studies provide new insight into FDBR loss, duplication, and diversification.

METHODS

Construction of FDBR sequence alignments and phylogenies

Additional bacterial and phage FDBR sequences were identified using BLAST (Altschul et al., 1997) searches of the DOE-IMG database and of the Genbank protein and GSS (genomic survey sequences) databases. Eukaryotic FDBR sequences were identified using

BLAST searches of the Genbank protein, EST (expressed sequence tag), TSA (transcriptome shotgun assembly), and SRA (sequence read archive) databases, of publicly available OneKP (one thousand plants, (Matasci et al., 2014)) transcriptomes, and of transcriptomes generated by the Marine Microeukaryotes Transcriptome Sequencing Project (MMETSP, (Keeling et al., 2014)). In some cases, including EST and SRA data, multiple hits were combined to yield a complete FDBR sequence using the approach we previously applied to PCYA from *Cyanophora paradoxa* CCMP329 (Rockwell et al., 2017). Accession information for FDBR sequences used in this paper is provided in Supplemental Table 1.

Sequence alignments constructed with MAFFT (Katoh et al., 2002) were used to identify FDBR sequences with complete or nearly complete catalytic core regions, and an in-house script was used to remove positions with ≥5% gaps and to add secondary structure and solvent accessibility calculated using DSSP (Kabsch and Sander, 1983). Positions that adopted different structural environments in different FDBR crystal structures were assigned as exposed/other for use with the EX_EHO substitution model (Le and Gascuel, 2010). Each resulting alignment was used to infer a structurally informed ML phylogeny using PhyML-structure (Le and Gascuel, 2010) with the Shimodaira-Hasegawa-like approximate likelihood ratio test (SH-aLRT) to assess statistical support (Anisimova and Gascuel, 2006). Command-line settings for MAFFT were (--genafpair --maxiterate 16 --clustalout --reorder). Command-line settings for PhyML-structure were (-m EX_EHO -M PART -a e -c 4 -v e -o tlr). Structures with the PDB accessions 2D1E, 2G18, 2X9O, and 2VCK were used as references. In analyzing phylogenetic relationships within the three FDBR lineages, statistical support is reported in the figures using SH-aLRT for the individual tree and in the text using the percent recovery for the 64 trees that were manually analyzed in lieu of conventional bootstrap analysis.

RESULTS AND DISCUSSION

Alignment and phylogenetic analysis of FDBRs

We used our previous phylogenetic analysis of HY2 from streptophyte algae (Rockwell et al., 2017) as a starting point for constructing new alignments including both eukaryotic sequences and sequences from more diverse bacteria. Alignments were then used to infer structurally informed ML phylogenies using PhyML-structure (Le and Gascuel, 2010). We initially generated 80 trees. These trees had a range of 186–365 sequences and 154–179 characters after removal of gap-enriched positions. Alignments with fewer characters typically included incomplete sequences. We tested two possible outgroups: CPO sequences from cyanobacteria and RCCR sequences from cyanobacteria and land plants. Both of these enzymes share the overall fold of FDBRs, despite limited sequence similarity, and hence provide suitable outgroups for structurally informed phylogenetic analysis (Hagiwara et al., 2006; Sugishima et al., 2009). The RCCR-based outgroup gave more consistent root placement and is similar to that used in our previous FDBR phylogeny (Rockwell et al., 2017).

We next focused on sequences with complete or nearly complete coverage of the catalytic core region, generating 64 trees from alignments having 189–219 sequences and 186–207 characters after removal of gap-enriched positions. These 64 trees were examined manually

at nodes of interest in lieu of conventional bootstrap analysis to avoid misleading results caused by inclusion of particular sequences. The most common topology in these 64 trees (47%) recovered a monophyletic PcyA lineage diverging first relative to the outgroup, with monophyletic PebA and PebB lineages diverging from each other subsequently (Fig. 3). Initial divergence of a monophyletic PebA lineage followed by separation of monophyletic PcyA and PebB lineages was observed in 22% of trees (Fig. 3). These results illustrate the current uncertainty in our understanding of how the three lineages arose during FDBR evolution: the three lineages are probably monophyletic, but we have not reconstructed the relationships among them. The behavior of FDBRs within the three lineages was more robustly recovered. Representative views of these lineages (Fig. 4–6) thus provide new insight into the distribution and diversification of FDBRs during evolution.

FDBRs in bacteria and phages

PcyA is ubiquitous in photosynthetic cyanobacteria and is thought to be essential in *Synechococcus* sp. PCC 7002 (Alvey et al., 2011), whereas PebA and PebB are specifically present in organisms that synthesize phycoerythrobilin (Dammeyer and Frankenberg-Dinkel, 2008). FDBR sequences are absent in the genomes of nonphotosynthetic Melainabacteria, sister to cyanobacteria (Soo et al., 2014), but present in the genomes of *Candidatus* Atelocyanobacterium spp. (also known as UCYN cyanobacteria), photoheterotrophs derived from unicellular nitrogen-fixing photosynthetic cyanobacteria (Bombar et al., 2014). *Ca.* Atelocyanobacterium contained only PcyA, and these sequences were robustly associated with other cyanobacterial PcyA sequences (90% support; Fig. 4). We also found examples of all three FDBRs in the recently reported genomes of cyanobacterial sponge symbionts (*Ca.* Synechococcus spongiarum: (Burgsdorf et al., 2015)). It is well established that marine cyanobacteria with greatly reduced genomes belonging to the genera *Cyanobium, Synechococcus,* and *Prochlorococcus* (hereafter, the CSP clade) form a monophyletic group connected to other cyanobacteria by a characteristically long branch (Shih et al., 2013). Cyanobacterial sponge symbionts have recently been shown to be sister to the CSP clade (Burgsdorf et al., 2015), and indeed sponge symbiont PebA and PebB sequences were sister to PebA and PebB sequences from the CSP clade (100% support; Figs. 5–6).

PcyA sequences from *Ca.* S. spongiarum and from the CSP clade also clustered with phage PcyA sequences, metagenomic sequences from marine and brackish environments, PCYA sequences from photosynthetic *Paulinella* spp., and PCYA sequences from photosynthetic euglenids (Fig. 4). We were able to define two groups within this cluster. One group comprised PcyA from CSP cyanobacteria, PCYA from *Paulinella* spp., PcyA from phages, and PcyA from metagenomic samples. The second comprised PcyA from *Ca.* S. spongiarum and PCYA from euglenids. The (sponge + euglenid) PcyA lineage was sister to (CSP + *Paulinella* + phage + metagenome) lineage (67% support), and phage PcyA sequences were derived from CSP PcyA sequences (72% support). These two lineages together formed a monophyletic clade (the expanded CSP PcyA clade) that was not robustly related to PcyA sequences from other cyanobacteria.

In addition to PcyA, phage genomes can also encode PebS or PcyX (Dammeyer et al., 2008; Ledermann et al., 2016). PebS and PcyX both catalyze the 4-electron reduction of BV to

PEB, but PebS belongs to the PebA lineage and PcyX belongs to the PcyA lineage (Figs. 4–5). PebS sequences formed a monophyletic clade most often (56% support) placed as sister to a clade of FDBRs from phaeophyte algae (phaeophyte PEBZ; Fig. 5 and see below). PcyX sequences also formed a monophyletic clade derived not from cyanobacterial FDBRs but from PcyA sequences from other bacteria, consistent with previous work (Ledermann et al., 2016).

FDBRs from the nonoxygenic photosynthetic bacterium *Bradyrhizobium* sp. ORS278 and the metagenomically sequenced actinobacterium SCGC AAA041-L13 were previously recovered as sister to phage PcyX sequences (Jaubert et al., 2007; Ledermann et al., 2016). *Bradyrhizobium* PcyA produces PCB (Jaubert et al., 2007), so the transition from bacterial PcyA to phage PcyX is a change in regiospecificity. We observed an additional closely related FDBR in the genome of *Roseomonas stagni* DSM 19981 (Rhodospirillales, α-proteobacteria). In both *Bradyrhizobium* sp. ORS278 and *R. stagni*, the *pcyA* gene was associated with a phytochrome gene, a heme oxygenase gene, and a large cluster of gas vesicle genes (Fig. S1). The FDBR from actinobacterium SCGC AAA041-L13 is also adjacent to a heme oxygenase, but phytochrome and gas vesicle genes are apparently absent (Fig. S1). The sequences from *Bradyrhizobium,* actinobacterium SCGC AAA041-L13, and *R. stagni* were recovered as part of a paraphyletic group that were the closest relatives to phage PcyX (Fig. 4), as expected (Ledermann et al., 2016).

We also observed two additional cases of possible FDBRs in nonphotosynthetic bacteria. In the first case, both available genomes for *Chondromyces* spp. (Myxococcales, δ-proteobacteria) encode possible FDBRs. These sequences were typically recovered as sister to the PcyA lineage (52% support) or sister to a combined (PebA + PebB) lineage (30% support, Fig. 7). These sequences could thus belong to a fourth FDBR lineage. *Chondromcyes* FDBRs are adjacent not to HO genes but to genes implicated in corrin metabolism and signal transduction. The function of these proteins is thus unclear. FDBRs may also be present in δ-proteobacterium DG_8 and members of *Candidatus* Schekmanbacteria, all known only from environmental metagenome sequencing. BLAST searches identified FDBRs as the closest known relatives to these sequences, but this relationship was distant. These sequences were most frequently recovered (63%) as sister to all FDBRs in phylogenetic analyses. We were unable to find cognate heme oxygenases in the available genomes for *Ca.* Schekmanbacteria and δ-proteobacterium DG_8. We therefore consider these proteins to be less promising FDBR candidates and chose to place the arbitrary root of the phylogeny between (RCCR + δ-proteobacterium DG_8/*Ca.* Schekmanbacteria sequences) and all other sequences (Fig. 7).

Eukaryotic FDBRs and photosynthesis

BLAST searches identified candidate FDBRs in transcriptomes and/or genomes of every photosynthetic eukaryotic taxon for which data are available. FDBRs were present in 90% of examined algal transcriptomes, and phylogenetic analysis placed all eukaryotic FDBR sequences within the three previously recognized FDBR lineages (Figs. 4–6). Cyanobacterial FDBRs were not found to be basal to eukaryotic sequences in any lineage,

indicating that cyanobacterial sequences have diverged significantly since primary endosymbiosis gave rise to Archaeplastida.

The absence of an FDBR gene in a given transcriptome does not preclude the presence of such a gene in the underlying genome. To date, the only photosynthetic eukaryotic genomes lacking annotated FDBRs are those of *Aureococcus anophagefferens* CCMP1794 and *Chrysochromulina* sp. CCMP291. Strikingly, reexamination of the archived sequence reads for both genomes revealed partial FDBR sequences (Fig. S2). For *A. anophagefferens,* the C-terminus of the candidate FDBR sequence could also be identified in transcriptomes. For *Chrysochromulina* sp. CCMP291, a closely related FDBR was observed in the transcriptome of *C. rotalis* UIO044 (Keeling et al., 2014), confirming that FDBRs are transcribed in *Chrysochromulina* spp. It thus seems plausible that FDBRs are ubiquitous in photosynthetic eukaryotes.

FDBRs are rare in nonphotosynthetic eukaryotes. FDBRs are absent in fungi, and metazoans, and the previously photosynthetic apicomplexans. The cryptophyte genus *Cryptomonas* includes both photosynthetic and nonphotosynthetic strains (Hoef-Emden, 2005). Analysis of transcriptomes relvealed multiple FDBRs in photosynthetic *Cryptomonas curvata* CCAP 979/52 but not in nonphotosynthetic *C. paramecium* CCAP 977/2a. Cryptophytes are members of the Cryptista, and FDBRs were not detected in transcriptomes for the nonphotosynthetic Cryptistans *Roombia truncata* (a katablepharid), *Goniomonas* spp., and *Palpitomonas bilix* (Burki et al., 2016).

Surprisingly, we were able to detect a candidate FDBR sequence in the transcriptome for the nonphotosynthetic chlorophyte *Polytomella parva* SAG 63-3 (Keeling et al., 2014). This sequence was not closely related to PCYA sequences from photosynthetic chlorophytes and was excluded from the detailed phylogenetic analysis. However, incomplete closely related sequences could be found in other *Polytomella* transcriptomes (Fig. S3). *P. parva* SAG 63-3 also retained a candidate ortholog for at least one of the two heme oxygenases found in the *C. reinhardtii* genome. It is thus possible that *Polytomella* spp. have retained a complete phycobilin biosynthesis pathway.

Loss and diversification of FDBRs in Archaeplastida

A single primary endosymbiosis has been proposed to give rise to the Archaeplastida supergroup, comprising glaucophytes, rhodophytes, and Viridiplantae (Price et al., 2012), although monophyly of the Archaeplastida is still under debate (Burki et al., 2016). Glaucophytes retain phycobilisomes (Price et al., 2012). *PCYA* genes were found in all glaucophyte transcriptomes and genomes. These PCYA sequences were sister to a subset of extant cyanobacterial PcyA sequences (100% support, Fig. 4).

Extant rhodophytes also retain phycobiliproteins for light harvesting. PEBA and PEBB genes were found in all rhodophytes examined, with the exception of *Cyanidioschyzon merolae. C. merolae* belongs to the early-diverging Cyanidiales; its genome encodes a *PCYA* gene (Fig. 4). This finding is in contrast to the presence of *PEBA* and *PEBB* genes in the genome of another member of the Cyanidiales, *Galdieria sulphuraria* (Figs. 5–6). PCYA from *C. merolae* was recovered as sister to cyanobacterial PcyA sequences but was not

robustly associated with glaucophyte PCYA enzymes (Fig. 4). Interestingly, PCYA sequences from the cryptophyte algal genera *Chroomonas* and *Hemiselmis* formed a monophyletic group descended from *C. merolae* PCYA with modest (73%) support.

Rhodophyte PEBA sequences formed a paraphyletic group that may be ancestral to a monophyletic group of cryptophyte PEBA sequences (80% support; Fig. 5 and see below). Rhodophyte PEBB enzymes also formed a paraphyletic group that was clearly ancestral to cryptophyte PEBB sequences (100% support; Fig. 6 and see below). In both cases, (rhodophyte + cryptophyte) sequences were closely related to cyanobacterial sequences (86% support for PEBA and 88% support for PEBB).

The third Archaeplastida lineage is the Viridiplantae, comprising the streptophyte lineage and the prasinophyte/chlorophyte lineage (Duanmu et al., 2014; Leliaert et al., 2016). *PCYA* genes were found in prasinophyte and chlorophyte algae but not in streptophytes. Similarly, *HY2* genes were only found in streptophytes. Prasinophyte/chlorophyte PCYA sequences and streptophyte HY2 sequences each formed a monophyletic clade with 75% support (Figs. 4 & 6). The PebA lineage is represented by PUBS in both Viridiplantae lineages (Fig. 5). PUBS sequences from the two Viridiplantae lineages do not form a monophyletic clade. Candidate PUBS sequences from *Mesostigma viride* strains were not closely related to other PUBS sequences (Fig. S4), and complete PUBS sequences could not be obtained for other early-diverging streptophytes. It thus may prove possible that a Viridiplantae PUBS clade will be established in the future.

We also observed an additional FDBR in transcriptomes from the prasinophyte genera *Tetraselmis* and *Pyramimonas,* which we designate PCYZ. Phylogenetic analysis most often (61% support) placed these sequences sister to a clade comprising phage PcyX and non-cyanobacterial PcyA (Fig. 4). Key catalytic residues in cyanobacterial PcyA (Tu et al., 2007) and eukaryotic PCYA (Table 1) were not conserved in this group. Interestingly, we were unable to detect PUBS in *Tetraselmis* and *Pyramimonas,* raising the possibility that PCYZ has functionally replaced PUBS in these prasinophyte genera. Multiple strains clearly contained both PCYA and PCYZ sequences (Fig. 4).

FDBRs in *Paulinella* species

A second primary endosymbiosis has occurred in some species of the testate Rhizarian amoeba *Paulinella* (Lhee et al., 2017). Photosynthetic *Paulinella* species acquired photosynthesis and carbon fixation via endosymbiosis of a member of the CSP clade. The three available *Paulinella* plastid genomes contain *PCYA* genes, although two of the protein sequences were identical (Fig. S5A). *Paulinella* PCYA sequences belonged to the expanded CSP clade (Fig. 4), consistent with the evolutionary history of these algae.

Single-cell genome sequencing of six cells of the nonphotosynthetic *P. ovalis* has been reported (Bhattacharya et al., 2012). Sequencing reads from one of these cells include a possible FDBR sequence (Fig. S5B). This sequence was associated with *Bradyrhizobium* PcyA and related bacterial sequences (83% support; Fig. 4). BLAST searches did not detect related sequences for the other five cells. It remains unclear whether this is sequence is a

bona fide FDBR from *P. ovalis* or instead is a contaminating sequence from a marine bacterium.

FDBRs in chlorarachniophytes

Chlorarachniophytes, another Rhizarian lineage, acquired photosynthesis via endosymbiosis of a chlorophyte alga (Curtis et al., 2012), predicting the presence of PCYA in chlorarachniophytes. However, available chlorarachniophyte transcriptomes contain 1–2 sequences belonging to the PebA lineage (Fig. 5). Two FDBRs are present in the genome of *Bigelowiella natans* CCMP2755 (Curtis et al., 2012), implicating a similar situation in other chlorarachniophytes. Several chlorarachniophyte transcriptomes contained multiple candidate FDBRs. Two groups can be distinguished by examination of candidate catalytic residues (Table 1), but the available complete sequences are too inabundant to conclude that there are two chlorarachniophyte FDBRs.

FDBRs in photosynthetic euglenids

Some euglenids have acquired photosynthesis via endosymbiosis of a prasinophyte alga related to extant Pyramimonadales (Turmel et al., 2009). Euglenids belong to the Excavata, an early-diverging eukaryotic supergroup (Burki et al., 2016). A prasinophyte plastid ancestor would be expected to contain PCYA, and we identified PCYA sequences in euglenid transcriptomes from *Euglena gracilis* and *Eutreptiella* isolates (Ebenezer et al., 2017; Keeling et al., 2014).

Euglenid PCYA sequences belonged to the expanded CSP clade (Fig. 4). This result could indicate that these FDBR sequences are associated with bacterial contaminants. However, closely related sequences are found in all three transcriptomes, so it is possible that these euglenid FDBR sequences could be encoded by previously undetected cyanobacterial symbionts or via acquisition of cyanobacterial FDBR sequences via HGT. Further studies will be required to resolve this question.

FDBRs in haptophytes

Haptophyte algae acquired photosynthesis via secondary endosymbiosis of a rhodophyte alga. Recent work established haptophytes as sister to the nonphotosynthetic centrohelids in the Haptista (Burki et al., 2016), with (Haptista + SAR) sister to (Cryptista + Archaeplastida). We found FDBR sequences in haptophyte transcriptomes but not in centrohelid transcriptomes (Burki et al., 2016; Keeling et al., 2014). Most haptophyte FDBRs were part of a clade belonging to the PebB lineage that also included PEBB sequences from the cryptophyte genus *Hemiselmis* (Fig. 6). This relationship did not hold for other cryptophyte PEBB sequences.

Two haptophyte FDBR sequences from *Pavlova* sp. CCMP459 and from the unclassified prymnesiophyte CCMP2436 were associated with the PebA lineage (Fig. 5). This difference could arise due to variable levels of gene expression and/or transcriptomic coverage. However, a PebA-related gene could not be readily detected in the genome sequence reads for *Chrysochromulina* sp. CCMP291. It is possible that there are differences in tetrapyrrole and bilin biology among haptophytes.

FDBRs in stramenopiles

Stramenopiles (also known as heterokonts) are a diverse eukaryotic group including both nonphotosynthetic and photosynthetic lineages. Together with alveolates and rhizarians, they form a monophyletic assemblage known as the SAR clade (Burki et al., 2016). A stramenopile clade, the Ochrophyta, is proposed to include all photosynthetic species (Derelle et al., 2016; Sevcikova et al., 2015). Such ochrophytes acquired photosynthesis via endosymbiosis of a eukaryotic rhodophyte alga. One would thus expect the presence of *PEBA* and *PEBB* genes, but not *PCYA* genes, in ochrophytes.

As expected, PCYA sequences were not detected in ochrophytes. PEBA genes were widespread. The diatoms *Corethron pennatum* L29A3 and *Proboscia alata* PI_D3 apparently had multiple PEBA transcripts (Fig. 5), although it is not clear whether this arises due to alternative splicing or gene duplication. PEBB sequences were less abundant but were still found in diatoms, phaeophytes, and other ochrophyte lineages (Fig. 6). It is currently unclear whether the relative inabundance of stramenopile PEBB sequences arises due to lower expression levels or loss of the *PEBB* gene. Stramenopile PEBA sequences were not monophyletic, but stramenopile PEBB sequences were monophyletic with modest support (75%).

We also observed two cases in which stramenopile PEBA genes had apparently duplicated and diverged, potentially giving rise to new FDBRs (Fig. 5). In one case, a subset of diatom sequences were shown to form a small clade most frequently recovered (58%) as descended from diatom PEBA sequences. These sequences also had characteristic differences at potential catalytic residues (Table 1), leading us to designate these sequences as diatom PEBX (Fig. 5). In the second case, a group of phaeophyte FDBR sequences formed a clade most frequently recovered as sister to phage PebS sequences (56% support; Fig. 5). Potential catalytic residues in this clade were distinct from those of stramenopile PEBA sequences and diatom PEBX sequences, so we designate these sequences as phaeophyte PEBZ (Table 1).

FDBRs in chromerids

Chromerid algae (Alveolata) are the closest known photosynthetic relatives of the formerly photosynthetic apicomplexans. Two chromerids, *Chromera velia* CCMP2878 and *Vitrella brassicaformis* CCMP3155, have been characterized genomically and transcriptomically (Woo et al., 2015). Chromerids are thought to have acquired photosynthesis via secondary endosymbiosis of a rhodophyte alga or via tertiary endosymbiosis of an ancient ochrophyte (Sevcikova et al., 2015). In either case, one would expect to find PEBA and/or PEBB in chromerids. Both *C. velia* and *V. brassicaformis* contain candidate PEBA sequences (Fig. 5). *C. velia* encodes 2 PEBA genes, whereas *V. brassicaformis* encodes only one. Genomes are available for both organisms, so the difference in the number of FDBR genes is likely to be reliable. However, it is not currently clear whether this difference is significant. *C. velia* and *V. brassicaformis* are not closely related, and additional chromerid diversity may exist (Janouskovec et al., 2012). Future studies may thus provide additional information.

FDBRs in dinoflagellates

Dinoflagellates (Alveolata) are among the most abundant organisms in surface marine waters (Janouskovec et al., 2017). Photosynthetic dinoflagellates initially acquired photosynthesis via endosymbiosis of a rhodophyte, but the resulting secondary plastid has been replaced in some lineages. Dinoflagellates such as *Karenia brevis* and *Karlodinium micrum* carried out tertiary endosymbiosis of a haptophyte, whereas the dinotoms *Durinskia baltica* CSIRO CS-38*, Kryptoperidinium foliaceum* CCMP 1326, and *Glenodinium foliaceum* CCAP 1116/3 carried out tertiary endosymbiosis of a diatom. At least one dinoflagellate, *Noctiluca scintillans,* can live photosynthetically with the assistance of a chlorophyte endosymbiont or can live heterotrophically without it. Other dinoflagellates transiently acquire photosynthesis via kleptoplasty, as in *Dinophysis acuminata*. Phylotranscriptomic studies implicate both *N. scintillans* and *D. acuminata* as having lost functional red-derived plastids in the past (Janouskovec et al., 2017).

We did not observe FDBR sequences in transcriptomes for heterotrophic dinoflagellates such as *Oxyrrhis marina* CCMP1795 and *Crypthecodinium cohnii* Seligo, organisms which retain nonphotosynthetic plastids. FDBR sequences were readily detected in transcriptomes for photosynthetic dinoflagellates. Most such sequences formed a clade in the PebA lineage (Fig. 5), with one such sequence per dinoflagellate transcriptome. *Gymnodinium catenatum* GC744 contained two such sequences, both belonging to the core dinoflagellate FDBR clade. FDBR sequences from the dinotoms *D. baltica* and *K. foliaceum* were instead grouped with PEBA and PEBB sequences from diatoms or ochrophytes (Figs. 5–6) and hence are likely to originate in the diatom endosymbiont. Core dinoflagellate FDBRs were recovered as descended from (69% support) or sister to (30% support) prasinophyte PUBS sequences, so we designate them as PUBS rather than PEBA (Fig. 5).

We also observed FDBR sequences in the nonphotosynthetic *N. scintillans* and *D. acuminata*. The candidate FDBR sequence in *N. scintillans* belonged to the core dinoflagellate PUBS clade. The two FDBR sequences in *D. acuminata* were closely related to those from the cryptophyte alga *Geminigera cryophila* CCMP2564 (Fig. S6). Phylogenetic analysis clearly placed the *D. acuminata* PEBB gene with cryptophyte PEBB sequences (Fig. 6). Plastids from *G. cryophila* are first acquired via kleptoplasty by the ciliate *Myrionecta rubra*, which in turn is subject to kleptoplasty by *D. acuminata* (Wisecaver and Hackett, 2010). FDBR genes are not encoded on cryptophyte plastid genomes reported to date, and *D. acuminata* FDBR sequences are not identical to those from *G. cryophila* (Fig. S6). It seems likely that *D. acuminata* FDBR sequences are derived from cryptophyte prey but incorporated into the *D. acuminata* nuclear genome, as for the cryptophyte-derived gene for PsbM (photosystem II subunit M) in the nuclear genome of *D. acuminata* (Wisecaver and Hackett, 2010).

FDBRs in cryptophytes

Recent work has shown that cryptophytes are members of the Cryptista, along with nonphotosynthetic katablepharids, goniomonads, and *P. bilix* (Burki et al., 2016). Cryptophytes acquired photosynthesis via secondary endosymbiosis of a rhodophyte alga. One would therefore expect cryptophyte algae to contain PEBA and PEBB. The model

cryptophyte *Guillardia theta* CCMP2712 is known to contain a functional PEBB enzyme and a putative PEBA enzyme (Overkamp et al., 2014). The presence of PCYA in the extremophilic rhodophyte *C. merolae* could also implicate the presence of PCYA in cryptophytes.

PEBA was ubiquitous in photosynthetic cryptophytes but absent in at least one nonphotosynthetic isolate of *Cryptomonas* (see above). Cryptophyte PEBA sequences were closely related to rhodophyte PEBA sequences (Fig. 5). PEBB was also widespread in photosynthetic cryptophytes. However, cryptophyte PEBB sequences were not monophyletic: PEBB sequences from the genus *Hemiselmis* were part of a clade also including haptophyte PEBB sequences, whereas other cryptophyte PEBB sequences were descended from rhodophyte PEBB sequences as expected (Fig. 6). Uniquely, PEBB from *Chroomonas mesostigmatica* CCMP1168 was present as a tandem fusion with a PEBA gene (Fig. S7). Both *C. mesostigmatica* PEBA sequences were part of the cryptophyte PEBA clade, and *C. mesostigmatica* PEBB was part of the cryptophyte PEBB clade (Figs. 5–6).

Transcriptomes from *C. mesostigmatica* and *Hemiselmis* spp. also contained PCYA sequences not found in other cryptophytes. These sequences were recovered as a monophyletic clade (Fig. 4) that was most often (73%) recovered as descended from the rhodophyte PCYA sequence from *C. merolae.* The evolutionary relationship between PCYA enzymes in cryptophytes and in *C. merolae* raises the possibility that the ancestral rhodophyte giving rise to the cryptophyte plastid retained all three cyanobacterial FDBRs, with most modern rhodophytes losing *PCYA* genes at a later date. Characterization of additional members of the Cyanidiales will be important to address this ambiguity.

Future research on bilin biosynthesis

Even the great expansion in available data for eukaryotic algae has left us with an incomplete picture of several key transitions in algal evolution. We have very few PCYA and PUBS sequences from early-diverging prasinophyte algae, and even fewer PUBS sequences from early-diverging streptophyte algae. We also do not have any information on FDBRs from Palmophyllales (Leliaert et al., 2016). Transcriptomic or genomic data are available for only a few chromerids, glaucophytes, and euglenids, and FDBR sequences have not been reported at all for more recently described secondarily photosynthetic lineages (Choi et al., 2017; Janouskovec et al., 2012). It will thus be interesting to revisit this analysis as more data become available.

Our work highlights a disparity between the FDBRs that have been experimentally characterized and the actual diversity of known FDBRs. To date, approximately sixteen FDBRs have been characterized using *in vitro* enzymology and/or heterologous reconstitution of bilin biosynthesis. Of these enzymes, only one was found in a secondarily photosynthetic alga (Overkamp et al., 2014). Moreover, recent work has shown that closely related FDBRs can carry out different reactions (Rockwell et al., 2017). Our current study provides a starting point in remedying this deficiency by illustrating the extent to which FDBR diversity remains undersampled experimentally.

We have also identified several potentially new FDBRs in algae. Diatom PEBX is apparently a duplication of an ancestral PEBA gene, but changes in potential catalytic residues (Table 1) implicate a different enzymatic function. Phaeophyte PEBZ presents a similar case, but PEBZ is not robustly derived from phaeophyte PEBA sequences and its origins are less clear. The presence of PCYZ only in *Pyramimonas* and *Tetraselmis* spp. raises several questions. These genera are not closely related (Duanmu et al., 2014), and PUBS could not be detected in transcriptomes for either genus. This situation could arise through loss of PCYZ or PUBS after early emergence of *PCYZ* in prasinophytes or through later emergence of *PCYZ* in one of these two genera with subsequent transfer to the other.

Our studies support the hypothesis that bilin biosynthesis is ubiquitous in oxygenic photosynthetic organisms. With the possible exception of euglenids, all photosynthetic eukaryotes apparently contain FDBRs. Euglenid PCYA sequences clade with cyanobacterial PcyA sequences in the expanded CSP clade and hence may arise from bacterial contaminants or symbionts. However, these sequences could also have been acquired from a cyanobacterial genome independently of plastid acquisition. It is even possible that a cyanobacterial symbiont could supply bilin to euglenids *in trans* during phototrophic growth. Our studies provide a new palette of FDBRs for understanding the diversity and ubiquity of bilins in photosynthetic eukaryotes and presage similar analysis of heme oxygenase, the other enzyme in the bilin biosynthesis pathway.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by NIH grant R01 GM068552 (to J. C. L.). The authors thank Kevin Rockwell for assistance in constructing *M. viride* PUBS sequences from SRA data.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res. 1997; 25(17): 3389–3402. [PubMed: 9254694]

Alvey RM, Biswas A, Schluchter WM, Bryant DA. Effects of modified Phycobilin biosynthesis in the Cyanobacterium *Synechococcus* sp Strain PCC 7002. J Bacteriol. 2011; 193(7):1663–1671. [PubMed: 21296968]

Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol. 2006; 55(4):539–552. [PubMed: 16785212]

Bhattacharya D, Price DC, Yoon HS, Yang EC, Poulton NJ, Andersen RA, Das SP. Single cell genome analysis supports a link between phagotrophy and primary plastid endosymbiosis. Sci Rep. 2012; 2:356. [PubMed: 22493757]

Bombar D, Heller P, Sanchez-Baracaldo P, Carter BJ, Zehr JP. Comparative genomics reveals surprising divergence of two closely related strains of uncultivated UCYN-A cyanobacteria. ISME J. 2014; 8(12):2530–2542. [PubMed: 25226029]

Burgsdorf I, Slaby BM, Handley KM, Haber M, Blom J, Marshall CW, Gilbert JA, Hentschel U, Steindler L. Lifestyle evolution in cyanobacterial symbionts of sponges. mBio. 2015; 6(3):e00391–00315. [PubMed: 26037118]

Burki F, Kaplan M, Tikhonenkov DV, Zlatogursky V, Minh BQ, Radaykina LV, Smirnov A, Mylnikov AP, Keeling PJ. Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. Proc Royal Soc B Biol Sci. 2016; 283(1823)

Choi CJ, Bachy C, Jaeger GS, Poirier C, Sudek L, Sarma VV, Mahadevan A, Giovannoni SJ, Worden AZ. Newly discovered deep-branching marine plastid lineages are numerically rare but globally distributed. Curr Biol. 2017; 27(1):R15–R16. [PubMed: 28073013]

Curtis BA, Tanifuji G, Burki F, Gruber A, Irimia M, Maruyama S, Arias MC, Ball SG, Gile GH, Hirakawa Y, Hopkins JF, Kuo A, Rensing SA, Schmutz J, Symeonidi A, Elias M, Eveleigh RJ, Herman EK, Klute MJ, Nakayama T, Obornik M, Reyes-Prieto A, Armbrust EV, Aves SJ, Beiko RG, Coutinho P, Dacks JB, Durnford DG, Fast NM, Green BR, Grisdale CJ, Hempel F, Henrissat B, Hoppner MP, Ishida K, Kim E, Koreny L, Kroth PG, Liu Y, Malik SB, Maier UG, McRose D, Mock T, Neilson JA, Onodera NT, Poole AM, Pritham EJ, Richards TA, Rocap G, Roy SW, Sarai C, Schaack S, Shirato S, Slamovits CH, Spencer DF, Suzuki S, Worden AZ, Zauner S, Barry K, Bell C, Bharti AK, Crow JA, Grimwood J, Kramer R, Lindquist E, Lucas S, Salamov A, McFadden GI, Lane CE, Keeling PJ, Gray MW, Grigoriev IV, Archibald JM. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature. 2012; 492(7427):59–65. [PubMed: 23201678]

Dammeyer T, Bagby SC, Sullivan MB, Chisholm SW, Frankenberg-Dinkel N. Efficient phage-mediated pigment biosynthesis in oceanic cyanobacteria. Curr Biol. 2008; 18(6):442–448. [PubMed: 18356052]

Dammeyer T, Frankenberg-Dinkel N. Function and distribution of bilin biosynthesis enzymes in photosynthetic organisms. Photochemical and Photobiological Sciences. 2008; 7(10):1121–1130. [PubMed: 18846276]

Derelle R, Lopez-Garcia P, Timpano H, Moreira D. A Phylogenomic Framework to Study the Diversity and Evolution of Stramenopiles (=Heterokonts). Mol Biol Evol. 2016; 33(11):2890–2898. [PubMed: 27512113]

Duanmu D, Bachy C, Sudek S, Wong CH, Jimenez V, Rockwell NC, Martin SS, Ngan CY, Reistetter EN, van Baren MJ, Price DC, Wei CL, Reyes-Prieto A, Lagarias JC, Worden AZ. Marine algae and land plants share conserved phytochrome signaling systems. Proc Natl Acad Sci USA. 2014; 111:15827–15832. [PubMed: 25267653]

Duanmu D, Casero D, Dent RM, Gallaher S, Yang W, Rockwell NC, Martin SS, Pellegrini M, Niyogi KK, Merchant SS, Grossman AR, Lagarias JC. Retrograde bilin signaling enables *Chlamydomonas* greening and phototrophic survival. Proc Natl Acad Sci USA. 2013; 110(9): 3621–3626. [PubMed: 23345435]

Ebenezer TE, Carrington M, Lebert M, Kelly S, Field MC. Euglena gracilis Genome and Transcriptome: Organelles, Nuclear Genome Assembly Strategies and Initial Features. Adv Exp Med Biol. 2017; 979:125–140. [PubMed: 28429320]

Hagiwara Y, Sugishima M, Takahashi Y, Fukuyama K. Crystal structure of phycocyanobilin:ferredoxin oxidoreductase in complex with biliverdin IXalpha, a key enzyme in the biosynthesis of phycocyanobilin. Proc Natl Acad Sci U S A. 2006; 103(1):27–32. [PubMed: 16380422]

Hoef-Emden K. Multiple independent losses of photosynthesis and differing evolutionary rates in the genus Cryptomonas (Cryptophyceae): combined phylogenetic analyses of DNA sequences of the nuclear and the nucleomorph ribosomal operons. J Mol Evol. 2005; 60(2):183–195. [PubMed: 15785847]

Hu W, Franklin KA, Sharrock RA, Jones MA, Harmer SL, Lagarias JC. Unanticipated regulatory roles for *Arabidopsis* phytochromes revealed by null mutant analysis. Proc Natl Acad Sci USA. 2013; 110(4):1542–1547. [PubMed: 23302690]

Janouskovec J, Gavelis GS, Burki F, Dinh D, Bachvaroff TR, Gornik SG, Bright KJ, Imanian B, Strom SL, Delwiche CF, Waller RF, Fensome RA, Leander BS, Rohwer FL, Saldarriaga JF. Major transitions in dinoflagellate evolution unveiled by phylotranscriptomics. Proc Natl Acad Sci U S A. 2017; 114(2):E171–E180. [PubMed: 28028238]

Janouskovec J, Horak A, Barott KL, Rohwer FL, Keeling PJ. Global analysis of plastid diversity reveals apicomplexan-related lineages in coral reefs. Curr Biol. 2012; 22(13):R518–519. [PubMed: 22789997]

Jaubert M, Lavergne J, Fardoux J, Hannibal L, Vuillet L, Adriano JM, Bouyer P, Pignol D, Giraud E, Vermeglio A. A singular bacteriophytochrome acquired by lateral gene transfer. J Biol Chem. 2007; 282(10):7320–7328. [PubMed: 17218312]

Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983; 22(12):2577–2637. [PubMed: 6667333]

Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002; 30(14):3059–3066. [PubMed: 12136088]

Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, Armbrust EV, Archibald JM, Bharti AK, Bell CJ, Beszteri B, Bidle KD, Cameron CT, Campbell L, Caron DA, Cattolico RA, Collier JL, Coyne K, Davy SK, Deschamps P, Dyhrman ST, Edvardsen B, Gates RD, Gobler CJ, Greenwood SJ, Guida SM, Jacobi JL, Jakobsen KS, James ER, Jenkins B, John U, Johnson MD, Juhl AR, Kamp A, Katz LA, Kiene R, Kudryavtsev A, Leander BS, Lin S, Lovejoy C, Lynn D, Marchetti A, McManus G, Nedelcu AM, Menden-Deuer S, Miceli C, Mock T, Montresor M, Moran MA, Murray S, Nadathur G, Nagai S, Ngam PB, Palenik B, Pawlowski J, Petroni G, Piganeau G, Posewitz MC, Rengefors K, Romano G, Rumpho ME, Rynearson T, Schilling KB, Schroeder DC, Simpson AG, Slamovits CH, Smith DR, Smith GJ, Smith SR, Sosik HM, Stief P, Theriot E, Twary SN, Umale PE, Vaulot D, Wawrik B, Wheeler GL, Wilson WH, Xu Y, Zingone A, Worden AZ. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLoS Biol. 2014; 12(6):e1001889. [PubMed: 24959919]

Le SQ, Gascuel O. Accounting for solvent accessibility and secondary structure in protein phylogenetics is clearly beneficial. Syst Biol. 2010; 59(3):277–287. [PubMed: 20525635]

Ledermann B, Beja O, Frankenberg-Dinkel N. New biosynthetic pathway for pink pigments from uncultured oceanic viruses. Environ Microbiol. 2016

Leliaert F, Tronholm A, Lemieux C, Turmel M, DePriest MS, Bhattacharya D, Karol KG, Fredericq S, Zechman FW, Lopez-Bautista JM. Chloroplast phylogenomic analyses reveal the deepest-branching lineage of the Chlorophyta, Palmophyllophyceae class. nov. Scientific reports. 2016; 6:25367. [PubMed: 27157793]

Lhee D, Yang EC, Kim JI, Nakayama T, Zuccarello G, Andersen RA, Yoon HS. Diversity of the Photosynthetic Paulinella Species, with the Description of Paulinella micropora sp nov and the Chromatophore Genome Sequence for strain KR01. Protist. 2017; 168(2):155–170. [PubMed: 28262587]

Matasci N, Hung LH, Yan Z, Carpenter EJ, Wickett NJ, Mirarab S, Nguyen N, Warnow T, Ayyampalayam S, Barker M, Burleigh JG, Gitzendanner MA, Wafula E, Der JP, dePamphilis CW, Roure B, Philippe H, Ruhfel BR, Miles NW, Graham SW, Mathews S, Surek B, Melkonian M, Soltis DE, Soltis PS, Rothfels C, Pokorny L, Shaw JA, DeGironimo L, Stevenson DW, Villarreal JC, Chen T, Kutchan TM, Rolf M, Baucom RS, Deyholos MK, Samudrala R, Tian Z, Wu X, Sun X, Zhang Y, Wang J, Leebens-Mack J, Wong GK. Data access for the 1,000 Plants (1KP) project. GigaScience. 2014; 3:17. [PubMed: 25625010]

Overkamp KE, Gasper R, Kock K, Herrmann C, Hofmann E, Frankenberg-Dinkel N. Insights into the biosynthesis and assembly of cryptophycean phycobiliproteins. J Biol Chem. 2014; 289(39): 26691–26707. [PubMed: 25096577]

Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber AP, Schwacke R, Gross J, Blouin NA, Lane C, Reyes-Prieto A, Durnford DG, Neilson JA, Lang BF, Burger G, Steiner JM, Loffelhardt W, Meuser JE, Posewitz MC, Ball S, Arias MC, Henrissat B, Coutinho PM, Rensing SA, Symeonidi A, Doddapaneni H, Green BR, Rajah VD, Boore J, Bhattacharya D. *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science. 2012; 335(6070):843–847. [PubMed: 22344442]

Rockwell NC, Martin SS, Li FW, Mathews S, Lagarias JC. The phycocyanobilin chromophore of streptophyte algal phytochromes is synthesized by HY2. New Phytol. 2017 in press.

Sevcikova T, Horak A, Klimes V, Zbrankova V, Demir-Hilton E, Sudek S, Jenkins J, Schmutz J, Pribyl P, Fousek J, Vlcek C, Lang BF, Obornik M, Worden AZ, Elias M. Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte? Scientific reports. 2015; 5:10134. [PubMed: 26017773]

Shih PM, Wu D, Latifi A, Axen SD, Fewer DP, Talla E, Calteau A, Cai F, Tandeau de Marsac N, Rippka R, Herdman M, Sivonen K, Coursin T, Laurent T, Goodwin L, Nolan M, Davenport KW, Han CS, Rubin EM, Eisen JA, Woyke T, Gugger M, Kerfeld CA. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. Proc Natl Acad Sci USA. 2013; 110(3):1053–1058. [PubMed: 23277585]

Soo RM, Skennerton CT, Sekiguchi Y, Imelfort M, Paech SJ, Dennis PG, Steen JA, Parks DH, Tyson GW, Hugenholtz P. An expanded genomic representation of the phylum cyanobacteria. Genome Biol Evol. 2014; 6(5):1031–1045. [PubMed: 24709563]

Spang A, Saw JH, Jorgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJ. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature. 2015; 521(7551):173–179. [PubMed: 25945739]

Sugishima M, Kitamori Y, Noguchi M, Kohchi T, Fukuyama K. Crystal structure of red chlorophyll catabolite reductase: enlargement of the ferredoxin-dependent bilin reductase family. J Mol Biol. 2009; 389(2):376–387. [PubMed: 19374909]

Tu SL, Rockwell NC, Lagarias JC, Fisher AJ. Insight into the radical mechanism of phycocyanobilin-ferredoxin oxidoreductase (PcyA) revealed by X-ray crystallography and biochemical measurements. Biochemistry. 2007; 46(6):1484–1494. [PubMed: 17279614]

Turmel M, Gagnon MC, O'Kelly CJ, Otis C, Lemieux C. The chloroplast genomes of the green algae Pyramimonas, Monomastix, and Pycnococcus shed new light on the evolutionary history of prasinophytes and the origin of the secondary chloroplasts of euglenids. Mol Biol Evol. 2009; 26(3):631–648. [PubMed: 19074760]

Wisecaver JH, Hackett JD. Transcriptome analysis reveals nuclear-encoded proteins for the maintenance of temporary plastids in the dinoflagellate Dinophysis acuminata. BMC Genom. 2010; 11:366.

Woo YH, Ansari H, Otto TD, Klinger CM, Kolisko M, Michalek J, Saxena A, Shanmugam D, Tayyrov A, Veluchamy A, Ali S, Bernal A, del Campo J, Cihlar J, Flegontov P, Gornik SG, Hajduskova E, Horak A, Janouskovec J, Katris NJ, Mast FD, Miranda-Saavedra D, Mourier T, Naeem R, Nair M, Panigrahi AK, Rawlings ND, Padron-Regalado E, Ramaprasad A, Samad N, Tomcala A, Wilkes J, Neafsey DE, Doerig C, Bowler C, Keeling PJ, Roos DS, Dacks JB, Templeton TJ, Waller RF, Lukes J, Obornik M, Pain A. Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. eLife. 2015; 4:e06974. [PubMed: 26175406]

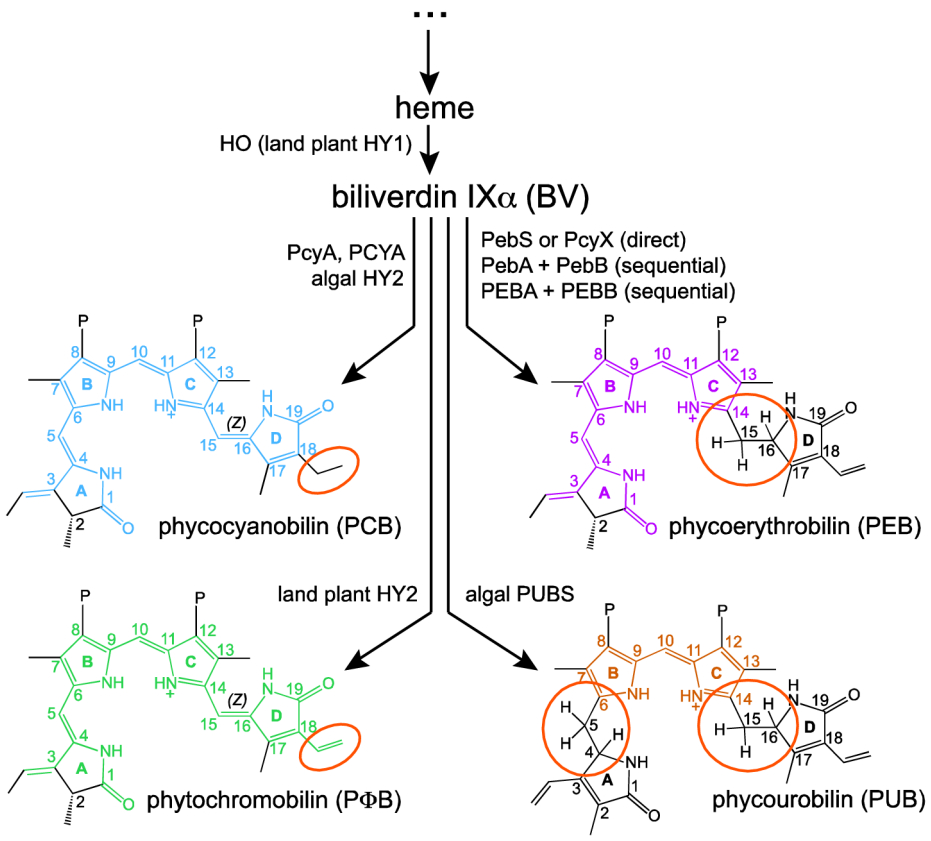

Figure 1. Bilin biosynthesis by ferredoxin-dependent bilin reductases (FDBRs)

After initial breakdown of heme by heme oxygenase (encoded by the *HY1* gene in streptophytes) to yield biliverdin IXα (BV), different FDBRs produce different bilins. PcyA and HY2 from streptophyte algae reduce BV to phycocyanobilin (PCB) in a 4-electron reduction, whereas HY2 from land plants reduces BV to phytochromobilin (PΦB) in a 2-electron reduction. PebA and PebB reduce BV to phycoerythrobilin (PEB) in a 2-enzyme, 4-electron reduction. The same reaction can be carried out by PebS or PcyX as 1-enzyme reactions. PUBS from Viridiplantae reduces BV to phycourobilin (PUB) in a 4-electron reduction. The chromophoric conjugated π systems of different bilins are colored, approximately matching their apparent visual colors. Orange, moieties whose structure differs in different bilins. P, propionate.

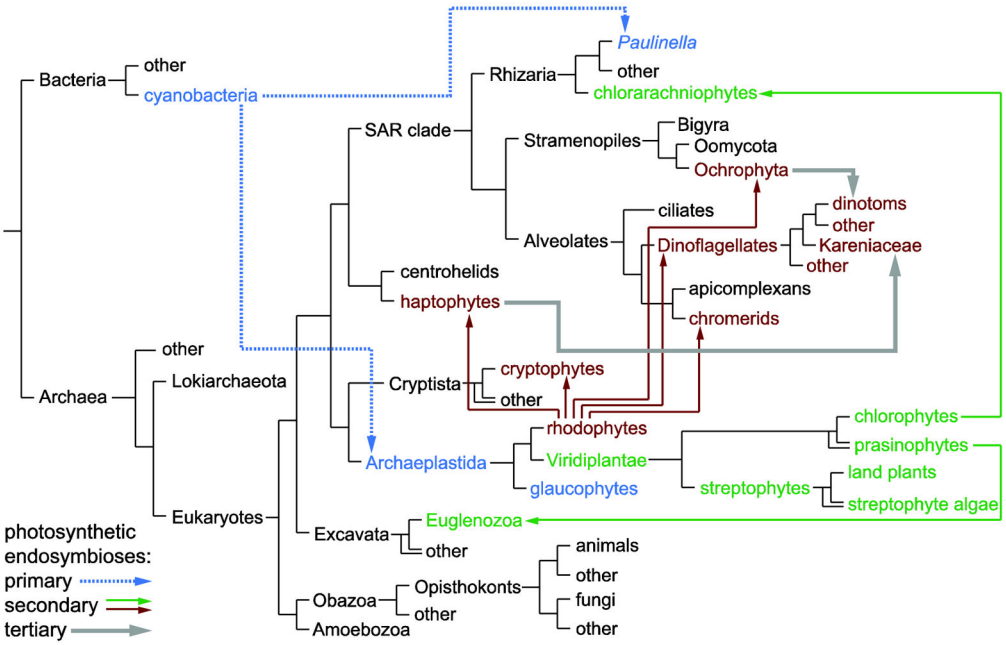

Figure 2. Evolution of photosynthetic eukaryotes

A simplified view of the tree of life is shown, based on recent studies and assuming a monophyletic Archaeplastida for simplicity (Burki et al., 2016; Price et al., 2012; Spang et al., 2015). Primary (dashed blue lines), secondary (thin solid lines), and tertiary (thick grey lines) endosymbioses are shown. Oxygenic photosynthetic organisms are color-coded by light-harvesting strategy and plastid ancestry. Only endosymbioses resulting in creation of a photosynthetic, carbon-fixing organelle are shown.

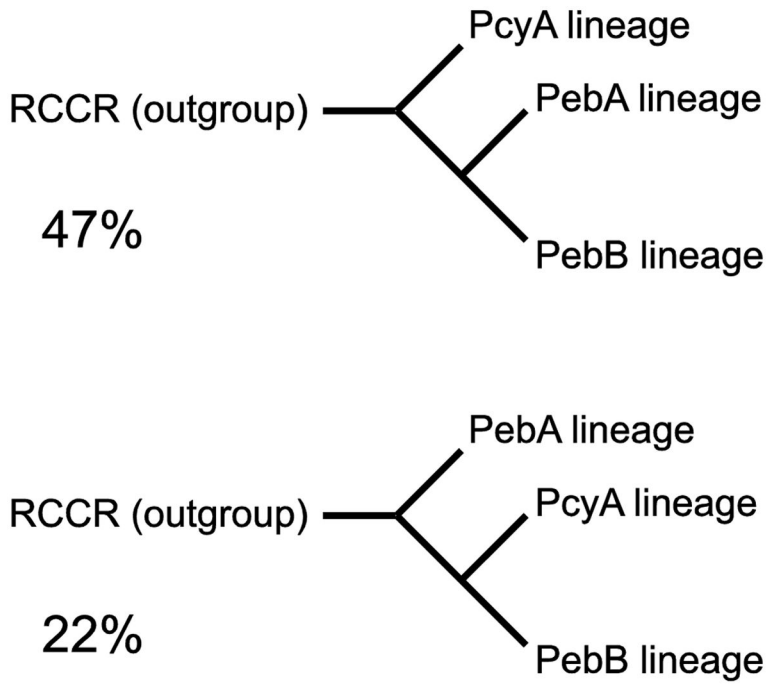

Figure 3. Phylogenetic analysis of FDBR evolution

Divergence of the three FDBR lineages relative to the RCCR outgroup is shown for the two topologies most frequently recovered in the final set of 64 phylogenies.

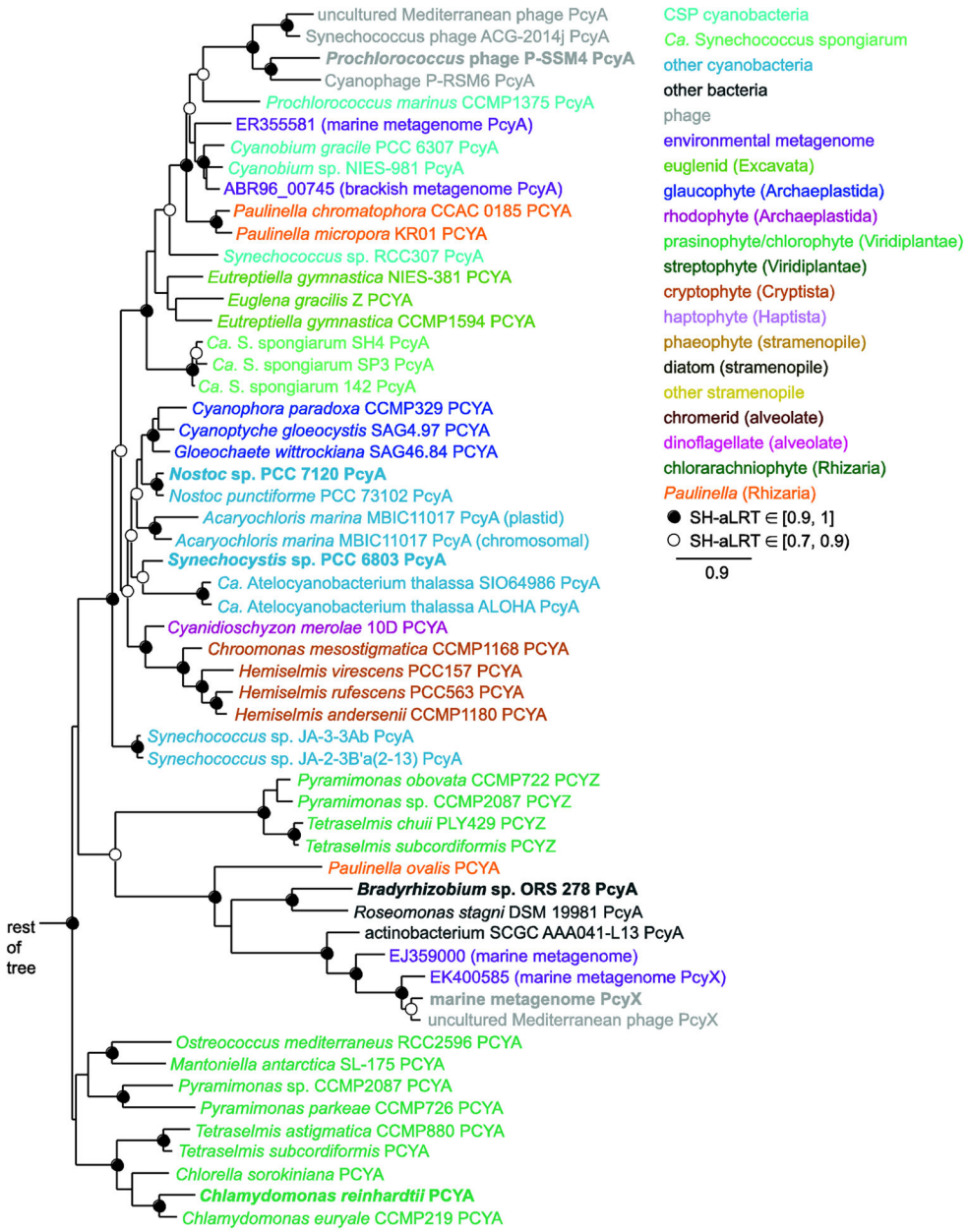

Figure 4. Diversification of the PcyA lineage

A detailed view of a phylogenetic analysis of members of the PcyA lineage is shown, with sequences color-coded by origin. Circles denote SH-aLRT confidence (black, SH-aLRT ≥0.9; white, 0.9 >SH-aLRT ≥0.7. Experimentally characterized FDBRs are in bold.

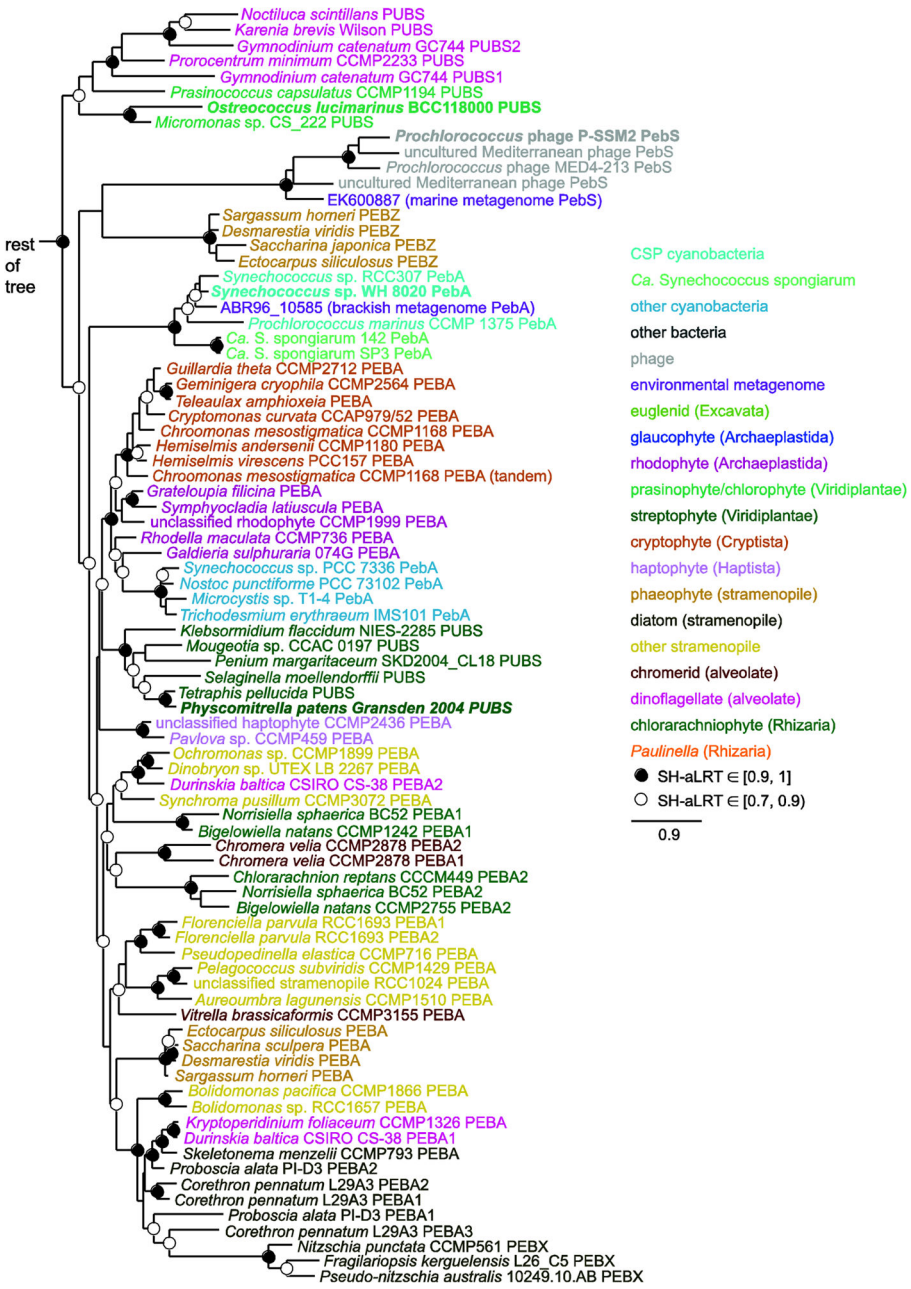

Figure 5. Diversification of the PebA lineage

A detailed view of a phylogenetic analysis of members of the PebA lineage is shown, with sequences color-coded by origin. Circles denote SH-aLRT confidence (black, SH-aLRT ≥0.9; white, 0.9 >SH-aLRT ≥0.7. Experimentally characterized FDBRs are in bold.

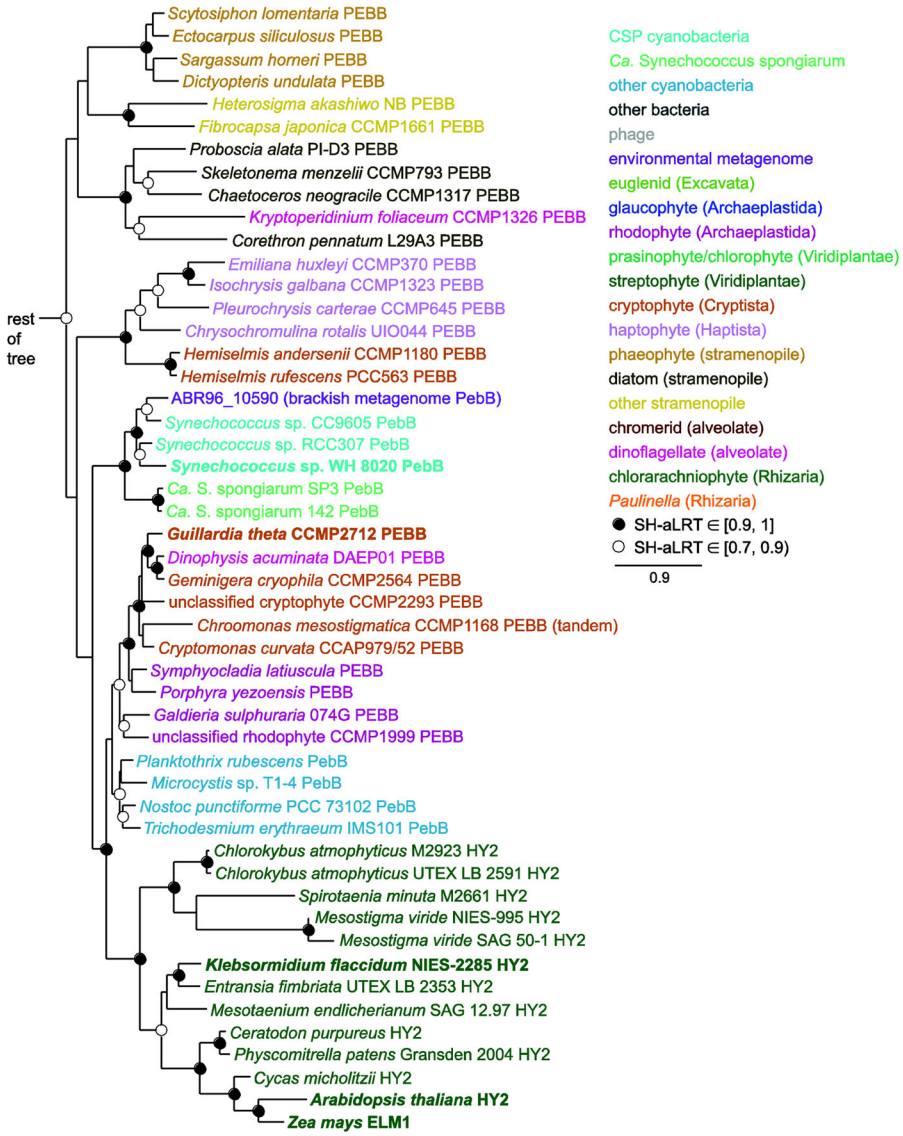

Figure 6. Diversification of the PebB lineage

A detailed view of a phylogenetic analysis of members of the PebB lineage is shown, with sequences color-coded by origin. Circles denote SH-aLRT confidence (black, SH-aLRT ≥0.9; white, 0.9 >SH-aLRT ≥0.7. Experimentally characterized FDBRs are in bold.

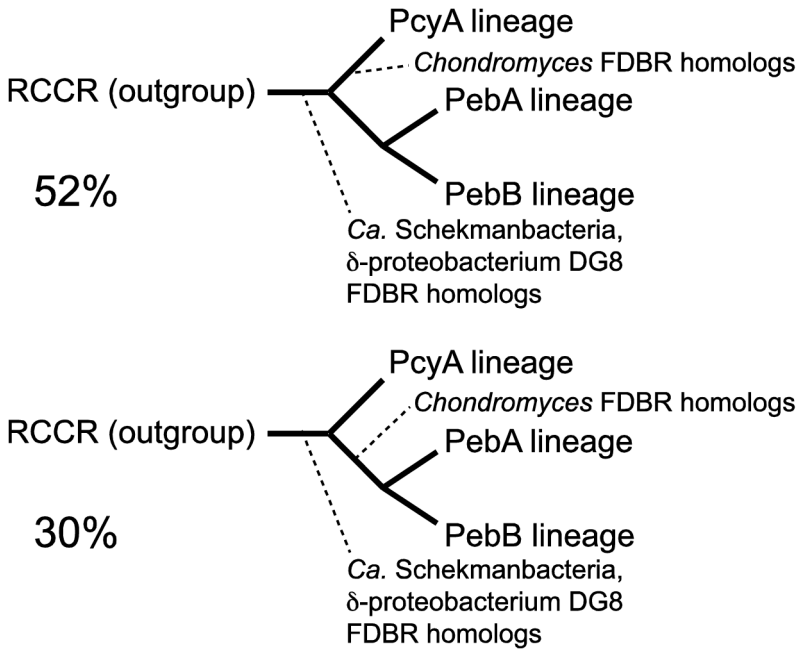

Figure 7. Phylogenetic analysis of possible novel FDBR lineages

The most frequently observed topology for overall FDBR evolution (Fig. 3) is shown with the most likely placements of candidate FDBRs from *Chondromyces* spp., *Ca.* Schekmanbacteria, and δ-proteobacterium DG_8. Percentages indicate support for placement of *Chondromyces* sequences. This placement of sequences from *Ca.* Schekmanbacteria and δ-proteobacterium DG_8 was observed in 63% of phylogenies.

Table 1

Potential catalytic residues of FDBRs[1]

Taxon	Enzyme	H_{71}	E_{73}	H_{85}	D_{102}	D_{217}
cyanobacteria	PcyA	H	E	H	D	D
glaucophytes	PCYA	H	E	H	D	D
Cyanidioschyzon merolae	PCYA	H	E	H	D	D
cryptophytes	PCYA	H	E	H	D	D/E
euglenids	PCYA	H	E	H	D	D
photosynthetic *Paulinella*	PCYA	H	E	H	D	D
P. ovalis	PCYA	H	E	H	D	E
prasinophyte/chlorophyte	PCYA	H	E	H	D	varies
prasinophyte/chlorophyte	PCYZ	T/P	V/A	L	E	K
bacteria	PcyA	H	E/D	H	D	E/D/P
phage	PcyA	H	E	H	D	D/K
phage	PcyX	H	D	H	D	P
phage	PebS	R	A/V/T	N	D	D
phaeophytes	PEBZ	R	S	S	S	E
diatoms	PEBX	R	T	R	G	E
prasinophytes	PUBS	R	T	N	D	D
dinoflagellates	PUBS	R	T/S	N	D/E	D
streptophyte	PUBS	R	T	N	D	D
chlorarachniophytes	PEBA1	R	E	N	D	D
chlorarachniophytes	PEBA2	R	T	N	N	D
chromerids	PEBA	R	T	N	D	D
stramenopile	PEBA	R	T/L	N	D	D
haptophytes	PEBA	R	T	N	D	D
cyanobacteria	PebA	R	T/S	N	D	D
cryptophytes	PEBA	R	T	N	D	D
rhodophytes	PEBA	R	T	N	D	D

Taxon	Enzyme	H_{71}	E_{73}	H_{85}	D_{102}	D_{217}
cyanobacteria	PebB	R	A	N	D	D
rhodophytes	PEBB	R	A	N	D	D
cryptophytes	PEBB	R	A/C	N	D	D
streptophytes	HY2	R	C/M/L/A	N/D	D/N	D
haptophytes, *Hemiselmis*	PEBB	R	A/L	N	D	D
stramenopiles	PEBB	R	A/L/C	N	D	D
Chondromyces spp.	FDBR	R	V	N	D	D

[1] Based on mutagenesis studies reported in (Tu et al., 2007).