RESEARCH PAPER

Mobile Genetic Elements 3, e27313; September/October 2013; © 2013 Landes Bioscience

Foreign gene recruitment to the fatty acid biosynthesis pathway in diatoms

Cheong Xin Chan,[1,*] Francesca L Baglivi,[2] Christina E Jenkins,[2,†] and Debashish Bhattacharya[3]

[1]Institute for Molecular Bioscience, and ARC Centre of Excellence in Bioinformatics; The University of Queensland; Brisbane, QLD Australia; [2]Department of Biology; University of Iowa; Iowa City, IA USA; [3]Department of Ecology, Evolution and Natural Resources, and Institute of Marine and Coastal Sciences; Rutgers University; New Brunswick, NJ USA

[†]Current affiliation: School of Biological Sciences; Washington State University; Pullman, WA USA

Keywords: diatoms, fatty acid biosynthesis, horizontal gene transfer, lipid metabolism, phylogenomics, endosymbiosis

Diatoms are highly successful marine and freshwater algae that contribute up to 20% of global carbon fixation. These species are leading candidates for biofuel production owing to ease of culturing and high fatty acid content. To assist in strain improvement and downstream applications for potential use as a biofuel, it is important to understand the evolution of lipid biosynthesis in diatoms. The evolutionary history of diatoms is however complicated by likely multiple endosymbioses involving the capture of foreign cells and horizontal gene transfer into the host genome. Using a phylogenomic approach, we assessed the evolutionary history of 12 diatom genes putatively encoding functions related to lipid biosynthesis. We found evidence of gene transfer likely from a green algal source for seven of these genes, with the remaining showing either vertical inheritance or evolutionary histories too complicated to interpret given current genome data. The functions of horizontally transferred genes encompass all aspects of lipid biosynthesis (initiation, biosynthesis, and desaturation of fatty acids) as well as fatty acid elongation, and are not restricted to plastid-targeted proteins. Our findings demonstrate that the transfer, duplication, and subfunctionalization of genes were key steps in the evolution of lipid biosynthesis in diatoms and other photosynthetic eukaryotes. This target pathway for biofuel research is highly chimeric and surprisingly, our results suggest that research done on related genes in green algae may have application to diatom models.

Introduction

Diatoms are one of the most common phytoplankton in aquatic environments, with an estimated diversity of 100000 species.[1] These free-living, unicellular primary producers provide oxygen via photosynthesis and are crucial for regulating the biogeochemical cycle of silicon.[2] The hard siliceous structure (frustule) surrounding diatoms has been utilized in bionanotechnology applications[3] as a cross-linking agent for active biomolecules; e.g., in immunoprecipitation.[4] In recent years, diatoms have also been targeted for biofuel production owing to the ease of mass culturing and their high fat content.[5] Given their ecological and potential economic value, it is important to understand the evolution of fatty acid biosynthesis in diatoms. These data may also assist in experimental and industrial designs of downstream biofuel applications relying on these organisms or the products that their genes encode.

Fatty acids are used in many essential cellular processes, from energy production (e.g., triglycerides),[6] to the synthesis of membrane (e.g., phospholipid)[7] and hormones.[8] Fatty acid biosynthesis (FAB) occurs in all living organisms, with the exception of some parasites that have highly reduced genomes.[9] There are two main types of FAB systems: Type I that commonly utilize a single multifunctional protein complex and Type II that involves multiple monofunctional enzymes. Both of these FAB systems are present in prokaryotes and plastid-lacking eukaryotes. In animals and fungi, the Type I system important for palmitate synthesis is cytosolic, whereas Type II is involved in the production of eight-carbon chains within the mitochondrial matrix.[10,11] In comparison, photosynthetic eukaryotes, e.g., plants, algae and the polyphyletic group of photosynthetic protists often referred to as "chromalveolates" (including diatoms) use only Type II FAB[12] specifically within the plastid (**Fig. 1**).

The origin and evolution of the plastid is explained by endosymbiosis, during which genes are transferred from the engulfed endosymbiont to the host genome.[13,14] Secondary endosymbiosis that involves the engulfment of an existing (in this case, a red) alga is thought to be the landmark event that gave rise to photosynthesis in "chromalveolates" (e.g., diatoms, dinoflagellates, cryptophytes, haptophytes), and the complex plastid structure (e.g., 3–4 bounding membranes, remnant nucleomorph in cryptophytes) in these photosynthetic lineages.[14] Nevertheless, a

*Correspondence to: Cheong Xin Chan; Email: c.chan@imb.uq.edu.au
Submitted: 09/13/2013; Revised: 10/23/2013; Accepted: 11/21/2013
Citation: Chan CX, Baglivi FL, Jenkins CE, Bhattacharya D. Foreign gene recruitment to the fatty acid biosynthesis pathway in diatoms. Mobile Genetic Elements 2013; 3:e27313; http://dx.doi.org/10.4161/mge.27313

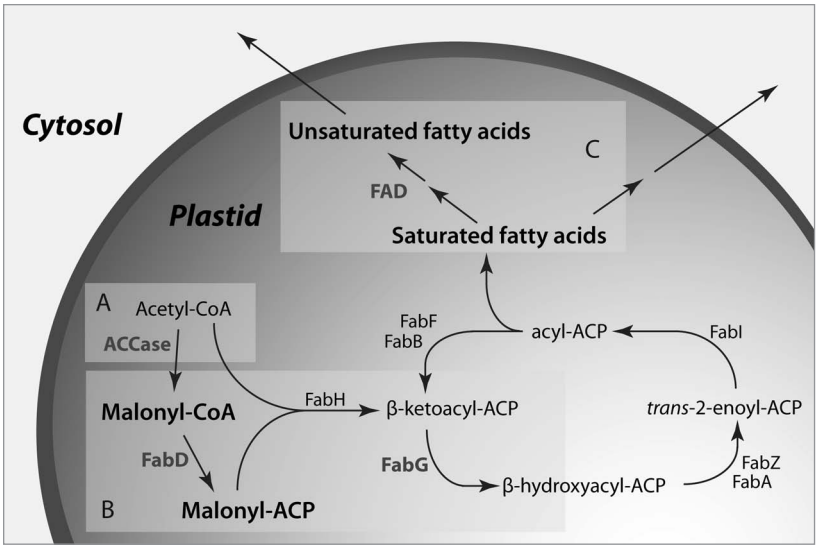

Figure 1. Simplified illustration of Type II fatty acid biosynthesis (FAB) in photosynthetic organisms. The synthesis of fatty acids takes place within the plastid. (**A**) The initial steps involve the synthesis of malonyl-CoA by the enzyme acetyl-CoA carboxylase (ACCase). (**B**) Malonyl-CoA is converted into malonyl-ACP by enzyme FabD, following which a series of Fab enzymes are engaged in the synthesis of saturated fatty acid chains. (**C**) The production of unsaturated fatty acids is catalyzed by a number of fatty acid desaturases (FADs).

phylogenomic study[15] demonstrated that a substantial number of genes in diatoms have arisen from green algal sources. This surprising result may be explained by an additional (cryptic) endosymbiosis in the ancestor of diatoms and other "chromalveolates." Under this (still controversial) scenario, the plastid of the captured green alga was lost but endosymbiont genes that were transferred to the "chromalveolate" nucleus remain as "footprints" of this association. Alternatively, all "green genes" that have been discovered in diatoms and in other taxa such as the stramenopile *Ectocarpus siliculosus*,[16] the dinoflagellate *Alexandrium tamarense*,[17] the cryptophyte *Bigelowiella natans* and the chlorarachniophyte *Guillardia theta*,[18] and the haptophyte *Emiliania huxleyi*,[15,19] may be the by-product of dozens of independent horizontal gene transfer (HGT) events.[20,21] Although proving unambiguously either of these hypotheses is challenging with current data, the cryptic endosymbiosis hypotheses provides a testable prediction: there should be a variety of plastid-targeted proteins encoded by green genes, suggesting these pre-existing functions associated with a green plastid have been co-opted to expand the metabolic capacity of the red algal-derived organelle. An example of this is provided by the existence of prasinophyte (green algal)-derived genes for the photo-protective xanthophyll cycle in "chromalveolates" that provides high photosynthetic efficiency under fluctuating light conditions.[22]

Beyond plastid functions, a recent analysis shows a mosaic (red and or green algal) origin of membrane transporters in diatoms with the majority of genes putatively derived from green sources.[23,24] These results demonstrate more broadly the role of genetic transfer as a driver of environmental adaption and cell evolution. In addition, the genomes of microbial eukaryotes, particularly "chromalveolates," appear to host a large number of horizontally transferred genes from various prokaryote and eukaryote lineages.[17,25-28] Therefore, the evolutionary history of FAB in diatoms is expected to be complex with potentially multiple different genetic inputs including red and/or green algal donors, as a result of endosymbiotic gene transfer (EGT) as well as HGT from these and other sources. Here we applied a phylogenetic approach to examine the evolutionary history of a representative set of genes in diatoms that are implicated in the plastid FAB system to assess the impact of E/HGT on FAB evolution in diatoms and microbial eukaryotes in general.

Results and Discussion

We identified 12 genes in diatoms that have putative functions implicated across all three major phases in FAB: (a) the initiation of FAB, (b) the synthesis of fatty acid chains, and (c) the desaturation of fatty acid chains, as well as fatty acid elongation. **Table 1** shows the list of these genes with their putative origins and presence/absence of protein-targeting signals to the plastid. The phylogeny of each of these gene families is shown in **Figure S1**, following the order in **Table 1**. Interestingly using our approach, seven of the 12 proteins have a putative green algal origin, and the remainder shows convoluted evolutionary histories that are too difficult to decipher with available data. We found evidence of a plastid-targeting signal in most of the enzymes encoded by these nuclear encoded FAB-related genes, consistent with the hypothesis that de novo FAB occurs in the plastid in photosynthetic eukaryotes.

Initiation of fatty acid biosynthesis

One of the most striking examples of green algal derived genes is the *Acc* gene that encodes acetyl-CoA carboxylase (ACCase, EC 6.4.1.2), an enzyme critical for the first dedicated step of Type II FAB. Two of the key intermediate molecules in FAB are acetyl-CoA and malonyl-CoA. Malonyl-CoA is required for the production of the backbone structure of fatty acid chains.[29] The enzyme ACCase is involved in the carboxylation of acetyl-CoA, yielding malonyl-CoA. The inhibition of ACCase production leads to cell death, making the enzyme a useful target for commercial herbicides in plants.[30] Expression of the *Acc* gene has been reported to regulate fatty acid composition in plant seeds.[31]

ACCase consists of four protein subunits: biotin carboxylase, biotin carboxyl carrier protein, and two subunits of carboxyltransferase.[32] In most plastid-bearing organisms, two types of ACCase exist in the cells: (a) a plastidic, heteromeric ACCase that contains two different subunits of carboxyltransferase (α and β), and (b) a cytosolic, homomeric ACCase that contains two identical carboxyltransferase subunits. Plastidic heteromeric ACCase is essential for the synthesis of fatty acids in the majority of photosynthetic organisms,[33] except in the grass family in which it is lacking.[34] In contrast, homomeric ACCase that is exclusively cytosolic in most photosynthetic organisms is

Table 1. List of FAB-related genes in diatoms that are used in this study

Gene	Encoded protein or putative function	Query (GI)	EC number	Plastid-targeting signal
*Acc**	Acetyl-CoA carboxylase	224004864	6.4.1.2	Yes
*ACS**	Acyl-CoA synthethase	224003657	2.3.1.86	No
fabD	Malonyl-CoA:ACP transacylase	224001858	2.3.1.39	Yes
*fabG**	3-ketoacyl-ACP reductase	224005350	1.1.1.100	No
fabH	β-ketoacyl-ACP synthase III	224013337	2.3.1.41	Yes
*fad3**	Omega-3 fatty acid desaturase	224002771	1.14.99.-	Yes
fad?	Fatty acid desaturase (predicted)	224014800	1.14.99.-	No
Delta-6 FAD-like*	Delta-6 FAD-like protein	223999591	1.14.99.-	Yes
*fad11**	Delta-11 palmitoyl CoA desaturase	224000772	1.14.19.5	No
ELO1	Polyunsaturated fatty acid elongase 1	75108642	1.14.99.-	No
*ELO2**	Polyunsaturated fatty acid elongase 2	224005955	1.14.99.-	No
ELO3	Polyunsaturated fatty acid elongase 3	220970795	1.14.99.-	No

For each gene, the corresponding encoded/predicted protein from *Thalassiosira pseudonana* is used as query (GenBank GI number shown). The corresponding Enzyme Commission number and prediction of plastid-targeting signal for each of the encoded proteins is shown. Genes marked with an asterisk (*) show evidence of putative green algal origin based on our phylogenetic analysis.

important for the synthesis of a number of other metabolites; e.g., flavonoids, anthocyanins, malonated amino acids, and ethylene precursors.[33]

Figure 2A shows the phylogeny of the diatom ACCase protein family (complete tree shown in **Fig. S1A**). For the FAB-related plastid-targeted ACCase, we observe monophyly (bootstrap support 86%) of diatoms (*Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, and *Fragilariopsis cylindrus*) and prasinophyte green algae (*Ostreococcus* and *Micromonas*), in the presence of other algal lineages such as the red alga *Porphyridium purpureum*. This phylogeny is consistent with (i.e., does not prove) an algal origin of plastid-targeted ACCase in diatoms with the prasinophytes being a putative source. The absence of bootstrap support for the order of divergence within this clade however makes it impossible to infer the direction of gene transfer between red/green algae and "chromalveolates." Interestingly, we find no evidence of HGT in the non-plastid-targeted ACCase. These proteins in diatoms form a separate group with the other stramenopiles (including *Phytophthora* and *Ectocarpus*) external to the clade containing the plastid-targeted isoform. The distinct evolutionary history of these two *Acc* isoforms suggests that the genes were independently acquired in diatoms, other "chromalveolates," and in the green-plastid containing rhizarians *Bigelowiella natans* and *Euglena gracilis*, and that this gene implicated in FAB has a putative green algal origin in all of these taxa.

The origin of the Type II FAB system in diatoms (and shared by other "chromalveolates") can be traced back to the origin of plastid itself. The plastid in "chromalveolates" is postulated to have arisen from an endosymbiotic relationship with a red alga,[13,14] and as proposed more recently, also potentially via an earlier cryptic endosymbiosis with a green alga, in particular, a prasinophyte.[15] Focusing on the prasinophyte lineages in a tree that excludes other members in the clade within which plastid-targeted diatom ACCases are found, and the highly diverged prokaryotic outgroups (**Fig. 2B**), we find all prasinophyte *Acc* genes to be divided into two subgroups with each of them robustly supported (bootstrap 100%). This suggests that the *Acc* gene underwent duplication in the common ancestor of *Micromonas* and *Ostreococcus* (and potentially other prasinophytes). The lack of support for monophyly of the two paralogs is likely explained by sequence divergence that occurred after the duplication event, because there is no reason to believe that the prasinophyte host lineage is polyphyletic among eukaryotes. Under this scenario, plastidic ACCase in "chromalveolates" arose via the transfer (e.g., EGT) of the *Acc* gene that had undergone subfunctionalization in prasinophytes and was tailored specifically for Type II FAB within the plastid. In contrast, the gene copy encoding cytosolic ACCase within plastid-bearing organisms appears to have been vertically inherited in "chromalveolates."

Synthesis of fatty acid chains

In the Type II system, the synthesis of fatty acid chains is a sequential repetitive process involving a series of enzymes, using acyl carrier protein (ACP) as a carrier molecule and the malonyl side chain as donor to sequentially add two-carbon units to the chain.[35,36] Multiple pathways are involved in the initiation of FAB[36] and the one relevant to this study is shown in **Figure 1**. **Figure S1D and S1E** show the phylogenies of two enzymes in diatoms that are involved in this process, respectively, for 3-ketoacyl-ACP reductase (FabG, EC 1.1.1.100) and β-ketoacyl-ACP synthase III (FabH, EC 2.3.1.41). FabH is the key enzyme in the synthesis of β-ketoacyl-ACP, which, in an intermediate step in FAB, is then converted into β-hydoxyacyl-ACP by FabG,[36] as shown in **Figure 1**. We found no clear evidence of HGT in the evolutionary history of FabH, indicating vertical inheritance (**Fig. S1E**), in which the diatom genes were within a strongly supported clade with the other stramenopiles, rhizarians, and the haptophyte *Emiliania huxleyi* (the "chromalveolates"; bootstrap

Figure 2. Phylogenies of *Acc* gene families in diatoms, showing (**A**) for all lineages and (**B**) with focus on the prasinophyte lineages. Only bootstrap support ≥50% is shown on the internal branches. The diatoms and prasinophytes are highlighted in bold and each plastid-targeted protein is marked with a dot. The full tree is shown in **Figure S1A**.

76%), with substantial support of a sister lineage of alveolates, and a strongly supported Viridiplantae clade (bootstrap 80%) elsewhere on the tree. Interestingly, the FabG proteins in diatoms (together with other stramenopiles and Rhizaria) are clustered within a strongly supported monophyletic group (bootstrap 86%) with prasinophytes (**Fig. S1D**). In the absence of red algal and other green algal homologs in this tree, our results suggest that FabG in diatoms shares an affiliation with prasinophytes, an observation that is plausibly explained by HGT or less likely, by gene loss events in all other lineages. The third important enzyme in this group is the malonyl-CoA:ACP transacylase (FabD, EC 2.3.1.39) that converts malonyl-CoA into malonyl-ACP.[36,37] This gene, while showing likely vertical inheritance in diatoms (**Fig. S1C**) occurs as non-plastid-targeted proteins in

prasinophytes that are clustered in a well-supported monophyletic relationship (85%) with the dinoflagellate *Alexandrium tamarense* (and the choanoflagellates *Monosiga*), suggesting a putative E/HGT association between the green algal and the alveolate lineages. The prasinophytes, which represent a basal lineage of green algae, may be sources of these genes, although this aspect remains to be validated.

Desaturation of fatty acid chains

Unsaturated fatty acids are pivotal components in cell membranes and therefore crucial for cell survival. In photosynthetic organisms, a variety of fatty acid desaturases (FADs) are present in their genomes. FADs are highly labile,[38] many of which have stringent specificity; i.e., location on the carbon chain (regiospecificity) and select substrates.[39] At the molecular level, an amino acid difference of as few as five residues can change the regiospecificity of the enzymatic reaction of the protein. Previous work has demonstrated that different sets of FADs can operate in different pathways and subcellular compartments,[39,40] therefore enzymes with the same or very similar functions within the diatoms (or any organism) can have different evolutionary histories.[41] Therefore, even using our stringent phylogenomic approach, some of the highly similar proteins (e.g., of different groups of FADs) could be included in a phylogenetic tree.

As shown in **Figure S1F**, we found evidence of red and/or green algal origin of the diatom *fad3* gene, which encodes omega-3 FAD, an important enzyme that converts linoleic (18:2) into linolenic (18:3) acids, as shown in two separate monophyletic clades at bootstrap support respectively, of 99% and 93%. Upon closer inspection, the proteins encoded in *T. pseudonana* that show a putative green algal origin are also annotated as hypothetical proteins with putative omega-6 FAD function, as encoded by the *fad6* gene. The phylogenetic tree of another predicted fatty acid desaturase in diatoms (**Fig. S1G**) shows no clear evidence of HGT. This finding suggests divergence of protein functions based on acquired genetic material from different sources, but this aspect remains to be validated as more genome data and better annotations become available. **Figure S1H** shows the phylogeny of the delta-6 FAD-like gene family (also known as the FADS2). The delta-6 FAD-like domain (GenBank accession CD03506) includes integral membrane enzymes of both delta-6 and delta-8. We find within this gene family an association (bootstrap 69%), although not as strong as the commonly accepted threshold of ≥70%, between prasinophytes and "chromalveolate" (diatoms and haptophytes) lineages, with other diatom and green algal copies elsewhere in the tree. Our findings highlight the complicated issues associated with inferring phylogenetic trees from the divergent protein families of FADs.

There are two distinct types of FADs: (a) ACP (soluble) desaturases found only in plants and certain bacteria, and (b) membrane-bound (insoluble) desaturases found in most aerobic organisms including bacteria, fungi, plants, and animals.[42] In plants, ACP desaturases are associated with the plastid, whereas membrane-bound desaturases are found in both the plastid and the endoplasmic reticulum.[43] It is intriguing that ACP desaturases are not found in extant cyanobacteria (**Fig. S1D**). Because cyanobacteria gave rise to plastids, ACP desaturases in plastid-bearing organisms can be explained by HGT from non-cyanobacterial lineages into ancestral algal lineages or gene loss within cyanobacteria after establishment of the photosynthetic eukaryotes.[42] Here we report a number of other green algal derived genes in diatoms that are involved in fatty acid biosynthesis, desaturation, and elongation, including *fad11*, encoding FAD11 that desaturates palmitic acids (**Fig. S1I**), *ACS* encoding acyl-CoA synthethase (**Fig. S1B**), and *ELO2* (**Fig. S1K**) encoding membrane-bound polyunsaturated fatty acid elongase (the ELO superfamily). The other two diatom genes within the ELO superfamily, *ELO1* (**Fig. S1J**) and *ELO3* (**Fig. S1L**), show possible vertical inheritance or an evolutionary history that currently precludes an easy explanation.

The "fat revolution" in diatoms

Our findings demonstrate that HGT, duplication, and subfunctionalization of genes are key evolutionary processes in the evolution of the Type II FAB system within photosynthetic eukaryotes. In diatoms, and in potentially most, if not all stramenopiles, some of the key genes involved in Type II FAB system and desaturation of fatty acids trace their origin to green algal (likely prasinophyte) sources. The finding that components of key plastid functions such as FAB and photo-protection[22] in diatoms and other "chromalveolates" are prasinophyte-derived is consistent with (but does not prove) the cryptic endosymbiosis hypothesis.[15] It is also conceivable that prasinophytes provided a rich and easily accessible source of genes for ancestral "chromalveolates" lineage(s) and underwent massive levels of HGT into these chlorophyll *c*-containing taxa. This resulted in a highly chimeric red plastid proteome and "chromalveolates" nuclear genome. Additional sources of evidence are needed to test these ideas.

It should be noted that by relying on phylogenomics we make the implicit assumption that genes are transferred as a whole during an E/HGT event. The modularity of HGT,[44] degree of conservation within the gene family, and genome rearrangement following HGT could affect the delineation of HGT history using phylogenetic comparisons; in the latter cases, alternative phylogenomic approach might be useful.[45] Furthermore, evolutionary analysis of FAD is complicated by the duplicated nature of these genes and the retention of high sequence similarity among homologs with different regiospecificities, which is the basis for gene annotation. Comparative biochemical validation of FAB pathways among photosynthetic "chromalveolate" and algal species will be a useful test of our results. Assessing functional biases among the horizontally transferred genes into the diatom genomes; e.g., whether genes implicated in the FAB system are more likely to have been transferred than those genes involved in other metabolic or cellular processes, can provide invaluable insights into how FAB evolved in diatoms. Nevertheless, given that FAB is a process crucial to the survival of organisms (in addition to photosynthesis in algae and plants), our results clearly demonstrate that endosymbiosis plays a more significant role in genome evolution and innovation of "chromalveolates" than previously thought. This includes the pathways that hold promise for providing biofuels in the near future.

Materials and Methods

Data

Twelve protein sequences with functions implicated in FAB, as predicted from the genome of *Thalassiosira pseudonana* CMP1335[46] (**Table 1**), were used to query genome data. The resulting alignments were used for phylogenetic analysis.

Phylogenetic analysis

We applied a phylogenetic approach adopted from Chan et al.[47] for inferring E/HGT events. For each of the protein sequences, we searched for putative homologs within a local database similar to the one used a earlier study[23] but with the updated NCBI RefSeq release 51 (http://www.ncbi.nlm.nih.gov/RefSeq). This database, comprising ca. 17 million protein sequences, include other genomic sources of predicted proteins (http://www.jgi.doe.gov/) and EST data (http://www.ncbi.nlm.nih.gov/dbEST/; http://tbestdb.bcm.umontreal.ca/) from other algae and protists. Other published transcriptome (or genome, where available) data from red algae *Cyanidioschyzon merolae*,[48] *Porphyridium purpureum*,[49] *Calliarthron tuberculosum*,[47] *Chondrus crispus*,[50] and *Galdieria sulphuraria*,[51] the stramenopile *Ectocarpus siliculosus*,[16] and the dinoflagellate *Alexandrium tamarense*[17] were also included in the database. Homologous protein families were aligned using MUSCLE[52] at default settings, and phylogenies were reconstructed using RAxML[53] using the WAG[54] model of amino acid substitution. E/HGT in diatoms from other algae were inferred when strongly supported monophyly (bootstrap value ≥75%) was observed for a sister group relationship between diatoms (with or without other "chromalveolates") and lineages of green and/or red algae.

Prediction of protein subcellular targets

Subcellular targets of all stramenopile proteins (including those of the diatoms) were determined using HECTAR (http://www.sb-roscoff.fr/hectar/).[55]

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

CXC is supported by a ResTeach Fellowship from The University of Queensland. This research was partially supported by a grant from the National Science Foundation awarded to DB (0936884).

Supplemental Materials

Supplemental materials may be found here:
www.landesbioscience.com/journals/mge/article/27313

References

1. Montsant A, Jabbari K, Maheswari U, Bowler C. Comparative genomics of the pennate diatom *Phaeodactylum tricornutum*. Plant Physiol 2005; 137:500-13; PMID:15665249; http://dx.doi.org/10.1104/pp.104.052829
2. Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. Primary production of the biosphere: integrating terrestrial and oceanic components. Science 1998; 281:237-40; PMID:9657713; http://dx.doi.org/10.1126/science.281.5374.237
3. Kröger N, Poulsen N. Diatoms-from cell wall biogenesis to nanotechnology. Annu Rev Genet 2008; 42:83-107; PMID:18983255; http://dx.doi.org/10.1146/annurev.genet.41.110306.130109
4. Townley HE, Parker AR, White-Cooper H. Exploitation of diatom frustules for nanotechnology: tethering active biomolecules. Adv Funct Mater 2008; 18:369-74; http://dx.doi.org/10.1002/adfm.200700609
5. Dismukes GC, Carrieri D, Bennette N, Ananyev GM, Posewitz MC. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. Curr Opin Biotechnol 2008; 19:235-40; PMID:18539450; http://dx.doi.org/10.1016/j.copbio.2008.05.007
6. Young RA. Fat, energy and mammalian survival. Am Zool 1976; 16:699-710
7. Kent C. Eukaryotic phospholipid biosynthesis. Annu Rev Biochem 1995; 64:315-43; PMID:7574485; http://dx.doi.org/10.1146/annurev.bi.64.070195.001531
8. Miller WL. Molecular biology of steroid hormone synthesis. Endocr Rev 1988; 9:295-318; PMID:3061784; http://dx.doi.org/10.1210/edrv-9-3-295
9. Razin S, Yogev D, Naot Y. Molecular biology and pathogenicity of mycoplasmas. Microbiol Mol Biol Rev 1998; 62:1094-156; PMID:9841667
10. Schweizer E, Hofmann J. Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. Microbiol Mol Biol Rev 2004; 68:501-17; PMID:15353567; http://dx.doi.org/10.1128/MMBR.68.3.501-517.2004
11. Zhang L, Joshi AK, Hofmann J, Schweizer E, Smith S. Cloning, expression, and characterization of the human mitochondrial beta-ketoacyl synthase. Complementation of the yeast CEM1 knock-out strain. J Biol Chem 2005; 280:12422-9; PMID:15668256; http://dx.doi.org/10.1074/jbc.M413686200
12. White SW, Zheng J, Zhang YM, Rock CO. The structural biology of type II fatty acid biosynthesis. Annu Rev Biochem 2005; 74:791-831; PMID:15952903; http://dx.doi.org/10.1146/annurev.biochem.74.082803.133524
13. Gould SB, Waller RF, McFadden GI. Plastid evolution. Annu Rev Plant Biol 2008; 59:491-517; PMID:18315522; http://dx.doi.org/10.1146/annurev.arplant.59.032607.092915
14. Reyes-Prieto A, Weber AP, Bhattacharya D. The origin and establishment of the plastid in algae and plants. Annu Rev Genet 2007; 41:147-68; PMID:17600460; http://dx.doi.org/10.1146/annurev.genet.41.110306.130134
15. Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D. Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science 2009; 324:1724-6; PMID:19556510; http://dx.doi.org/10.1126/science.1172983
16. Cock JM, Sterck L, Rouzé P, Scornet D, Allen AE, Amoutzias G, Anthouard V, Artiguenave F, Aury JM, Badger JH, et al. The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. Nature 2010; 465:617-21; PMID:20520714; http://dx.doi.org/10.1038/nature09016
17. Chan CX, Soares MB, Bonaldo MF, Wisecaver JH, Hackett JD, Anderson DM, Erdner DL, Bhattacharya D. Analysis of *Alexandrium tamarense* (Dinophyceae) genes reveals the complex evolutionary history of a microbial eukaryote. J Phycol 2012; 48:1130-42; PMID:23066170; http://dx.doi.org/10.1111/j.1529-8817.2012.01194.x
18. Curtis BA, Tanifuji G, Burki F, Gruber A, Irimia M, Maruyama S, Arias MC, Ball SG, Gile GH, Hirakawa Y, et al. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature 2012; 492:59-65; PMID:23201678; http://dx.doi.org/10.1038/nature11681
19. Read BA, Kegel J, Klute MJ, Kuo A, Lefebvre SC, Maumus F, Mayer C, Miller J, Monier A, Salamov A, et al. Emiliania huxleyi Annotation Consortium. Pan genome of the phytoplankton *Emiliania* underpins its global distribution. Nature 2013; 499:209-13; PMID:23760476; http://dx.doi.org/10.1038/nature12221
20. Burki F, Flegontov P, Oborník M, Cihlár J, Pain A, Lukeš J, Keeling PJ. Re-evaluating the green versus red signal in eukaryotes with secondary plastid of red algal origin. Genome Biol Evol 2012; 4:738-47; PMID:22593553; http://dx.doi.org/10.1093/gbe/evs049
21. Deschamps P, Moreira D. Reevaluating the green contribution to diatom genomes. Genome Biol Evol 2012; 4:795-800; PMID:22684208; http://dx.doi.org/10.1093/gbe/evs053
22. Frommolt R, Werner S, Paulsen H, Goss R, Wilhelm C, Zauner S, Maier UG, Grossman AR, Bhattacharya D, Lohr M. Ancient recruitment by chromists of green algal genes encoding enzymes for carotenoid biosynthesis. Mol Biol Evol 2008; 25:2653-67; PMID:18799712; http://dx.doi.org/10.1093/molbev/msn206
23. Chan CX, Reyes-Prieto A, Bhattacharya D. Red and green algal origin of diatom membrane transporters: insights into environmental adaptation and cell evolution. PLoS One 2011; 6:e29138; PMID:22195008; http://dx.doi.org/10.1371/journal.pone.0029138
24. Chan CX, Bhattacharya D, Reyes-Prieto A. Endosymbiotic and horizontal gene transfer in microbial eukaryotes: Impacts on cell evolution and the tree of life. Mob Genet Elements 2012; 2:101-5; PMID:22934244; http://dx.doi.org/10.4161/mge.20110

25. Nosenko T, Bhattacharya D. Horizontal gene transfer in chromalveolates. BMC Evol Biol 2007; 7:173; PMID:17894863; http://dx.doi.org/10.1186/1471-2148-7-173
26. Baurain D, Brinkmann H, Petersen J, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V, Roger AJ, Burger G, Lang BF, Philippe H. Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. Mol Biol Evol 2010; 27:1698-709; PMID:20194427; http://dx.doi.org/10.1093/molbev/msq059
27. Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber AP, Schwacke R, Gross J, Blouin NA, Lane C, et al. *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science 2012; 335:843-7; PMID:22344442; http://dx.doi.org/10.1126/science.1213561
28. Mackiewicz P, Bodył A, Moszczyński K. The case of horizontal gene transfer from bacteria to the peculiar dinoflagellate plastid genome. Mob Genet Elements 2013; 3:e25845; PMID:24195014; http://dx.doi.org/10.4161/mge.25845
29. Harwood JL. Fatty acid metabolism. Annu Rev Plant Physiol Plant Mol Biol 1988; 39:101-38; http://dx.doi.org/10.1146/annurev.pp.39.060188.000533
30. Zhang XQ, Powles SB. The molecular bases for resistance to acetyl co-enzyme A carboxylase (ACCase) inhibiting herbicides in two target-based resistant biotypes of annual ryegrass (*Lolium rigidum*). Planta 2006; 223:550-7; PMID:16133206; http://dx.doi.org/10.1007/s00425-005-0095-x
31. Roesler K, Shintani D, Savage L, Boddupalli S, Ohlrogge J. Targeting of the *Arabidopsis* homomeric acetyl-coenzyme A carboxylase to plastids of rapeseeds. Plant Physiol 1997; 113:75-81; PMID:9008389; http://dx.doi.org/10.1104/pp.113.1.75
32. Cronan JE Jr., Waldrop GL. Multi-subunit acetyl-CoA carboxylases. Prog Lipid Res 2002; 41:407-35; PMID:12121720; http://dx.doi.org/10.1016/S0163-7827(02)00007-3
33. Sasaki Y, Nagano Y. Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. Biosci Biotechnol Biochem 2004; 68:1175-84; PMID:15215578; http://dx.doi.org/10.1271/bbb.68.1175
34. Konishi T, Shinohara K, Yamada K, Sasaki Y. Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. Plant Cell Physiol 1996; 37:117-22; PMID:8665091; http://dx.doi.org/10.1093/oxfordjournals.pcp.a028920
35. Heath RJ, Rock CO. Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in *Escherichia coli*. J Biol Chem 1996; 271:1833-6; PMID:8567624; http://dx.doi.org/10.1074/jbc.271.4.1833
36. Magnuson K, Jackowski S, Rock CO, Cronan JE Jr. Regulation of fatty acid biosynthesis in *Escherichia coli*. Microbiol Rev 1993; 57:522-42; PMID:8246839
37. Verwoert II, Verbree EC, van der Linden KH, Nijkamp HJ, Stuitje AR. Cloning, nucleotide sequence, and expression of the *Escherichia coli fabD* gene, encoding malonyl coenzyme A-acyl carrier protein transacylase. J Bacteriol 1992; 174:2851-7; PMID:1314802
38. Cahoon EB, Lindqvist Y, Schneider G, Shanklin J. Redesign of soluble fatty acid desaturases from plants for altered substrate specificity and double bond position. Proc Natl Acad Sci U S A 1997; 94:4872-7; PMID:9144157; http://dx.doi.org/10.1073/pnas.94.10.4872
39. Hashimoto K, Yoshizawa AC, Okuda S, Kuma K, Goto S, Kanehisa M. The repertoire of desaturases and elongases reveals fatty acid variations in 56 eukaryotic genomes. J Lipid Res 2008; 49:183-91; PMID:17921532; http://dx.doi.org/10.1194/jlr.M700377-JLR200
40. Hashimoto K, Yoshizawa AC, Saito K, Yamada T, Kanehisa M. The repertoire of desaturases for unsaturated fatty acid synthesis in 397 genomes. Genome Inform 2006; 17:173-83; PMID:17503367
41. Domergue F, Spiekermann P, Lerchl J, Beckmann C, Kilian O, Kroth PG, Boland W, Zähringer U, Heinz E. New insight into *Phaeodactylum tricornutum* fatty acid metabolism. Cloning and functional characterization of plastidial and microsomal delta12-fatty acid desaturases. Plant Physiol 2003; 131:1648-60; PMID:12692324; http://dx.doi.org/10.1104/pp.102.018317
42. Sperling P, Ternes P, Zank TK, Heinz E. The evolution of desaturases. Prostaglandins Leukot Essent Fatty Acids 2003; 68:73-95; PMID:12538072; http://dx.doi.org/10.1016/S0952-3278(02)00258-2
43. Okuley J, Lightner J, Feldmann K, Yadav N, Lark E, Browse J. *Arabidopsis FAD2* gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. Plant Cell 1994; 6:147-58; PMID:7907506; http://dx.doi.org/10.1105/tpc.6.1.147
44. Chan CX, Darling AE, Beiko RG, Ragan MA. Are protein domains modules of lateral genetic transfer? PLoS One 2009; 4:e4524; PMID:19229333; http://dx.doi.org/10.1371/journal.pone.0004524
45. Chan CX, Ragan MA. Next-generation phylogenomics. Biol Direct 2013; 8:3; PMID:23339707; http://dx.doi.org/10.1186/1745-6150-8-3
46. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, et al. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 2004; 306:79-86; PMID:15459382; http://dx.doi.org/10.1126/science.1101156
47. Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, Bhattacharya D. Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. Curr Biol 2011; 21:328-33; PMID:21315598; http://dx.doi.org/10.1016/j.cub.2011.01.037
48. Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Nishida K, et al. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 2004; 428:653-7; PMID:15071595; http://dx.doi.org/10.1038/nature02398
49. Bhattacharya D, Price DC, Chan CX, Qiu H, Rose N, Ball S, Weber AP, Arias MC, Henrissat B, Coutinho PM, et al. Genome of the red alga *Porphyridium purpureum*. Nat Commun 2013; 4:1941; PMID:23770768; http://dx.doi.org/10.1038/ncomms2931
50. Collén J, Porcel B, Carré W, Ball SG, Chaparro C, Tonon T, Barbeyron T, Michel G, Noel B, Valentin K, et al. Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. Proc Natl Acad Sci U S A 2013; 110:5247-52; PMID:23503846; http://dx.doi.org/10.1073/pnas.1221259110
51. Schönknecht G, Chen WH, Ternes CM, Barbier GG, Shrestha RP, Stanke M, Bräutigam A, Baker BJ, Banfield JF, Garavito RM, et al. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 2013; 339:1207-10; PMID:23471408; http://dx.doi.org/10.1126/science.1231707
52. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004; 32:1792-7; PMID:15034147; http://dx.doi.org/10.1093/nar/gkh340
53. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 2006; 22:2688-90; PMID:16928733; http://dx.doi.org/10.1093/bioinformatics/btl446
54. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 2001; 18:691-9; PMID:11319253; http://dx.doi.org/10.1093/oxfordjournals.molbev.a003851
55. Gschloessl B, Guermeur Y, Cock JM. HECTAR: a method to predict subcellular targeting in heterokonts. BMC Bioinformatics 2008; 9:393; PMID:18811941; http://dx.doi.org/10.1186/1471-2105-9-393