http://www.elsevier.de/protis
Published online date 5 August 2011

Protist

ORIGINAL PAPER

Analysis of Expressed Sequence Tags from the Marine Microalga *Pseudochattonella farcimen* (Dictyochophyceae)

Simon M. Dittami[a,1], Ingvild Riisberg[a], Uwe John[b], Russell J.S. Orr[c], Kjetill S. Jakobsen[d], and Bente Edvardsen[a,1]

[a]Program for Marine Biology, Department of Biology, University of Oslo, P.O. Box 1066 Blindern, 0316 Oslo, Norway
[b]Alfred Wegner-Institute for Marine Research, Am Handelshafen 12, D-27570 Bremerhaven, Germany
[c]Microbial Evolution Research Group, Department of Biology, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway
[d]Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway

Submitted January 21, 2011; Accepted June 28, 2011
Monitoring Editor: Michael Melkonian

Pseudochattonella farcimen **(Eikrem, Edvardsen, et Throndsen) is a unicellular alga belonging to the Dictyochophyceae (Heterokonta). It forms recurring blooms in Scandinavian coastal waters, and has been associated to fish mortality. Here we report the sequencing and analysis of 10,368 expressed sequence tags (ESTs) corresponding to 8,149 unique gene models from this species. Compared to EST libraries from other heterokonts, *P. farcimen* contains a high number of genes with functions related to cell communication and signaling. We found several genes encoding proteins related to fatty acid metabolism, including eight fatty acid desaturases and two phospholipase A2 genes. Three desaturases are highly similar to △4-desaturases from haptophytes. *P. farcimen* also possesses three putative polyketide synthases (PKSs), belonging to two different families. Some of these genes may have been acquired via horizontal gene transfer by a common ancestor of brown algae and dictyochophytes, together with genes involved in mannitol metabolism, which are also present in *P. farcimen*. Our findings may explain the unusual fatty acid profile previously observed in *P. farcimen,* and are discussed from an evolutionary perspective and in relation to the ichthyotoxicity of this alga.**

Key words: Harmful algae; heterokonts; fatty acid desaturases; polyketide synthases (PKS); mannitol metabolism; horizontal gene transfer (HGT).

Introduction

Heterokonts (stramenopiles) constitute a major eukaryotic lineage, which has evolved independently from the well-studied plant (including red- and green algae) and opisthokont (including animals and fungi) lineages (Fig. 1). Recently published genomes of heterokonts, such as diatoms (Armbrust et al. 2004; Bowler et al. 2008), brown algae (Cock et al. 2010), and oomycetes (Tyler et al. 2006) have generated valuable insights into the evolution and unique metabolic pathways of

[1]Corresponding authors; fax +47 22 85 44 38
e-mail simon.dittami@bio.uio.no (S.M. Dittami), bente.edvardsen@bio.uio.no (B. Edvardsen).

doi:10.1016/j.protis.2011.07.004

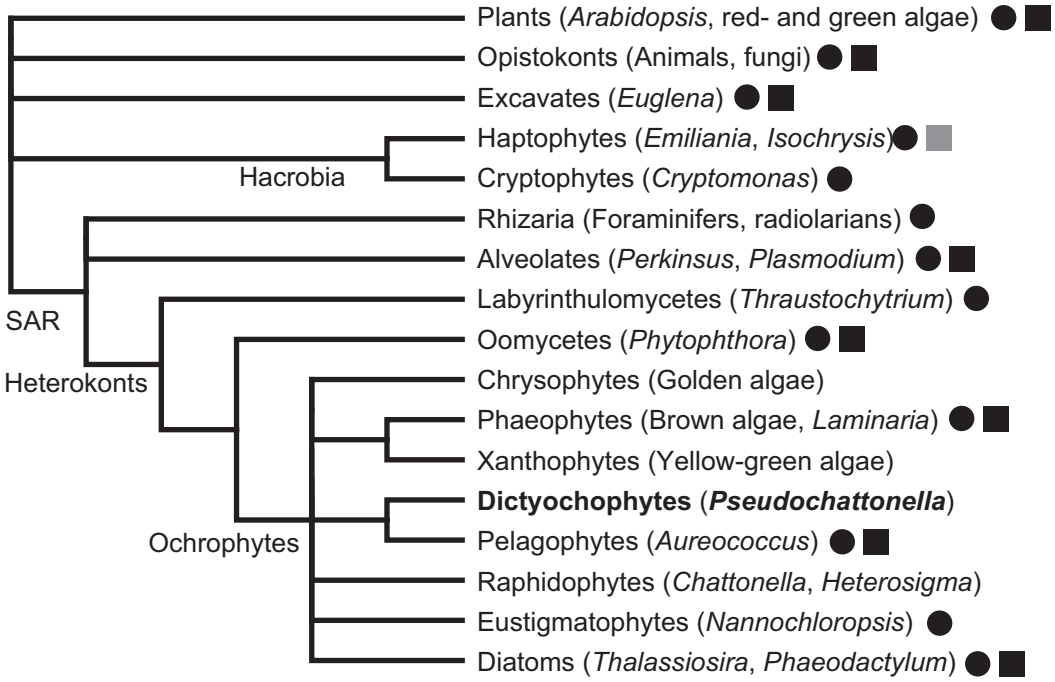

Figure 1. Reduced phylogenetic tree of eukaryotes according to Burki et al. (2007) and Riisberg et al. (2009), with a focus on heterokonts. Common names or examples of genera are given in brackets after the clade name. Circles indicate available ESTs (>1,000 sequences), squares available nuclear genome sequences according to the NCBI and JGI databases. The grey square for *Emiliania* indicates that this genome has been released but not officially published.

several of these organisms, yet several other classes within this sub-kingdom remain poorly covered by sequencing projects. Such studies could greatly further our understanding of the evolution of heterokonts.

Here, we focus on one representative of the heterokonts: the dictyochophyte *Pseudochattonella farcimen* Eikrem, Edvardsen et Throndsen (Edvardsen et al. 2007; Eikrem et al. 2009). Within the *Dictyochophyceae*, available molecular data prior to the submission of our dataset consisted of only 157 nucleotide and 43 protein sequences, which were almost exclusively taxonomic markers. The most closely related species with available EST or genome data was the pelagophyte *Aureococcus anophagefferens* Hargraves & Sieburth (Ben Ali et al. 2002; Riisberg et al. 2009). *Pseudochattonella farcimen* is a unicellular, ichthyotoxic alga, forming recurrent blooms in Scandinavian marine waters, at times causing severe fish mortality (Aure et al. 2001; Backe-Hansen et al. 2001; Edler 2006). Skjelbred et al. (2011) demonstrated *Pseudochattonella* spp. cultures to adversely affect metabolism of fish cells and to damage gills of cod fry and salmon smolts. Despite its impact on the aquaculture industry, little is known about the biology and toxicity of algae in the genus *Pseudochattonella*, encompassing the two species *P. farcimen* and *P. verruculosa*.

Other aspects of *Pseudchattonella* species that have been examined are chemical markers and potentially toxic or toxin-related substances such as sterols and fatty acids (Giner et al. 2008). Two strains of *Pseudochattonella* sp. were shown to produce a rare 27-nor sterol (occelasterol), for which the biosynthetic pathway is yet unknown. In addition, both strains contained a high proportion of polyunsaturated fatty acids (PUFAs). Two unusual features were particularly high ratios of docosahexaenoic acid (DHA) to eicosapentaenoic acid (EPA), as well as the presence of the rare PUFA C18:5n-3. The latter fatty acid was previously also found in a toxic species of the dinoflagellate genus *Prorocentrum* (Mansour 1999), the toxic raphidophyte *Heterosigma akashiwo* (Mostaert et al. 1998), and in several haptophytes including *Emiliania huxleyi* (Viso and Marty 1993), *Isochrysis galbana* (Renaud 1999), and *Chrysochromulina polylepis* (John et al.

Table 1. Overview of species and EST libraries analyzed in this study. Shown are the total number of ESTs analyzed, the number of non-redundant sequences (NRSs) obtained and the mean G/C content of the coding sequences.

Species	Class	ESTs	NRSs	G/C
Pseudochattonella farcimen	Dictyochophyceae	10,367	8,149	53.1%
Aureococcus anophagefferens	Pelagophyceae	51,271	18,668	68.1%
Nannochloropsis oculata	Eustigmatophyceae	1,961	1,858	52.1%
Phaeodactylum tricornutum	Bacillariophyceae	121,750	73,696	50.7%
Phytophthora capsici	Oomycota	56,457	11,448	54.0%
Ectocarpus siliculosus	Phaeophyceae	90,637	17,039	55.7%
Isochrysis galbana	Prymnesiophyceae	12,274	6,088	63.6%

2002). Just as for occelasterol, the molecular mechanisms underlying the synthesis of C18:5n-3 are still unknown.

In this study, the rationale was to improve our understanding of the biology, in particular the biochemical capacity of *P. farcimen* by generating and analyzing an EST library for this dictyochophyte. The results highlight several interesting features about the physiology of this alga, including the presence of numerous genes involved in signaling, fatty acid metabolism, polyketide synthesis, and mannitol metabolism. Finally, our data are also of interest regarding recent theories on the evolution of heterokonts, as they reveal diverse origins of fatty acid metabolism-related genes as well as polyketide synthases.

Results and Discussion

Characterization of the EST Library

After ligation and transformation of cDNA, 10,368 clones were sequenced, and a total of 10,042 sequences remained after cleaning. A total of 1,240 tentative contigs and 6,909 singletons were obtained during the assembly of these sequences. Overall, the low ratio of contigs to singletons indicates that higher sequencing depth could have led to a significant increase in the number of non-redundant sequences (NRSs, i.e. contigs + singletons). The distribution of the number of expressed sequence tags (ESTs) per NRS is available in Supplementary File S1. For each of these, a putative open reading frame was found: the overall G/C content in these open reading frames was 53.1%, and thus similar to that found in the EST libraries of most heterokonts excluding *Aureococcus* and the haptophyte *I. galbana* (Table 1).

Automatic GO annotations could be obtained for 1,394 (17%) of the 8,149 NRSs (e-value cutoff 1e-10), and a total of 3,327 sequences (41%) had homologs (e-value cutoff 1e-6) in the NCBI nr database. An overview of GOSlim annotations obtained for the *P. farcimen* library is shown in Figure 2. In most cases, best BLAST hits were found in *Ectocarpus siliculosus* (992 top hits i.e. 30%), *Thalassiosira pseudonana* (351 i.e. 11%), *Phaeodactylum tricornutum* (313 i.e. 9%), and *Phytophthora infestans* (280 i.e. 8%), followed by *Micromonas* (72 i.e. 2%) (Supplementary File S2).

For 4,822 (70%) of the sequences no homolog (e-value cutoff 1e-6) was found in the NCBI nr database. This number is unexpectedly high, considering that, at the time of the analysis, three genomes of heterokont algae (*T. pseudonana*, *P. tricornutum*, and *E. siliculosus*) had already been completed and in the database. The lack of homologous sequences may partially be due to the fact that some of the predicted coding sequences (CDSs) were incomplete (the mean NRS length was 1070 nucleotides, the mean predicted CDS length was 171 amino acids), but certainly also indicates the large discovery potential for genes with new functions in this species.

Contigs with High EST Support are Largely Unknown

One of the sequences of particular interest is, for example, the predicted protein with the highest EST support (15 reads): This CDS has no homologs in the NCBI nr protein database (e-value > 1), but is likely to contain a type 2 signal anchor as predicted by HECTAR (Gschlössl et al. 2008) and one transmembrane domain, the N-terminal end of the hypothetical protein being predicted to face towards the inside of the cell. Manual annotations for 16 predicted proteins with the next highest number of reads ($\geq$6) are detailed in Table 2. In total, a putative function could only be assigned to 5 of the 17 best supported sequences (i.e. 29%).

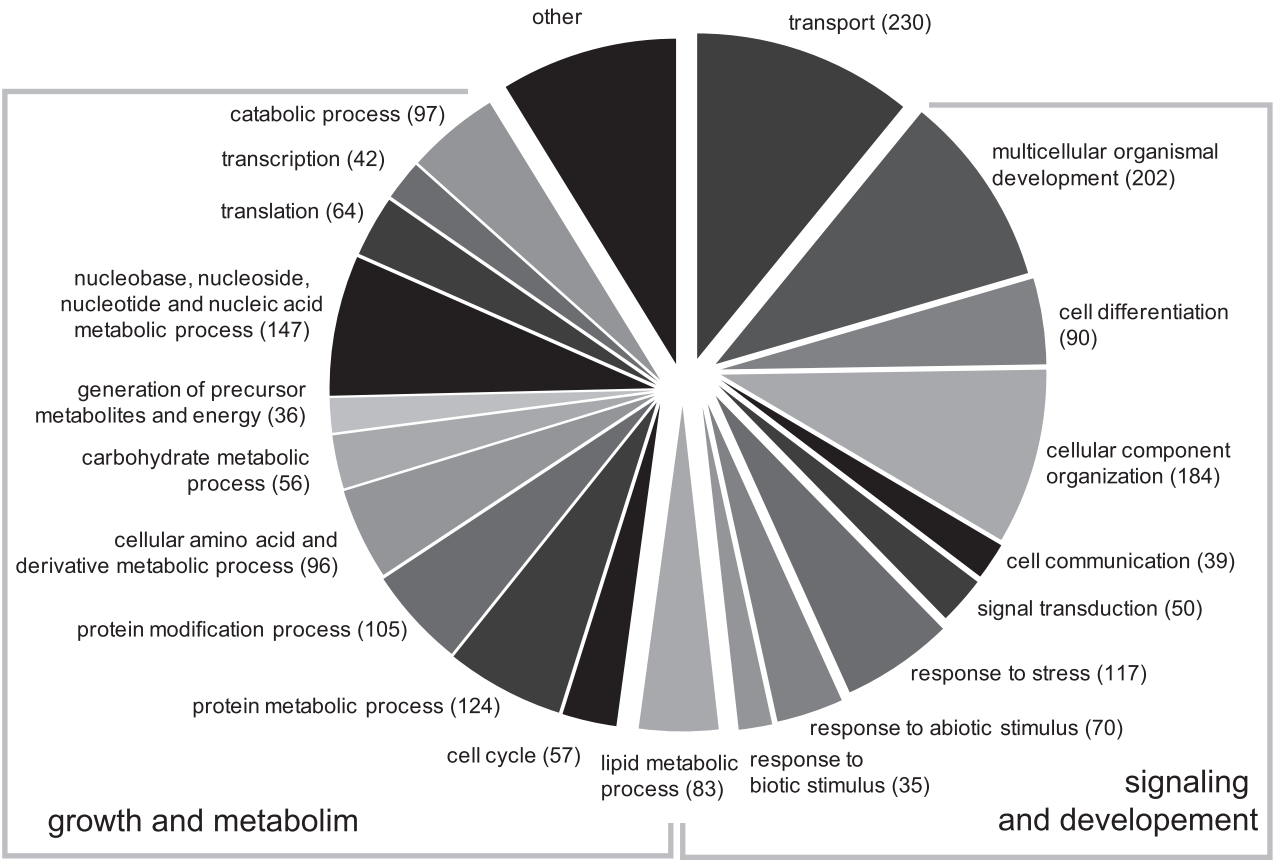

Figure 2. Overview of GOSlim annotations obtained for the *P. farcimen* EST library. The absolute number of sequences with annotations falling into each category is given in parentheses after the name of the category.

ESTs Related to Signaling and Multicellularity

In order to obtain a rough overview of which functional groups of genes were particularly abundant in the *P. farcimen* EST collection, automatic GO annotations obtained for this species were compared to those obtained for other public EST libraries of related organisms (listed in Table 1), limiting the false discovery rate to 5% using the Benjamini and Hochberg correction (Benjamini and Hochberg 1995). An overrepresentation of certain groups of ESTs in one library may be caused by a number of factors including the growth condition of the algal culture prior to RNA extraction, the method used for normalization, the quality of the sequences, the quality of the sequence assembly, and finally the quality of the annotations, the latter depending both on the length of the ESTs and the presence of well-annotated homologues, but may also reflect genomic and transcriptional differences between different species.

We did not detect any significant differences between the *P. farcimen* and the *I. galbana* EST libraries. However, compared to the other examined species, several functional categories were enriched in *P. farcimen*. Here we only considered GO terms significantly overrepresented compared to all five other examined species, except *I. galbana*. These GO-terms fell into three functional categories: protein binding, signaling, and multicellular organismal process (Table 3).

The finding of genes related to the development of multicellular organisms among the highly represented sequences may at first seem surprising, as both *P. farcimen* and *I. galbana* are unicellular. However, similarly high proportions of genes falling into this category (7.3%) have recently also been detected in the unicellular, protozoan parasite *Perkinsus marinus* (Joseph et al. 2010). Indeed, most of the genes annotated with the GO-term “multicellular organismal process” were related to signaling, the most abundant sequences

Table 2. Predominant transcripts in the *P. farcimen* EST library (≥6 reads support). Proteins were considered "hypothetical", if no blast hit with an e-value < 1e-10 was found in the NCBI nr database, and "conserved unknown", if blast hits (e-value < 1e-10) were available only with uncharacterized proteins. The features column contains information on predicted conserved domains or signal peptides obtained by InterProScan and HECTAR.

Sequence	Reads	Annotation	Best hit	e-value	Features
FR751558	15	hypothetical protein	*Fagus sylvatica*	2.8	Type II signal anchor
FR751561	7	conserved unknown protein	*Thalassiosira pseudonana*	1e-17	ZINC FINGER DHHC DOMAIN, 3 transmembrane domains
FR751562	7	hypothetical protein	–	>10	–
FR751563	7	hypothetical protein	–	>10	Type II signal anchor
FR751564	7	endoplasmic reticulum oxidoreduction	*Phytophthora infestans*	1e-43	ERO1 domain
FR751566	7	conserved unknown protein	*Ectocarpus siliculosus*	6e-23	–
FR751568	6	20S proteasome subunit alpha type 1	*E. siliculosus*	5e-56	PROTEASOME SUBUNIT, nucleophile aminohydrolases
FR751578	6	hypothetical protein	*Ixodes pacificus*	0.062	–
FR751584	6	hypothetical protein	*Marivirga tractuosa*	0.28	–
FR751572	6	Clp protease	*Phaeodactylum tricornutum*	1e-84	CLP_PROTEASE
FR751577	6	hypothetical protein	*E. siliculosus*	7e-07	ATG11 (Autophagy-related)
FR751571	6	hypothetical protein	Metagenome: Mediterranean deep chlorophyll max.	9e-45	–
FR751564	6	endoplasmic oxidoreduction	*P. infestans*	8e-43	ERO1 domain
FR751570	6	conserved unknown protein	*E. siliculosus*	4e-29	Mitochondrial transit peptide; CBS_pair domain
FR751574	6	hypothetical protein	*P. tricornutum*	2e-07	–
FR751567	6	sulfatase	*E. siliculosus*	3e-24	Sulfatase superfamily
FR751569	6	conserved unknown protein	*E. siliculosus*	3e-11	–

being protein kinases as well as genes related to protein recycling (Supplementary File S3). The importance of signaling for toxin-producing, bloom-forming algae seems plausible, as several of them are thought to regulate their toxin production (Granéli and Flynn 2006) as well as gene expression of toxin-related genes (Freitag et al. 2011; Wohlrab et al. 2010; Yang et al. 2011) depending on environmental conditions and specific signals e.g. from grazers. More importantly, toxins produced by microalgae benefit the entire bloom rather than the individual cell (Pohnert et al. 2007), thus, to a certain degree, resembling the behavior of cells in multicellular organisms. The mechanisms underlying the evolution of such "altruistic behavior" are poorly understood, but "chemical awareness" of surrounding cells and therefore cell-cell signaling may be important.

Polyunsaturated Fatty Acids

After having examined *P. farcimen* ESTs on a global scale, we focused our attention on specific metabolic pathways. Given the very particular

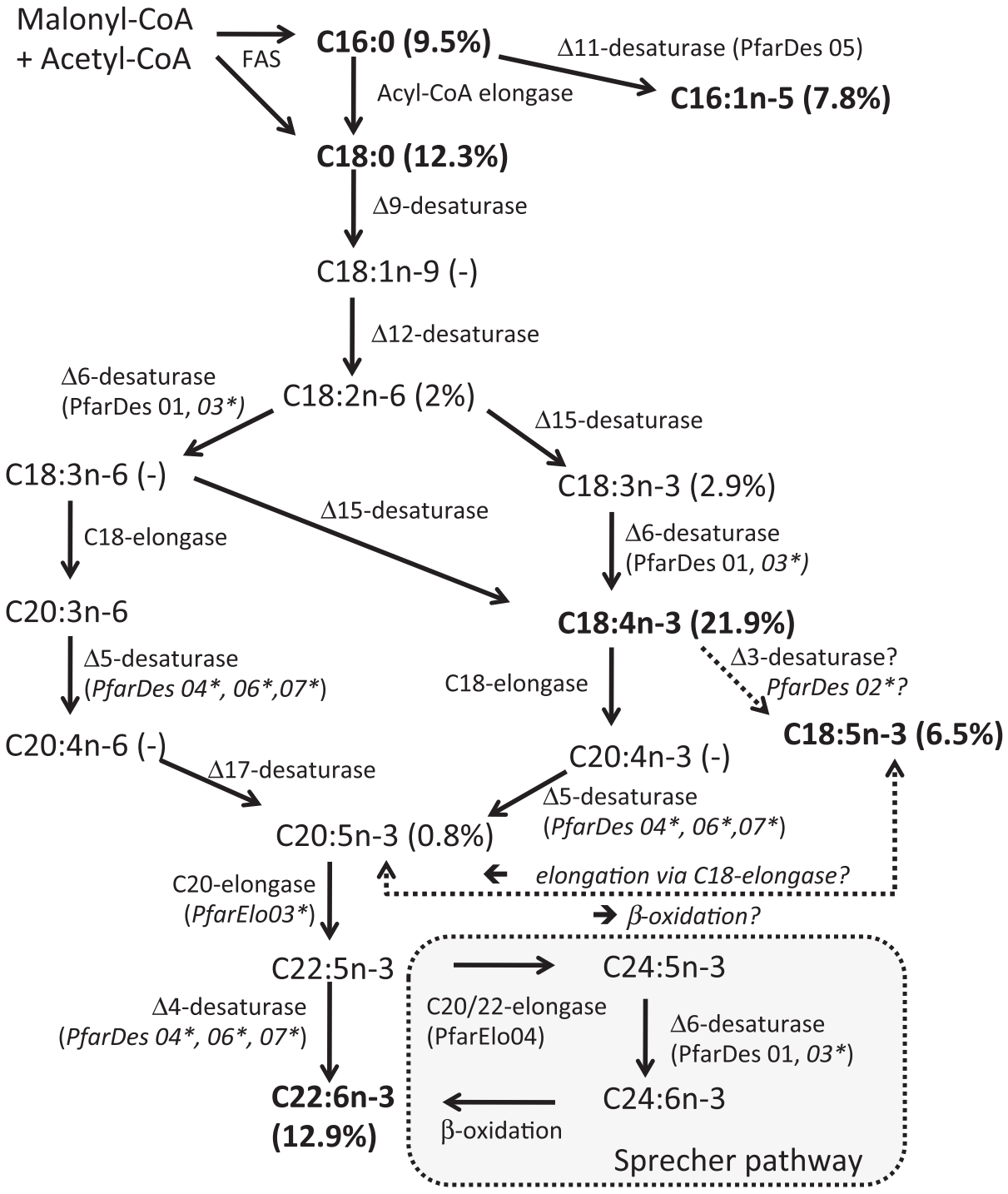

Figure 3. Putative pathways of fatty acid synthesis in *P. farcimen* based on Guschina and Harwood (2006), Meyer et al. (2003), Pereira et al. (2004), Qiu et al. (2001), and Tonon et al. (2003, 2005). The percentage of total fatty acids measured by Ginger et al. (2008) is given in parentheses after its name ("-" = not detected), uncharacterized candidate genes in *P. farcimen* are listed in parentheses under the name of the enzyme, where sequences in italics and marked with "*" grouped with haptophyte sequences in a phylogenetic analysis (see Table 4). Fatty acids in boldface were particularly abundant. C22:5n-3 and C20:3n-6 were not examined. Dotted arrows and question marks indicate hypothetical biosynthetic pathways that could be suggested by the PUFA composition of *P. farcimen,* but have not yet been described.

Table 3. Enriched GO annotations in *Pseudochattonella farcimen*. The table shows the total number of annotations as well as the number of annotations for each enriched category and their percentage with respect to the total number of annotations for each species. Pfa = *Pseudochattonella farcimen*, Iga = *Isochrysis galbana*, Noc = *Nannochloropsis oculata*. Aan = *Aureococcus anophagefferens*, Pca = *Phytophthora capsici*, Esi = *Ectocarpus siliculosus*, Ptr = *Phaeodactylum tricornutum*. The GO term Multicellular organismal process (GO:0032501) comprises also developmental process (GO:0032502), multicellular organismal development (GO:0007275), system development (GO:0048731), organ development (GO:0048513), and anatomical structure development (GO:0048856), which were also significantly overrepresented in *P. farcimen*. All differences except those between *P. farcimen* and *I. galbana* were statistically significant (binomial test, false discovery rate < 5%).

Term	Pfa	Iga	Noc	Aan	Pca	Esi	Ptr
Total annotations	2,921	1,633	509	4,761	1,2445	3,436	82,376
protein binding GO:0005515	230 (7.87%)	94 (5.76%)	15 (2.95%)	233 (4.89%)	812 (6.52%)	192 (5.59%)	4,238 (5.14%)
Signaling GO:0023052	90 (3.08%)	29 (1.78%)	1 (0.20%)	62 (1.30%)	245 (1.97%)	46 (1.34%)	1,278 (1.55%)
Multicellular organismal process GO:0032501	157 (5.37%)	55 (3.37%)	1 (0.20%)	60 (1.26%)	105 (0.84%)	83 (2.42%)	957 (1.16%)

PUFA composition of *Pseudochattonella* sp. (Giner et al. 2008), as well as the possible role of certain PUFAs as toxins (Jüttner 2001; Marshall et al. 2003; Yasumoto et al. 1990) or precursors for toxins (Pohnert 2002), PUFA metabolism was one of them. We did not find the *P. farcimen* EST library to contain any sequences coding for part of the 0 of the fatty acid elongase complex (FR752203), a 3-ketoacyl-CoA reductase with a similar sequence (CBJ30207.1) in *E. siliculosus* was found. These enzymes are normally involved in the synthesis of saturated or monounsaturated fatty acids, which serve as substrate for subsequent elongation and desaturation reactions (Fig. 3). This absence may, however, be explained by overall low expression levels of these genes, as supported by the observation that the same genes in *E. siliculosus* have little (FAS) or no EST-support despite their presence in the genome.

An interesting feature, however, is the presence of eight different cDNAs encoding fatty acid desaturases and six cDNAs encoding enzymes likely to be involved in the condensation reaction during the elongation of specific long-chain (C18, C20) fatty acids (elongases). In several cases the exact specificity of the corresponding enzymes was difficult to deduce based merely on sequence homology, and our EST library does not cover the complete set of expressed genes. This may explain why certain desaturases (Δ9, Δ12, and Δ15), were not identified in our dataset. Moreover, the fact that most of the sequences were incomplete also prevented us from performing one global phylogenetic analysis of all desaturases and elongases, respectively. This underlines the need for heterologous expression experiments and additional sequencing to completely describe fatty acid metabolism in *P. farcimen.* Nevertheless, individual phylogenies with sets of closely related sequences were possible with our data and are consequently described in the following sections.

DHA Synthesis via Δ4-Desaturases

One of the prominent features of *Pseudochattonella* sp. is the high ratio of DHA to EPA (Giner et al. 2008). In mammals, DHA is synthesized from EPA in low quantities via a pathway known as Sprecher's shunt (Sprecher et al. 1999), involving the elongation of C22:5n-3, desaturation via a Δ6-desaturase, and subsequent β-oxidation (Fig. 3). In contrast, in several unicellular eukaryotes, an alternative pathway exists, involving elongation of EPA via a highly specific C20 elongase and subsequent desaturation at the Δ4-position (Meyer

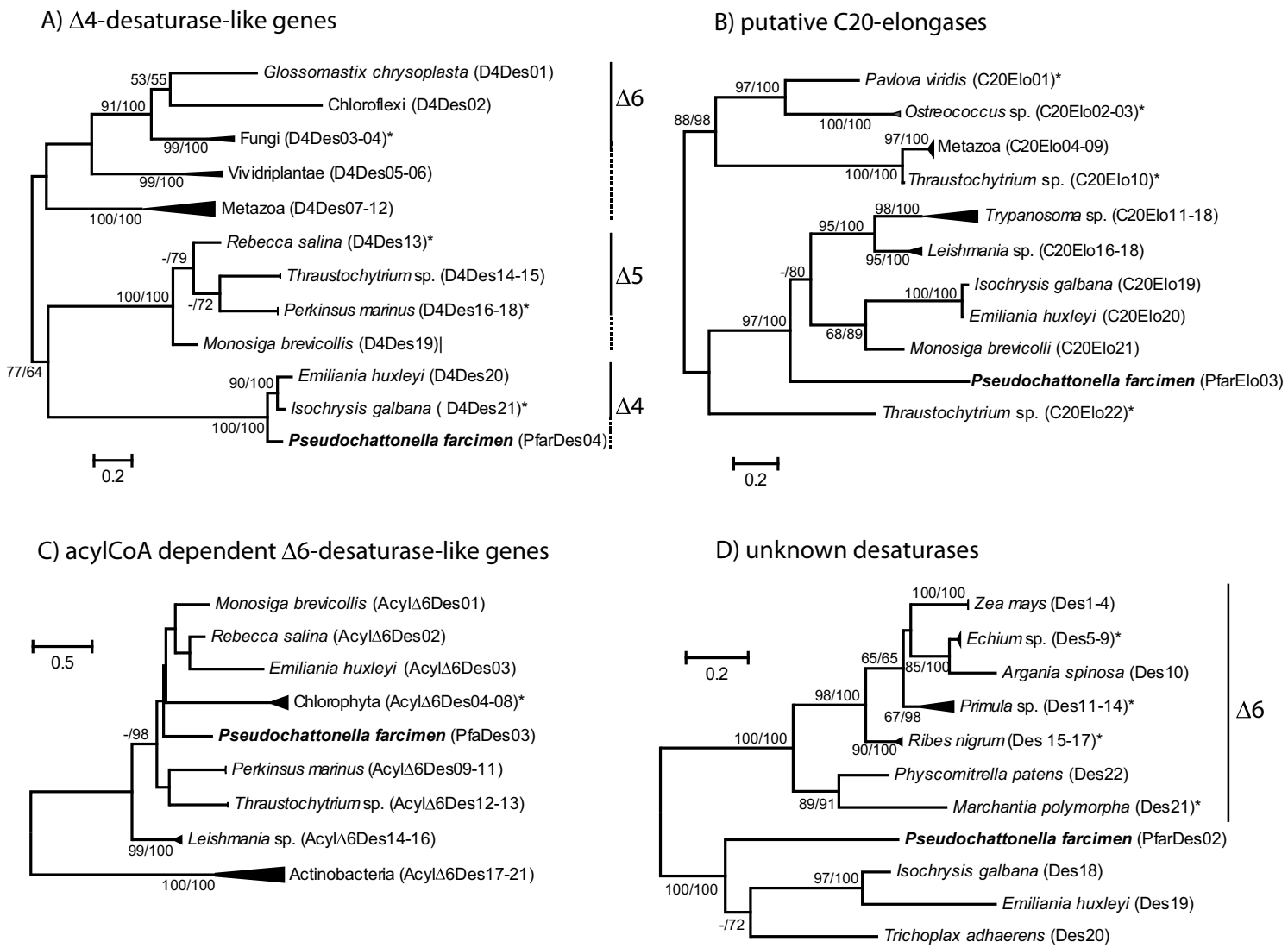

Figure 4. Unrooted maximum likelihood trees of four fatty acid-related enzymes from *P. farcimen* and similar sequences found in GenBank as well as the *E. huxleyi* genome. Confidence values correspond to bootstrap values (PhyML) and posterior probabilities (MrBayes), respectively. "*" indicates branches with functionally characterized sequences. Please note that none of the *P. farcimen* sequences have been functionally characterized so far. Sequence names are given in parenthesis and the corresponding accession numbers and references for the characterization are available in Supplementary file S4. In panels **A** and **C**, horizontal lines group sequences likely to have the same specificity as the characterized enzyme in the branch (given to the right of the line). Dotted lines indicate low confidence for the functional annotation.

Table 4. Fatty acid desaturases (PfarDes) and elongases (PfarElo) found in the *Pseudochattonella farcimen* EST library, as well as their closest relative(s) determined by phylogenetic analyses (PhyML, see Methods). IPR005804 = InterPro domain "Fatty acid desaturase, type 1"; IPR002076 InterPro domain "GNS1/SUR4 membrane protein", ELO family; cd03506 = "Delta6 fatty acid desaturase" domain; PLN03198 = "Delta6-acyl-lipid desaturase, provisional".

Name	Accession	Putative specificity	Domains	Closest relatives found in
PfarDes01	FR738286	Δ5, Δ6, or Δ8	IPR005804, cd03506	diatoms
PfarDes02	FR751741	unknown, Δ6-like	IPR005804, cd03506	haptophytes, *Trichoplax*
PfarDes03	FR751989	Δ6, acetyl-COA-dependent	IPR005804, cd03506, PLN03198	haptophytes, chlorophytes, choanoflagellates, heterokonts, alveolates
PfarDes04	FR739830	Δ5, Δ4	IPR005804, cd03506	haptophytes
PfarDes05	FR735162	Δ11	IPR005804, cd03506, PLN03198	*Thalassiosira pseudonana*
PfarDes06	FR736683	Δ5, Δ4	IPR005804, cd03506, PLN03198	haptophytes
PfarDes07	FR743137	Δ5, Δ4	IPR005804, cd03506, PLN03198	haptophytes
PfarDes08	FR752357	unknown	IPR005804	choanoflagellates
PfarElo01	FR740265	unknown	IPR002076	heterokonts
PfarElo02	FR735502	unknown	IPR002076	*Ectocarpus*
PfarElo03	FR737170	C20	IPR002076	haptophytes, choanoflagellates, excavates
PfarElo04	FR735622	unknown	IPR002076	haptophytes, heterokonts
PfarElo05	FR739139	unknown	IPR002076	*Nannochloropsis*
PfarElo06	FR742737	unknown	IPR002076	*Nannochloropsis*
PfarElo07	FR735397	unknown	IPR002076	*Nannochloropsis*

et al. 2003; Pereira et al. 2004; Qiu et al. 2001; Tonon et al. 2003). In *P. farcimen* homologous sequences to both of these genes were found, and in particular three very similar copies or splice variants, exhibiting a high percentage of identity with Δ4-desaturases previously characterized (Pereira et al. 2004), were identified: PfarDes04, PfarDes06, PfarDes07 (Table 4). This provides a plausible explanation for the comparatively low EPA and high DHA levels. In addition, it is interesting to note that for all three of these potential desaturases, the closest relatives were found in haptophytes (Table 4; Fig. 4A). Similar findings were also obtained for the elongase PfarElo03 (Table 4), although closely related sequences were also found in choanoflagellates and excavates (Fig. 4B).

Heterokont Acyl-CoA-Dependent Δ6-Desaturase and Δ11-Desaturase Genes

Another interesting feature is the presence of a putative acyl-CoA dependent Δ6-desaturase (PfarDes03). Acyl-CoA dependent Δ6-desaturases have been reported in green algae (Domergue et al. 2005; Hoffmann et al. 2008) as well as mammals and fungi, but are not a common feature (Tocher et al. 1998) in heterokonts, as illustrated

by their absence from the genomes of *E. siliculosus*, *P. tricornutum*, and *T. pseudonana*. Homologs of PfarDes03 were found in the choanoflagellate *Monosiga brevicollis,* the labyrinthulomycete *Thraustochytrium* sp. (Heterokonta), the alveolate *Perkinsus marinus* (Alveolata), prasinophytes, as well as haptophytes (Fig. 4C). An additional related sequence from *I. galbana* (gi|106827449 and gi|106819369) was not included in the phylogenetic tree, because of the short overlap with the available sequence from *P. farcimen*.

Acyl-CoA dependent desaturases are of particular interest when it comes to engineering plants with increased PUFA contents (Graham et al. 2007). Most desaturases known in terrestrial plants act primarily on phosphatidylcholine (PC)-bound fatty acids. In contrast, acyl-CoA dependent desaturases have the capacity to introduce double bonds in acyl-CoA-bound fatty acids. Since Δ6 elongases usually also act on acyl-CoA-bound substrates, acyl-CoA dependent desaturases circumvent the transfer of fatty acids from the PC pool to the CoA pool, which is usually the rate-limiting step in the production of PUFAs (Hoffmann et al. 2008).

Another interesting gene with similar sequences in other heterokonts was PfarDes05. This protein is closely related to a Δ11-desaturase previously characterized in *T. pseudonana* (Table 4), which was shown to specifically produce C16:1n5 from C16:0 (Tonon et al. 2004). Δ11-desaturases are frequently found in insects, and are rare among plants (Tonon et al. 2004), yet the presence of this gene in *P. farcimen* could explain why the rare fatty acid C16:1n5 was detected at relatively high levels in this organism (Giner et al. 2008).

Potential Candidates for C18:5n-3 Synthesis

Generally, the similarity between the PUFA profiles of *Pseudochattonella* and the haptophyte *I. galbana* (Giner et al. 2008) is likely to be related to the presence of a highly similar set of PUFA-metabolizing genes in both organisms, as reported above. This may also be true for C18:5n-3, a rare PUFA with a yet unknown biosynthetic pathway. Although it is possible that both organisms obtained (or retained) a biosynthetic pathway for this PUFA independently, the simplest explanation would be that this common feature is also related to an evolutionarily common set of enzymes. We can envision two pathways for the synthesis of C18:5n-3. The first is the ß-oxidation of C20:5n-3. This mechanism is employed in mammals to synthesize DHA via the Sprecher-pathway, but yields only very low quantities of PUFAs. The second pathway is the desaturation of C18:4n-3 by means of a Δ3-desaturase (Fig. 3). As *P. farcimen* and *I. galbana* both contain high concentrations of C18:5n-3, we focused on this latter option and searched for desaturase-like genes conserved between *P. farcimen* and haptophytes, with unknown specificity. Indeed, we found one candidate, PfarDes02, a sequence related to Δ6-desaturases, meeting this criterion (Fig. 4D). This gene is an excellent candidate for heterologous expression experiments.

The extensive set of genes related to fatty acid metabolism found in *P. farcimen* may also be, at least partially, related to its toxicity. Free PUFAs may, for example, act directly as toxins, as suggested for the haptophyte *Chrysochromulina polylepis* and the dinophyte *Gyrodinium aureolum* (Yasumoto et al. 1990), for benthic diatoms, where they constitute a possible grazer defense mechanism (Jüttner 2001), and, in combination with reactive oxygen species, for the raphidophyte *Chattonella marina* (Marshall et al. 2003). In diatoms, PUFAs may be released from phospholipids via the wound-activated activity of a Phospholipase A2 (PLPA2) (Pohnert 2002). Besides Ca-independent phospholipase genes, which can be found in a wide range of organisms, including *P. tricornutum*, *T. pseudonana* (both generally considered non-toxic diatoms), *E. huxleyi,* and *E. siliculosus*, the *P. farcimen* EST library contains two Ca-dependent PLPA2 genes (FR734514 and FR740898) with similar sequences in other heterokonts, mainly in *P. infestans*.

In diatoms, PLPA2 activity is also suggested to be responsible for the initiation of polyunsaturated aldehyde (PUA) production (Pohnert 2002), a class of compounds shown to have toxic effects in several taxonomic groups including bacteria, fungi, algae, mollusks, copepods, and human cell lines (Adolph et al. 2004), although the concentrations required for the observation of such effects were high in most cases. No putative homologs for known lipoxygenase genes were found in the sequence data available for *P. farcimen*, but the presence of several PLPA2 genes nevertheless provides a first indication that this alga may release free PUFAs and possibly also PUAs. Furthermore, PLPA2 s, together with acyltransferases (ATs), constitute key enzymes in the Lands cycle (Lands 1960), a metabolic pathway responsible for the remodeling of glycerophospholipids and mainly found in mammals, but also e.g. in excavates (Das 2001). Good candidates for ATs can indeed be found in the *P. farcimen* EST library, such as a putative 1-acyl-sn-glycerol-3-phosphate AT (FR738784), or two

putative lysophosphatidylcholine ATs (FR734706 and FR740991), providing an indication that the Lands cycle may also be active in *P. farcimen*.

Polyketides

A class of compounds of particular interest because of their possible implication in the toxicity of *P. farcimen,* are polyketides. These secondary metabolites are biosynthetically and evolutionarily closely related to fatty acids and have numerous functions in nature, including chemical defense and cell communication (Bender et al. 1999; Hopwood and Sherman 1990; Staunton and Weissman 2001). They have previously been related to toxicity, e.g. in *Karenia brevis, Alexandrium ostenfeldii, Chrysochromulina polylepis*, and *Prymnesium parvum* (John et al. 2010; MacKinnon et al. 2006; Monroe and Van Dolah 2008; Freitag et al. 2011). Polyketides are synthesized via the activity of polyketide synthases (PKSs). Three types of PKS are presently known. Type I PKSs were previously reported in several alveolate organisms and haptophytes, but not heterokonts (John et al. 2008, Monroe and Van Dolah 2008); type II PKSs are considered to be restricted to bacteria; and type III PKSs have been described in members of the green lineage, in the genome of *E. siliculosus* (Cock et al. 2010), and in bacteria (Gross et al. 2006). In the *P. farcimen* EST library, we found three cDNA sequences of putative PKSs, one of a putative type I PKS and two from type III PKSs.

Type I Polyketide Synthases

The fragment of the putative type I PKS (FR752500, termed PfarPKS1 hereafter) consists of two EST reads that cover mainly the acyl transferase (AT) domain. Although, except for one uncharacterized sequence generated as part of the zebra finch genome project, closely related sequences were found mainly in different bacteria (proteobacteria, verrucomicroabia, and actinobacteria, Fig. 5A), the overall identity of PfarPKS1 with the most similar bacterial sequence was only 28%. Direct comparison of the AT domains in randomly chosen bacterial PKSs yielded considerably higher levels of identity between sequences from different genera (between 62% and 76%). Furthermore, the G/C-content of PfarPKS1 was 55% and thus congruent to the average for *P. farcimen*. Together, this provides a strong indication that, in spite of its similarity to bacterial sequences, PfarPKS1 does probably not merely constitute bacterial contamination, at least not from a bacterial phylum currently represented in public protein databases.

Interestingly, in the case of PfarPKS1, no similar haptophyte sequences were found (e-value 4e-06 for *E. huxleyi*), although the presence of type I PKS sequences in this phylum has previously been reported (John et al. 2008). Unfortunately, our EST data cover only 259 amino acid residues comprising mainly the highly variable AT domain of the protein, and strong phylogenetic signals with this potential type I PKS were not possible to obtain. However, the lack of similarity with type I PKS sequences from haptophytes as well as the absence of these genes in the genomes of *T. pseudonana*, *P. tricornutum* (John et al. 2008), and *E. siliculosus* (Cock et al. 2010) could potentially point to an independent acquisition of the type I PKS in *P. farcimen* or one of its ancestors.

Type III Polyketide Synthases

In addition to the putative type I PKS fragment, we identified two type III PKS fragments (FR739730 and FR738409, named PfarPKS2 and PfarPKS3 hereafter). In contrast to PfarPKS1, these sequences were similar to sequences identified in the brown alga *E. siliculosus* and also *Sargassum muticum*. Due to the high sequence identity between PFarPKS2 and PfarPKS3 (95%), only the longer of the two sequences (PfarPKS3) was considered for a phylogenetic analysis (Fig. 5B). As for the type I PKS fragment, the heterokont type III PKSs seemed to be most closely related to bacterial sequences, in this case exclusively actinobacteria, and again these genes have not been detected in the genomes of the diatoms *P. tricornutum* or *T. pseudonana*. Similar observations have recently been made in cyanobacteria and dinoflagellates, where genes related to saxitoxin A production are likely to be derived from actinobacterial, proteobacterial, or, in the latter case, possibly cyanobacterial sources (Moustafa et al. 2009, Stüken et al. 2011). Moreover, in the brown alga *E. siliculosus*, genes involved in the synthesis of mannitol, alginate, and hemicellulose also grouped with actinobacterial sequences, and were not found in sequenced diatom genomes. This finding led Michel *et al.* (2010a, b) to propose that these metabolic pathways were acquired via horizontal gene transfer (HGT) after the separation of brown algae and diatoms. In this context, and given the phylogenetic tree presented in Figure 5, it would seem possible that type III PKSs in *P. farcimen*, and possibly also type I PKSs, could have been acquired during the same event.

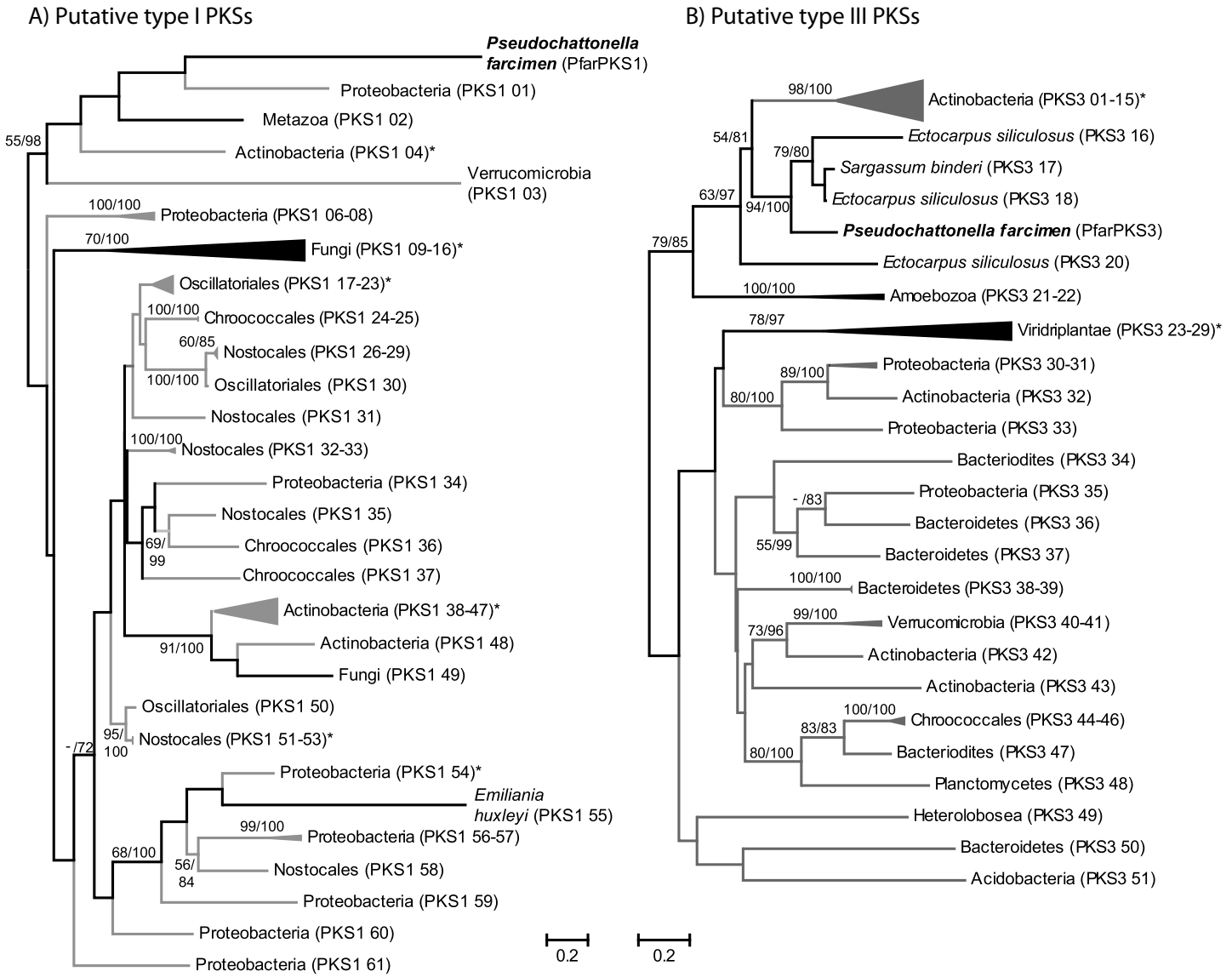

Figure 5. Unrooted maximum likelihood tree of PfarPKS1 (**A**), PfarPKS3 (**B**), and similar sequences found in GenBank. Confidence values correspond to bootstrap values (PhyML) and posterior probabilities (MrBayes) respectively. "*" indicates branches with functionally characterized type III PKSs. Branches with only bacterial sequences are displayed in grey. Sequence names are given in parenthesis and the corresponding accession numbers and references for the characterization are available in Supplementary file S4.

A Complete Set of Genes for Mannitol Metabolism

In order to further explore the evolutionary history of these genes, we searched the *P. farcimen* EST library for genes thought to have been acquired by brown algae during this event. We did not detect any GDP-mannose 6-dehydrogenases or mannuronan C5-epimerases - both enzymes that catalyze the last steps of alginate synthesis. However, our analysis revealed several genes related to mannitol metabolism in *P. farcimen*.

Mannitol is a sugar alcohol, suggested to function as a carbon storage compound, in osmoregulation (Dickson and Kirst 1987), and in stress response (Dittami et al. 2011a) by scavenging of active oxygen species (Iwamoto and Shiraiwa 2005). It is of particular importance in brown algae, where concentrations can reach 30% of the algal dry weight (Zubia et al. 2008), and where strong diurnal (Gravot et al. 2010) and seasonal (Iwao et al. 2008) changes have been observed. It has, however, also been detected in red algae (Karsten et al. 1992) and prasinophytes (Kirst 1975). In brown algae, mannitol is synthesized from Fructose-6P via the activity of a manntiol-1-phosphate dehydrogenase (M1PDH) and a mannitol-1-phosphatase (M1Pase), and converted back into fructose via the activity of a mannitol-2-dehydrogenase (M2DH) (Iwamoto and Shiraiwa 2005; Rousvoal et al. 2011).

In *P. farcimen,* we found candidates for each of the three enzymatic activities: a M1PDH gene consisting of two reads (FR752741, called PfarM1PDH hereafter), a M1Pase supported by two reads (FR751736, PfarM1Pase hereafter), and a M2DH (FR743166, PfarM2DH hereafter). We also searched for hexokinases, which are responsible for the activation of fructose to fructose-1-phosphate and found one possible candidate, FR742747. This latter sequence, however, encodes only a short fragment of the sequence, making a reliable annotation impossible.

Phylogenetic analyses with PfarM1PDH, PfarM1Pase, and PfarM2DH (Fig. 6), revealed all three sequences to be highly similar to genes found in *E. siliculosus* and *Micromonas*, the latter having itself acquired mannitol metabolism from heterokonts (Michel et al. 2010b). This supports the hypothesis that these genes may have been acquired via HGT in a common ancestor of *P. farcimen* and *E. siliculosus*. Since no traces of such an event were found in diatoms, this would imply that these genes could have been lost in the latter lineage, or that dictyochophyes and brown algae may be more closely related to each other than to diatoms. Although the phylogenetic relationships among heterokonts have not been completely resolved (Riisberg et al. 2009), the latter hypothesis is in agreement with a recent phylogenetic reconstruction by Brown and Sorhannus (2010).

The presence of mannitol-related genes in *P. farcimen* supports our hypothesis that HGT is the mechanism behind the origin of heterokont PKSs (type III and possibly also type I). It also suggests that *P. farcimen* may produce mannitol – a hypothesis, which was recently confirmed by GC/MS (Dittami et al. 2011b), although the detected levels were low under standard conditions. Further studies on this topic will be required to elucidate the physiological role of mannitol in *P. farcimen*.

Concluding Remarks

The analysis of the *P. farcimen* EST library points to several interesting features with respect to the evolution of heterokonts in general, as well as the biology and possible mechanisms underlying the toxicity of the species.

From an evolutionary point of view, our study shows that *P. farcimen* has probably obtained genes and metabolic pathways by HGT from at least two different organisms. First, a set of genes involved in PUFA metabolism, which is highly conserved with haptophytes, exists parallel to a set of typical heterokont genes. One plausible explanation for the presence of these genes in *P. farcimen* can be found in recent hypotheses that postulate that, contrary to the classical chromalveolate hypothesis (Cavalier-Smith 1999), at least some heterokonts may have obtained their plastids by one or more tertiary endosymbiosis events involving the uptake of a haptophyte symbiont (Archibald 2009; Sanchez-Puerta and Delwiche 2008). Although molecular evidence against the existence of a monophyletic kingdom of chromalvelolates is increasing (Baurain et al. 2010; Burki et al. 2007; Stiller et al. 2009), sampling of a wider range of heterokont and haptophyte taxa will be required to resolve when and how this set of haptophyte genes was obtained.

Our data also show traces of a HGT from gram-positive bacteria (probably actinobacteria) proposed by Michel et al. (2010b), notably the presence of a complete set of genes involved in mannitol metabolism. This HGT therefore probably predates the separation of dictyochophytes and brown algae, supporting the idea that the latter two classes may be more closely related to each other than to diatoms (Brown and Sorhannus 2010), where no traces of this event were detected. This

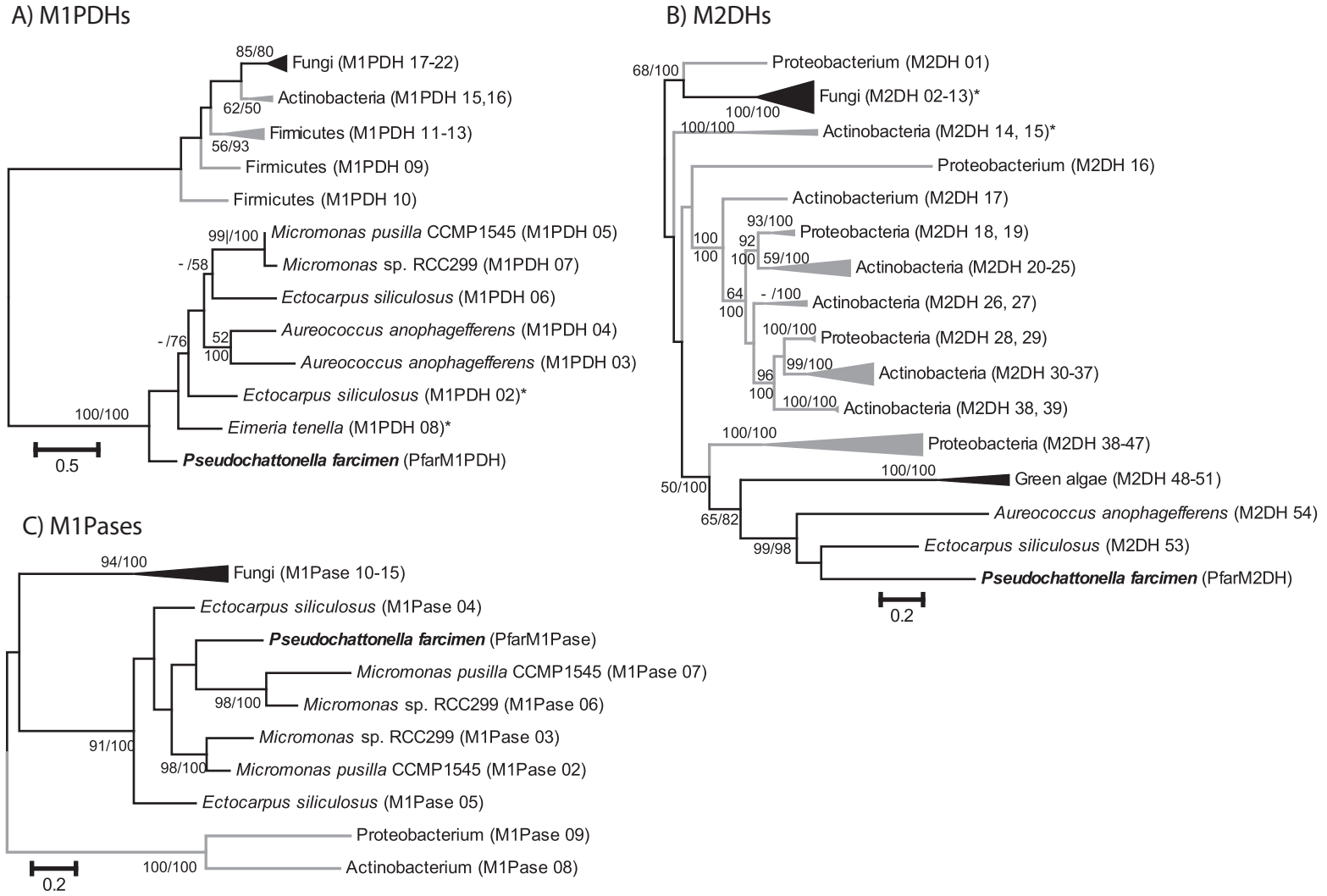

Figure 6. Unrooted maximum likelihood tree of M1PDHs (**A**), M2DHs (**B**), and MPases (**C**). (**A**) and (**C**) are based on phylogenies previously published by Michel et al. (2010b). Confidence values correspond to bootstrap values (PhyML) and posterior probabilities (MrBayes) respectively. "*" indicates branches with functionally characterized type III PKSs. Branches with only bacterial sequences are displayed in grey. Sequence names are given in parenthesis and the corresponding accession numbers and references for the characterization are available in Supplementary file S4.

or a similar event may also be at the origin of the putative type I and type III PKS fragments found in *P. farcimen*.

From a biological perspective, the gene content of *P. farcimen* explains some of the unusual features previously observed in this species, notably its PUFA profiles. Furthermore, we found genes and pathways that could potentially be related to the ichthyotoxic effects of *P. farcimen*, genes involved in the liberation of PUFAs and two sets of PKSs. Further experimental work will be required to determine, if free PUFAs, PUFA-derived substances such as PUAs, or PKS-derived toxins are produced, and if these molecules can explain the ichthyotoxic effects of this species. Finally, we detected all genes necessary to complete the mannitol cycle, providing a strong indication that this alga is, like brown algae, capable of producing and metabolizing this compound. The physiological role of mannitol in *P. farcimen*, however, remains to be explored.

Methods

Cell cultures and antibiotics treatments: *P. farcimen* strain UIO109 was isolated from water samples off the south coast of Norway on March 28th 2001, as described in detail by Edvardsen et al. (2007). Cultures of this strain were grown in IMR1/2 medium (Eppley et al. 1967) with added selenite (10 nM final concentration) at a salinity of 25 PSU and were kept at 14-15 °C under white fluorescent light with a photon flux rate of about 100 μmol photons m^{-2} s^{-1} and a 12:12 h light:dark cycle.

To obtain axenic cultures, 10 g L^{-1} penicillin, 2.5 g L^{-1} streptomycin, and 2.5 g L^{-1} gentamicin were dissolved in molecular grade water. Algal cultures in the exponential growth phase were treated with 0.5%, 1%, 2%, or 4% (v/v) of this antibiotics mixture and supplemented with one droplet IMR1/2 with added yeast extract and tryptone. After three days, treated cultures were used to inoculate fresh IMR1/2 medium as well as bacterial growth medium (IMR1/2 with 1 g L^{-1} tryptone and 0.25 g L^{-1} yeast extract). The latter was incubated for one week at room temperature in the dark. Cultures without detectable bacterial contamination were selected and the antibiotics treatment was repeated. Finally the efficiency of the antibiotics treatment was also verified by flow cytometry, where a decrease in the relative abundance of small particles was observed in the treated cultures.

RNA extraction, library construction, and sequencing: Axenic cultures were harvested in the exponential growth phase by gentle centrifugation (10 min) at 3,200 *g* and at 4 °C. The pellet was immediately frozen in liquid nitrogen, and total RNA was isolated using the Qiagen RNeasy plant mini kit (Qiagen, Hilden, Germany) and DNase treated with DNase I (Fermentas, Ontario, Canada) according to the manufacturer's recommendations.

cDNA was synthesized from total RNA using oligo-dT primer. Normalization was achieved by one round of denaturation and reassociation. Double-stranded cDNA was separated from the remaining single-stranded cDNA (normalized cDNA) by passing the mixture over a hydroxylapatite column. The EST library was constructed and transformed into electro-competent *Escherichia coli* cells by Vertis Biotechnologie AG (Freising-Weihenstephan, Germany). Colonies were picked, and the DNA was extracted by magnetic beads on a robot platform (Qiagen, Hilden, Germany). Plasmid inserts were sequenced from both sides using Big Dye Chemistry (Applied Biosystems, Darmstadt, Germany) and separated on an ABI Prism 3700 sequencing platform (Applied Biosystems).

Sequence analysis and functional annotation: Sequence and quality information was extracted from the sequencing trace files using Phred on the BioPortal (www.bioportal.uio.no) at the University of Oslo (Kumar et al. 2009). Low quality sequences as well as contaminants were removed by repeatedly running SeqClean in combination with NCBI's generic vector database UniVec. Cleaned EST sequences were deposited in the EMBL Nucleotide Sequence Database under accession numbers FR734407-FR744448 and assembled using TGICL (Quackenbush et al. 2000) and default parameters (hsp_length $\geq$40, frac_identical $\geq$0.95, unmatched_overhang $\leq$20). The assembled sequences (contigs) were submitted to the EMBL Transcriptome Shotgun Assembly (TSA) archive under accession numbers FR751558-FR752797. Both the contigs and the remaining singletons were blasted (blastx) against the NCBI nr protein database using the Bioportal (www.bioportal.uio.no, e-value cutoff 1e-6) and the results were used for the prediction of open reading frames using OrfPredictor (Min et al. 2005). Predicted protein sequences were automatically annotated with Gene Ontology (GO) terms using the Blast2GO software (Götz et al. 2008) with default parameters and an e-value cutoff of 1e-10; GO terms were then mapped to GOSlim terms using the AGBase GOSlimViewer (McCarthy et al. 2006) and the plant GOSlim set. To assist manual annotation all sequences were also submitted to InterProScan (Zdobnov and Apweiler 2001) using default parameters.

This entire procedure was carried out for *P. farcimen* ESTs but in parallel also for the publicly available ESTs from *A. anophagefferens* (JGI, unpublished), *Nannochloropsis oculata* (Shi et al. 2008), *P. tricornutum* (Maheswari et al. 2009), *Phytophthora capsici* (JGI, unpublished), *E. siliculosus* (Dittami et al. 2009) and *I. galbana* (Patron et al. 2006) (see Table 1 for details). The composition of the different EST libraries was compared in terms of the assigned GO annotations using the GOLEM software (Sealfon et al. 2006) and allowing a false discovery rate of 5%.

Phylogenetic analyses: Automatic protein alignments of sequences considered for phylogenetic analyses were generated using COBALT (Papadopoulos and Agarwala 2007), and manually refined using Jalview (Waterhouse et al. 2009). Poorly aligned positions and divergent regions were removed using the Gblocks server (Talavera and Castresana 2007), allowing smaller final blocks and less strict flanking positions. Phylogenetic trees were reconstructed by maximum likelihood using PhyML 3.0 (Guindon and Gascuel 2003). The LG substitution model (Le and Gascuel 2008) was selected using ProtTest (Abascal et al. 2005), and default parameters were used (4 substitution rate categories, estimated gamma distribution parameters). The reliability of the resulting trees was tested by bootstrap analysis, running 500 iterations. In addition, Bayesian interference analyses were performed using the parallel version of MrBayes 3.1 (Huelsenbeck et al. 2001). The analysis was run for 100,000 to 1,000,000 generations using default parameters (Poisson model, equal state frequencies), until the standard deviation of split frequencies was below 0.01. Trees were sampled every 100 generations and the first 25% of the trees were not considered for the calculation of Bayesian posterior probabilities (Burn-in). The resulting consensus trees were plotted using MEGA 4.0 (Kumar et al. 2008).

Phylogenetic analyses carried out to determine the closest relatives for sequences presented in Table 4 were also carried out as described above using PhyML, but bootstrap analyses were replaced by approximate likelihood-ratio tests.

Acknowledgements

We would like to thank Thierry Tonon, Kamran Shalchian-Tabrizi, and the anonymous reviewers for helpful discussions and critical reading of the manuscript. Furthermore, we are grateful to Berit Smestad Paulsen for testing the presence of mannitol in *P. farcimen*. SMD and BE received funding through the EU 7th Framework Programme (FP7-ENV-2007-1-*MIDTAL*-201724) and KSJ was supported by the Research Council of Norway.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.protis.2011.07.004.

References

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics **21**:2104–2105

Adolph S, Bach S, Blondel M, Cueff A, Moreau M, Pohnert G, Poulet SA, Wichard T, Zuccaro A (2004) Cytotoxicity of diatom-derived oxylipins in organisms belonging to different phyla. J Exp Biol **207**:2935–2946

Archibald JM (2009) The puzzle of plastid evolution. Curr Biol **19**:R81–R88

Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kroger N, Lau WW, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science **306**:79–86

Aure J, Danielssen DS, Skogen M, Svendsen E, Dahl E, Søiland H, Petterson L (2001) Environmental Conditions during the *Chattonella* Bloom in the North Sea and Skagerrak in May 1998. In Hallegraeff GM, Bolch CJS, Blackburn I, Lewis R (eds) Harmful Algal Blooms 2000. Intergovernmental Oceanographic Commission of UNESCO, Paris, pp 55–82

Backe-Hansen P, Dahl E, Danielssen DS (2001) On the Bloom of *Chattonella* in the North-Sea/Skagerrak in April–May 1998. In Hallegraeff GM, Bolch CJS, Blackburn I, Lewis R (eds) Harmful Algal Blooms 2000. Intergovernmental Oceanographic Commission of UNESCO, Paris, pp 78–81

Baurain D, Brinkmann H, Petersen J, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V, Roger AJ, Burger G, Lang BF, Philippe H (2010) Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. Mol Biol Evol **27**:1698–1709

Ben Ali A, De Baere R, De Wachter R, Van De Peer Y (2002) Evolutionary relationships among heterokont algae (the autotrophic stramenopiles) based on combined analyses of small and large subunit ribosomal RNA. Protist **153**:123–132

Bender C, Rangaswamy V, Loper J (1999) Polyketide production by plant-associated pseudomonads. Ann Rev Phytopathol **37**:175–196

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. J Roy Stat Soc B Met **57**:289–300

Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret J-P, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, Fournet J, Haruta M, Huysman MJJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert Y, Kaplan A, Kröger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-Jézéquel V, Lopez Pascal J, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A, Oudot-Le Secq M-P, Napoli C, Obornik M, Schnitzler-Parker M, Petit J-L, Porcel BM, Poulsen N, Robison M, Rychlewski L, Rynearson TA, Schmutz J, Shapiro H, Siaut M, Stanley M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van De Peer Y, Grigoriev IV (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. Nature **456**:239–244

Brown JW, Sorhannus U (2010) A molecular genetic timescale for the diversification of autotrophic stramenopiles (Ochrophyta): substantive underestimation of putative fossil ages. PloS ONE **5**:e12759

Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland A, Nikolaev SI, Jakobsen KS, Pawlowski J (2007) Phylogenomics reshuffles the eukaryotic supergroups. PloS ONE **2**:e790

Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol **46**:347–366

Cock JM, Sterck L, Rouzé P, Scornet D, Allen AE, Amoutzias G, Anthouard V, Artiguenave F, Aury J-M, Badger Jonathan H, Beszteri B, Billiau K, Bonnet E, Bothwell JH, Bowler C, Boyen C, Brownlee C, Carrano CJ, Charrier B, Cho GY, Coelho SM, Collén J, Corre E, Da Silva C, Delage L, Delaroque N, Dittami SM, Doulbeau S, Elias M, Farnham G, Gachon CMM, Gschloessl B, Heesch S, Jabbari K, Jubin C, Kawai H, Kimura K, Kloareg B, Küpper FC, Lang D, Le Bail A, Leblanc C, Lerouge P, Lohr M, Lopez PJ, Martens C, Maumus F, Michel G, Miranda-Saavedra D, Morales J, Moreau H, Motomura T, Nagasato C, Napoli CA, Nelson DR, Nyvall-Collén P, Peters AF, Pommier C, Potin P, Poulain J, Quesneville H, Read B, Rensing SA, Ritter A, Rousvoal S, Samanta M, Samson G, Schroeder DC, Ségurens B, Strittmatter M, Tonon T, Tregear JW, Valentin K, von Dassow

P, Yamagishi T, Van De Peer Y, Wincker P (2010) The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. Nature **465**:617–621

Das S (2001) Phospholipid remodeling/generation in *Giardia*: the role of the Lands cycle. Trends Parasitol **17**:316–319

Dickson DMJ, Kirst GO (1987) Osmotic adjustment in marine eukaryotic algae: the role of inorganic ions, quaternary ammonium, tertiary sulphonium and carbohydrate solutes: II. Prasinophytes and haptophytes. New Phytol **106**:657–666

Dittami SM, Aas HTN, Smestad Paulsen B, Boyen C, Edvardsen B, Tonon T (2011b) Mannitol in six autotrophic stramenopiles and *Micromonas*. Plant Signal Behav **6**(8): PMID: 21720212

Dittami SM, Gravot A, Renault D, Goulitquer S, Eggert A, Bouchereau A, Boyen C, Tonon T (2011a) Integrative analysis of metabolite and transcript abundance during the short-term response to saline and oxidative stress in the brown alga *Ectocarpus siliculosus*. Plant Cell Environ **34**:629–642

Dittami SM, Scornet D, Petit J-L, Ségurens B, Da Silva C, Corre E, Dondrup M, Glatting K-H, König R, Sterck L, Rouzé P, Van De Peer Y, Cock JM, Boyen C, Tonon T (2009) Global expression analysis of the brown alga *Ectocarpus siliculosus* (Phaeophyceae) reveals large-scale reprogramming of the transcriptome in response to abiotic stress. Genome Biol **10**:R66

Domergue F, Abbadi A, Zähringer U, Moreau H, Heinz E (2005) In vivo characterization of the first acyl-CoA Delta6-desaturase from a member of the plant kingdom, the microalga *Ostreococcus tauri*. Biochem J **389**:483–490

Edler L. (2006) Algal situation in marine waters surrounding Sweden. AlgAware Oceanographic Unit. 1

Edvardsen B, Eikrem W, Shalchian-Tabrizi K, Riisberg I, Johnsen G, Naustvoll L, Throndsen J (2007) *Verrucophora farcimen* gen. et sp. nov. (Dictyochophyceae, Heterokonta) - a bloom-forming ichthyotoxic flagellate from the Skagerrak, Norway. J Phycol **43**:1054–1070

Eikrem W, Edvardsen B, Throndsen J (2009) Research note: Renaming *Verrucophora farcimen* Eikrem, Edvardsen et Throndsen. Phycol Res **57**:170

Eppley RW, Holmes RW, Strickland JDH (1967) Sinking rates of marine phytoplankton measured with a fluorometer. J Exp Mar Biol Ecol **1**:191–208

Freitag M, Beszteri S, Vogel H, John U (2011) Induced toxicity and polyketide synthase gene expression following physiological shock in the toxigenic *Prymnesium parvum*. Eur J Phycol, **in press**

Giner J-L, Zhao H, Tomas C (2008) Sterols and fatty acids of three harmful algae previously assigned as *Chattonella*. Phytochemistry **69**:2167–2171

Graham IA, Larson T, Napier JA (2007) Rational metabolic engineering of transgenic plants for biosynthesis of omega-3 polyunsaturates. Curr Opin Biotech **18**:142–147

Granéli E, Flynn K (2006) Chemical and Physical Factors Influencing Toxin Content. In Granéli E, Turner JT (eds) Ecology of Harmful Algae. Springer, Berlin, Heidelberg, pp 229–241

Gravot A, Dittami SM, Rousvoal S, Lugan R, Eggert A, Collén J, Boyen C, Bouchereau A, Tonon T (2010) Diurnal oscillations of metabolite abundances and gene analysis provide new insights into central metabolic processes of the brown alga *Ectocarpus siliculosus*. New Phytol **188**: 98–110

Gross F, Luniak N, Perlova O, Gaitatzis N, Jenke-Kodama H, Gerth K, Gottschalk D, Dittmann E, Müller R (2006) Bacterial type III polyketide synthases: phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads. Arch Microbiol **185**:28–38

Gschlössl B, Guermeur Y, Cock JM (2008) HECTAR: a method to predict subcellular targeting in heterokonts. BMC Bioinformatics **9**:393

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52**:696–704

Guschina IA, Harwood JL (2006) Lipids and lipid metabolism in eukaryotic algae. Progr Lipid Res **45**:160–186

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res **36**:3420–3435

Hoffmann M, Wagner M, Abbadi A, Fulda M, Feussner I (2008) Metabolic engineering of omega3-very long chain polyunsaturated fatty acid production by an exclusively acyl-CoA-dependent pathway. J Biol Chem **283**:22352–22362

Hopwood DA, Sherman DH (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu Rev Genet **24**:37–66

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. Science **294**:2310–2314

Iwamoto K, Shiraiwa Y (2005) Salt-regulated mannitol metabolism in algae. Mar Biotechnol **7**:407–415

Iwao T, Kurashima A, Maegawa M (2008) Effect of seasonal changes in the photosynthates mannitol and laminaran on maturation of *Ecklonia cava* (Phaeophyceae, Laminariales) in Nishiki Bay, central Japan. Phycol Res **56**:1–6

John U, Beszteri S, Glöckner G, Singh R, Medlin L, Cembella A (2010) Genomic characterisation of the ichthyotoxic prymnesiophyte *Chrysochromulina polylepis*, and the expression of polyketide synthase genes in synchronized cultures. Eur J Phycol **45**:215–229

John U, Beszteri B, Derelle E, Van de Peer Y, Read B, Moreau H, Cembella A (2008) Novel insights into evolution of protistan polyketide synthases through phylogenomic analysis. Protist **159**:21–30

John U, Tillmann U, Medlin LK (2002) A comparative approach to study inhibition of grazing and lipid composition of a toxic and non-toxic clone of *Chrysochromulina polylepis* (Prymnesiophyceae). Harmful Algae **1**:45–57

Joseph SJ, Fernández-Robledo JA, Gardner MJ, El-Sayed NM, Kuo C-H, Schott EJ, Wang H, Kissinger JC, Vasta GR (2010) The alveolate *Perkinsus marinus*: biological insights from EST gene discovery. BMC Genomics **11**:228

Jüttner F (2001) Liberation of 5,8,11,14,17-eicosapentaenoic acid and other polyunsaturated fatty acids from lipids as a

grazer defense reaction in epilithic diatom biofilms. J Phycol **37**:744–755

Karsten U, West J, Mostaert A, King R (1992) Mannitol in the red algal genus *Caloglossa* (Harvey) J. Agardh. J Plant Physiol **140**:292–297

Kirst GO (1975) Correlation between content of mannitol and osmotic-stress in brackish-water alga *Platymonas subcordiformis* (Hazen). Z Pflanzenphysiol **76**:316–325

Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform **9**:299–306

Kumar S, Skjaeveland A, Orr RJS, Enger P, Ruden T, Mevik B-H, Burki F, Botnen A, Shalchian-Tabrizi K (2009) AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. BMC Bioinformatics **10**:357

Lands W (1960) Metabolism of glycerolipids: II. The enzymatic acylation of lysolecithin. J Biol Chem **235**:2233–2237

Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol **25**:1307–1320

MacKinnon SL, Cembella AD, Burton IW, Lewis N, LeBlanc P, Walter JA (2006) Biosynthesis of 13-desmethyl spirolide C by the dinoflagellate *Alexandrium ostenfeldii*. J Org Chem **71**:8724–8731

Maheswari U, Mock T, Armbrust EV, Bowler C (2009) Update of the Diatom EST Database: a new tool for digital transcriptomics. Nucleic Acids Res **37**:D1001–D1005

Mansour M (1999) Very-long-chain (C28) highly unsaturated fatty acids in marine dinoflagellates. Phytochemistry **50**:541–548

Marshall J, Nichols PD, Hamilton B, Lewis RJ, Hallegraef GM (2003) Ichthyotoxicity of *Chattonella marina* (Raphidophyceae) to damselfish (*Acanthochromis polycanthus*): the synergistic role of reactive oxygen species and free fatty acids. Harmful Algae **2**:273–281

McCarthy FM, Wang N, Magee GB, Nanduri B, Lawrence ML, Camon EB, Barrell DG, Hill DP, Dolan ME, Williams WP, Luthe DS, Bridges SM, Burgess SC (2006) AgBase: a functional genomics resource for agriculture. BMC Genomics **7**:229

Meyer A, Cirpus P, Ott C, Schlecker R, Zähringer U, Heinz E (2003) Biosynthesis of docosahexaenoic acid in *Euglena gracilis*: biochemical and molecular evidence for the involvement of a Delta4-fatty acyl group desaturase. Biochemistry **42**:9779–9788

Michel G, Tonon T, Scornet D, Cock JM, Kloareg B (2010a) The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in eukaryotes. New Phytol **188**:82–97

Michel G, Tonon T, Scornet D, Cock JM, Kloareg B (2010b) Central and storage carbon metabolism of the brown alga *Ectocarpus siliculosus*: insights into the origin and evolution of storage carbohydrates in eukaryotes. New Phytol **188**: 67–81

Min XJ, Butler G, Storms R, Tsang A (2005) OrfPredictor: predicting protein-coding regions in EST-derived sequences. Nucleic Acids Res **33**:W677–W680

Monroe EA, Van Dolah FM (2008) The toxic dinoflagellate *Karenia brevis* encodes novel type I-like polyketide synthases containing discrete catalytic domains. Protist **159**:471–482

Mostaert AS, Karsten U, Hara Y, Watanabe MM (1998) Pigments and fatty acids of marine raphidophytes: A chemotaxonomic re-evaluation. Phycologia **46**:213–220

Moustafa A, Loram JE, Hackett JD, Anderson DM, Plumley FG, Bhattacharya D (2009) Origin of saxitoxin biosynthetic genes in cyanobacteria. PloS ONE **4**:e5758

Papadopoulos JS, Agarwala R (2007) COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics **23**:1073–1079

Patron NJ, Waller RF, Keeling PJ (2006) A tertiary plastid uses genes from two endosymbionts. J Mol Biol **357**:1373–1382

Pereira SL, Leonard AE, Huang Y-S, Chuang L-T, Mukerji P (2004) Identification of two novel microalgal enzymes involved in the conversion of the omega3-fatty acid, eicosapentaenoic acid, into docosahexaenoic acid. Biochem J **384**:357–366

Pohnert G (2002) Phospholipase A2 activity triggers the wound-activated chemical defense in the diatom *Thalassiosira rotula*. Plant Physiol **129**:103–111

Pohnert G, Steinke M, Tollrian R (2007) Chemical cues, defence metabolites and the shaping of pelagic interspecific interactions. Trends Ecol Evol **22**:198–204

Qiu X, Hong H, MacKenzie SL (2001) Identification of a delta 4 fatty acid desaturase from *Thraustochytrium* sp. involved in the biosynthesis of docosahexanoic acid by heterologous expression in *Saccharomyces cerevisiae* and *Brassica juncea*. J Biol Chem **276**:31561–31566

Quackenbush J, Liang F, Holt I, Pertea G, Upton J (2000) The TIGR gene indices: reconstruction and representation of expressed gene sequences. Nucleic Acids Res **28**:141–145

Renaud S (1999) The gross chemical composition and fatty acid composition of 18 species of tropical Australian microalgae for possible use in mariculture. Aquaculture **170**:147–159

Riisberg I, Orr RJS, Kluge R, Shalchian-Tabrizi K, Bowers HA, Patil V, Edvardsen B, Jakobsen KS (2009) Seven gene phylogeny of heterokonts. Protist **160**:191–204

Rousvoal S, Groisillier A, Dittami SM, Michel G, Boyen C, Tonon T (2011) Mannitol-1-phosphate dehydrogenase activity in *Ectocarpus siliculosus*, a key role for mannitol synthesis in brown algae. Planta **233**:261–273

Sanchez-Puerta MV, Delwiche CF (2008) A hypothesis for plastid evolution in chromalveolates. J Phycol **44**:1097–1107

Sealfon RS, Hibbs MA, Huttenhower C, Myers CL, Troyanskaya OG (2006) GOLEM: an interactive graph-based gene-ontology navigation and analysis tool. BMC Bioinformatics **7**:443

Shi J, Pan K, Yu J, Zhu B, Yang G, Yu W, Zhang X (2008) Analysis of expressed sequence tags from the marine microalga *Nannochloropsis oculata* (Eustigmatophyceae). J Phycol **44**:99–102

Skjelbred B, Horsberg TE, Tollefsen KE, Andersen T, Edvardsen B (2011) Toxicity of the ichthyotoxic marine flagellate *Pseudochattonella* (Dictyochophyceae, Heterokonta) assessed by six bioassays. Harmful Algae **10**:144–157

Sprecher H, Chen Q, Yin FQ (1999) Regulation of the biosynthesis of 22:5n-6 and 22:6n-3: a complex intracellular process. Lipids **34(Suppl)**:S153–S156

Staunton J, Weissman KJ (2001) Polyketide biosynthesis: a millennium review. Nat Prod Rep **18**:380–416

Stiller JW, Huang J, Ding Q, Tian J, Goodwillie C (2009) Are algal genes in nonphotosynthetic protists evidence of historical plastid endosymbioses? BMC Genomics **10**:484

Stüken A, Orr RJS, Kellmann R, Murray SA, Neilan BA, Jakobsen KS (2011) Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. PloS ONE **6**:e20096

Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol **56**:564–577

Tocher DR, Leaver MJ, Hodgson PA (1998) Recent advances in the biochemistry and molecular biology of fatty acyl desaturases. Progr Lipid Res **37**:73–117

Tonon T, Harvey D, Larson TR, Graham IA (2003) Identification of a very long chain polyunsaturated fatty acid Delta4-desaturase from the microalga *Pavlova lutheri*. FEBS Lett **553**:440–444

Tonon T, Harvey D, Qing R, Li Y, Larson TR, Graham IA (2004) Identification of a fatty acid Δ11-desaturase from the microalga *Thalassiosira pseudonana*. FEBS Lett **563**:28–34

Tonon T, Harvey D, Larson TR, Li Y, Graham IA (2005) Acyl-CoA elongase activity and gene from the marine microalga *Pavlova lutheri* (Haptophyceae). J Appl Phycol **17**:111–118

Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RHY, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, Chapman J, Damasceno CMB, Dorrance AE, Dou D, Dickerman AW, Dubchak IL, Garbelotto M, Gijzen M, Gordon SG, Govers F, Grunwald NJ, Huang W, Ivors KL, Jones RW, Kamoun S, Krampis K, Lamour KH, Lee M-K, McDonald WH, Medina M, Meijer HJG, Nordberg EK, Maclean DJ, Ospina-Giraldo MD, Morris PF, Phuntumart V, Putnam NH, Rash S, Rose JKC, Sakihama Y, Salamov AA, Savidor A, Scheuring CF, Smith BM, Sobral BWS, Terry A, Torto-Alalibo TA, Win J, Xu Z, Zhang H, Grigoriev IV, Rokhsar DS, Boore JL (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science **313**:1261–1266

Viso A, Marty J (1993) Fatty acids from 28 marine microalgae. Phytochemistry **34**:1521–1533

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics **25**:1189–1191

Wohlrab S, Iversen MH, John U (2010) A molecular and co-evolutionary context for grazer induced toxin production in *Alexandrium tamarense*. PloS ONE **5**:e15039

Yang I, Selander E, Pavia H, John U (2011) Grazer-induced toxin formation in dinoflagellates: a transcriptomic model study. Eur J Phycol **46**:66–73

Yasumoto T, Underdal B, Aune T, Hormazabal V, Skulberg O, Oshima Y (1990) Screening for Hemolytic and Ichthyotoxic Components of *Chrysochromulina polylepis* and *Gyrodinium aureolum* from Norwegian Coastal Waters. In Granéli E, Sundström B, Edler L, Anderson DM (eds) Toxic Mmarine Phytoplankton. Elsevier, New York, pp 436–440

Zdobnov EM, Apweiler R (2001) InterProScan–an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17**:847–848

Zubia M, Payri C, Deslandes E (2008) Alginate, mannitol, phenolic compounds and biological activities of two range-extending brown algae, *Sargassum mangarevense* and *Turbinaria ornata* (Phaeophyta: Fucales), from Tahiti (French Polynesia). J Appl Phycol **20**:1033–1043