

OPEN

Diatoms synthesize sterols by inclusion of animal and fungal genes in the plant pathway

Carmela Gallo^{1,4}, Simone Landi^{1,2,4}, Giuliana d'Ippolito^{1,4*}, Genoveffa Nuzzo¹, Emiliano Manzo¹, Angela Sardo^{1,3} & Angelo Fontana^{1*}

Diatoms are ubiquitous microalgae that have developed remarkable metabolic plasticity and gene diversification. Here we report the first elucidation of the complete biosynthesis of sterols in the lineage. The study has been carried out on the bloom-forming species *Skeletonema marinoi* and *Cyclotella cryptica* that synthesise an ensemble of sterols with chemotypes of animals (cholesterol and desmosterol), plants (dihydrobrassicasterol and 24-methylene cholesterol), algae (fucosterol) and marine invertebrates (clonasterol). In both species, sterols derive from mevalonate through cyclization of squalene to cycloartenol by cycloartenol synthase. The pathway anticipates synthesis of cholesterol by enzymes of the phytosterol route in plants, as recently reported in Solanaceae. Major divergences stem from reduction of $\Delta 24(28)$ and $\Delta 24(25)$ double bonds which, in diatoms, are apparently dependent on sterol reductases of fungi, algae and animals. Phylogenetic comparison revealed a good level of similarity between the sterol biosynthetic genes of *S. marinoi* and *C. cryptica* with those in the genomes of the other diatoms sequenced so far.

Sterols are vital components of all eukaryotic cells where they modulate structure and function of membranes and, as precursors of signaling molecules, they control growth and development. Plant sterols, commonly named phytosterols, are involved in morphogenesis, development, reproduction and stress response^{1–3}. Analogously, microalgal sterols showed critical physiological roles related to photosynthesis, growth, light response and fatty acid metabolism⁴.

Sterols are terpenes deriving from a complex process of polymerization of six isoprene units. In animals and fungi, the mevalonic acid (MVA) pathway is the only route for the biosynthesis of the isoprene units isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). In higher plants and algae, IPP and DMAPP derive either from the MVA pathway in the cytoplasm or the methylerythritol phosphate (MEP) pathway in the plastids⁵. Without a clear cellular compartmentalization, both biochemical routes have been also reported in mosses and streptomyces. Sterol biosynthesis from IPP and DMAPP is generally taxa-specific and proceeds via lanosterol by lanosterol synthase (LSS) in nonphotosynthetic organisms (e.g. animals and fungi) or cycloartenol by cycloartenol synthase (CAS) in photosynthetic lineages (e.g. plants and algae)⁶. Cholesterol is the major animal sterol but its biosynthesis in tomato has been recently shown to derive from a cycloartenol-dependent pathway composed of enzymes either shared with phytosterols or evolved from phytosterol biosynthetic genes⁷. Furthermore, insects and lower eukaryotes, including marine invertebrates and a few microalgae, are suggested to convert phytosterols to cholesterol and *vice versa*^{8–11}.

Diatoms represent an important component of the aquatic ecosystem^{12,13} and form the largest biological group of marine phytoplankton^{14,15}. These microalgae are major global producers by contributing to fixation of CO₂ and geochemical cycles in the world oceans^{16,17}. Furthermore, they can biosynthesize a number of secondary metabolites¹⁸ including domoic acid^{19,20}, oxylipins^{21–24}, carotenoids^{25,26}, as well as organic and inorganic biopolymers such as chitin, silica and calcite^{27–30}. Recently, we showed that a specific class of sterol derivatives, namely sterol sulfates (StS), operate as chemical trigger of cell death programs in the diatom *Skeletonema marinoi*³¹. The report is new but underlines the possible role of sterols to control cell survival during phytoplankton blooms. Diatoms possess

¹National Research Council of Italy, Institute of Biomolecular Chemistry, Bio-Organic Chemistry Unit, Via Campi Flegrei 34, 80078, Pozzuoli (Napoli), Italy. ²University of Naples "Federico II", Department of Biology, Cupa Nuova Cintia 21, 80126, Napoli, Italy. ³Stazione Zoologica "A. Dohrn", Villa Comunale, 80121, Napoli, Italy. ⁴These authors contributed equally: Carmela Gallo, Simone Landi and Giuliana d'Ippolito. *email: gdippolito@icb.cnr.it; afontana@icb.cnr.it

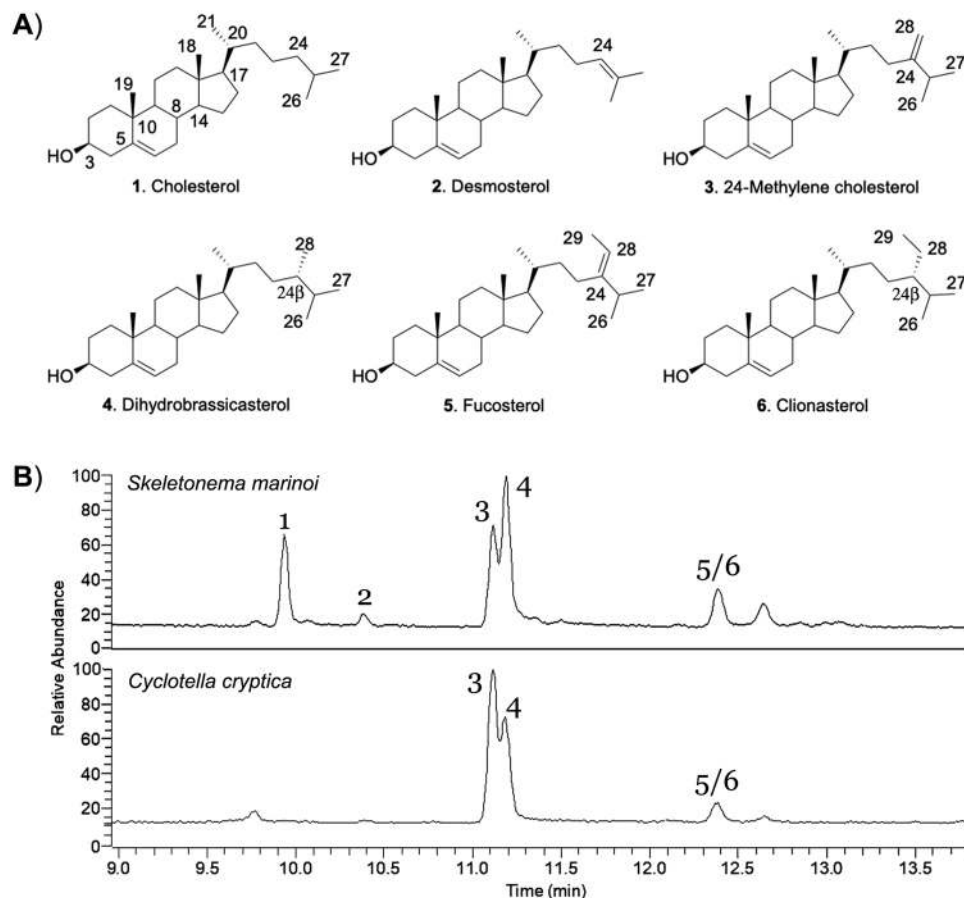


Figure 1. Sterols in *S. marinoi* and *C. cryptica*. **(A)** Structures and **(B)** GC-MS chromatograms of acetylated derivatives. β -chiral descriptor in side chain is in agreement with Nes⁶.

both plant (e.g. brassicasterol) and animal (e.g. cholesterol) sterols³². To explain this unconventional blend, Fabris *et al.* have recently suggested a mixed plant and fungal pathway on the basis of molecular studies in the diatom *Phaeodactylum tricornerutum*³³. However, diatoms usually do not produce ergosterol, the typical fungal sterol, as well as they lack homologs of the eukaryotic proteins required for its synthesis from lanosterol³⁴. On the whole, the “fungal” hypothesis does not have so far biochemical confirmation and sterol biogenesis is still an undisclosed question in these microalgae.

In the present study, we have investigated the biosynthesis of sterols, from acetyl-CoA to the final products, with the aim to find a unified pathway suitable to produce both plant and animal sterols in diatoms. Combination of transcriptomic approach with labelling methods and GC-MS profiling has been used to disclose the origin of the complex mixtures of sterols in the diatoms *Skeletonema marinoi* and *Cyclotella cryptica*. Post analysis of the molecular data, sterol profiling and comparison with other diatom species provides general significance to this study.

Results

Profiling and identification of diatom sterols. Four different Δ^5 -sterols, namely 24-methylene cholesterol, dihydrobrassicasterol, clionasterol (also named γ -sitosterol) and fucosterol were identified by standard comparison in the GC-MS profiles of *C. cryptica* (Fig. 1). The above phytosterols together with cholesterol and minor level of desmosterol were also observed in *S. marinoi*. Dihydrobrassicasterol (β -orientation of the methyl group at C-24) was distinguished from campesterol (α -orientation at C-24) on the basis of the ¹H- and ¹³C-chemical shift of Me-28^{35,36} after purification of the natural product. Analogously, NMR comparison with an original standard confirmed the *E* configuration of fucosterol^{37–39} while stereochemistry at C-24 of clionasterol (β -orientation of the ethyl group) required a complete NMR analysis after purification by HPLC (Supporting Material). The proton chemical shifts of this natural product were very similar to the values reported in the literature for clionasterol^{35,40} but showed differences with a standard of β -sitosterol for H₃-26 (0.8308 vs 0.8371 ppm), H₃-27 (0.8106 vs 0.8163 ppm) and H₃-29 (0.8578 vs 0.8462) (Supplementary Fig. S1). As shown in Table 1, ¹³C NMR data of the diatom product provided additional agreement with the literature³⁵ and corroborated the differences between the two 24-ethyl epimers at C-23 (26.4 vs 26.1 ppm), C-24 (46.1 vs 45.8), C-26 (19.0 vs 19.8) and C-27 (19.6 vs 19.1), thus giving definitive confirmation of the assignment of clionasterol in the diatom extracts.

To provide general significance to these results, we also performed additional GC-MS profiling of the diatoms *Skeletonema costatum*, *Thalassiosira weissflogii*, *Phaeodactylum tricornerutum*, *Pseudonitzschia arenysensis*.

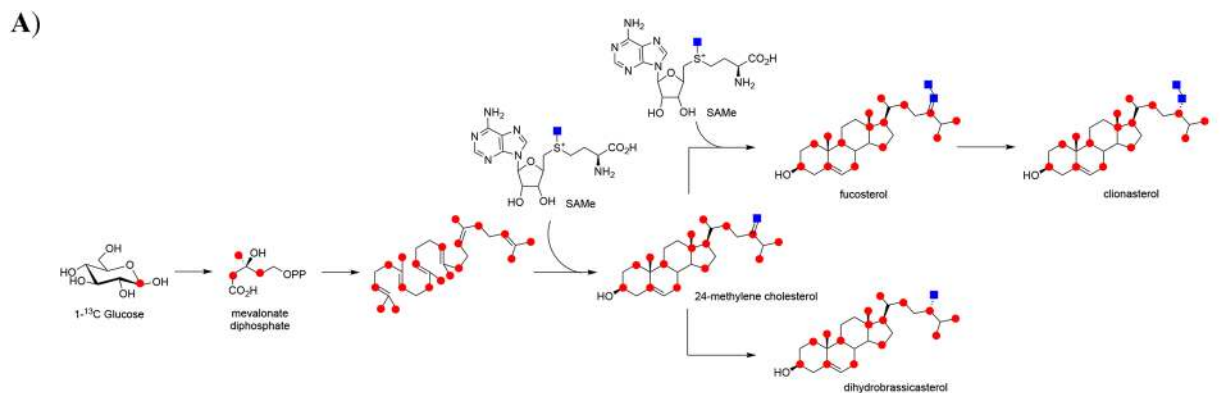
C	Clionasterol ^a	β -Sitosterol	C	Clionasterol ^a	β -Sitosterol
1	37.3	37.3	16	28.2	28.2
2	31.7	31.6	17	56.0	56.0
3	71.8	71.7	18	11.9	11.8
4	42.3 ^b	42.2 ^c	19	19.4	19.4
5	140.8	140.7	20	36.3	36.1
6	121.7	121.7	21	18.8	18.8
7	31.9 ^d	31.9 ^d	22	33.9	33.9
8	31.9 ^d	31.9 ^d	23	26.4	26.1
9	50.1	50.1	24	46.1	45.8
10	36.4	36.5	25	28.9	29.2
11	21.1	21.0	26	19.0	19.8
12	39.8	39.8	27	19.6	19.1
13	42.3 ^b	42.3 ^c	28	23.0	23.0
14	56.8	56.8	29	12.3	12.0
15	24.3	24.3			

Table 1. ¹³C NMR (CDCl₃, 125 MHz) data of clionasterol (24 β -ethyl epimer) from *C. cryptica* and an authentic standard of β -sitosterol (24 α -ethyl epimer). ^aNatural product; ^{b,c,d,e}Overlapped signals.

In agreement with the most recent reports, 24-methylene cholesterol was the recurrent sterol in these diatoms, often accompanied by dihydrobrassicasterol and cholesterol (Supplementary Table S1). In addition, desmosterol and 24-ethyl cholesterol (likely clionasterol) were found in *P. tricornutum* while desmosterol was present in *S. costatum* and *T. weissflogii*. In agreement with the literature^{32,41–44}, the results show an evident variability of the sterol composition in these diatoms but confirm the original assumption that diatoms can in principle biosynthesize sterols from cyclization of both cycloartenol and lanosterol. Another characteristic of the analyzed diatoms is the β configuration at C-24 of dihydrobrassicasterol and clionasterol which is commonly found in organisms, such as algae or fungi, that are considered lower than vascular plants in the evolutionary hierarchy^{45,46}.

Isoprene origin from mevalonate pathway. In order to clarify the sterol biosynthetic pathway of *S. marinoi* and *C. cryptica*, transcriptome assembly and sequence alignments were performed by triplicates of each species. The analysis revealed the whole set of genes of the MVA pathway from acetyl-CoA C-acetyltransferase to isopentenyl diphosphate isomerase (Fig. 2). Experimental support to these results was addressed by labelling studies with [1-¹³C]-glucose. Diatoms are photo-autotrophic organisms thus use of an organic substrate required acclimation of the species to heterotrophic conditions. To this aim, *C. cryptica* was grown heterotrophically as previously reported^{47,48}. After extraction and fractionation of the cell pellets, labeled and unlabeled dihydrobrassicasterol were compared by ¹³C NMR spectroscopy³⁶. Spectra of these products clearly revealed selective incorporation at C-1, C-3, C-5, C-7, C-9, C-13, C-15, C-17, C-18, C-19, C-21, C-22, C-24, C-26, C-27 and C-28 (Supplementary Fig. S2). A few positions such as C-6 (121.74 ppm) and C-16 (28.25 ppm) were not labeled while other carbon atoms (C-1, 37.3 ppm; C-7, 31.9 ppm; C-18, 11.9 ppm; C-19, 19.4 ppm) resulted highly enriched. Mapping of these positions permitted to conclude that labeling was specifically present at C-2, C-4 and C-5 of the six isoprene units that give origin to the rearranged terpene skeleton of dihydrobrassicasterol (Fig. 2). On the contrary, very low or no labeling were observed at carbon positions derived from C-1 of the isoprene units. According to the Rohmer's method⁴⁹, this enrichment study proved the synthesis of dihydrobrassicasterol unambiguously by the mevalonate pathway with no contribution from MEP. Similar results were obtained for fucosterol and clionasterol (Supplementary Figs. S3 and S4). Intense flanking doublets due to incorporation in vicinal carbons were very evident in those signals, namely C-13/C-18 and C-13/C-17, whose proximity is in agreement with the cyclization of the sterol skeleton from squalene. A strong coupling between C-24/C-28 supported labelling of these carbon atoms in dihydrobrassicasterol. Analogous effects were also observed between C-24/C-28 and C-28/C-29 in fucosterol and clionasterol. The carbon atoms C-28 and C-29 of these metabolites do not originate from the isoprene unit but derive from labelled S-adenosyl methionine (SAM) (Fig. 2). In plants, the two reactions are dependent on sterol 24C-methyl transferase 1 (SMT1) and sterol 24C-methyl transferase 2 (SMT2) that catalyze transfer of the methyl group from SAM to cycloartenol or 24-methylenelophenol, respectively. Orthologs of the sequences related to genes of the MVA pathway in *S. marinoi* and *C. cryptica* were also identified in the published genomes of *Thalassiosira oceanica*, *Thalassiosira pseudonana*, *Phaeodactylum tricornutum*, and *Fragilariopsis cylindrus* (Supplementary Table S2).

Cyclization of sterol skeleton by cycloartenol synthase. Table 2 reports the list of transcripts coding for putative enzymes related to synthesis of sterols from IPP and DMAPP in *S. marinoi* and *C. cryptica*. FPKM counts were obtained for the gene toolbox that oversees each biosynthetic step except for squalene epoxidase (SQE) that did not give confident coverage with any known sequence. Despite the conservation of this enzyme in animals, fungi and plants, more than one polyphyletic group of eukaryotes have shown the lack of the corresponding gene. Very recently, an alternative SQE (AltSQE) belonging to the fatty acid hydroxylase family has been described in diatoms⁵⁰. Analysis of the transcripts of *S. marinoi* and *C. cryptica* supported the presence of orthologs of AltSQE (Sm-TR7561 and Cc-TR29442) thus suggesting that this enzyme catalyzes epoxidation



B)

Predicted transcripts function by Blast2GO	<i>S. marinoi</i> identified transcripts	Expression (FPKM counts)			<i>C. cryptica</i> identified transcripts	Expression (FPKM counts)			
		Replicate 1	Replicate 2	Replicate 3		Replicate 1	Replicate 2	Replicate 3	
Acetyl-CoA C-acetyltransferase -1	AACT	TR12914	3590	1972	2204	TR6331	67632	54442	64482
Acetyl-CoA C-acetyltransferase -2	AACT	TR13665	50473	64707	69182	TR10152	7620	13556	14528
Acetyl-CoA C-acetyltransferase -3	AACT	TR19239	72485	119313	121484	-	-	-	-
Hydroxymethylglutaryl-CoA synthase -1	HMGAS	TR6533	17602	13726	14833	TR27129	567	676	823
Hydroxymethylglutaryl-CoA reductase -1	HMGAR	TR11383	36197	30150	33665	TR26098	36553	20240	43821
Mevalonate kinase	MEVK	TR10489	54045	71329	68578	TR35862	1858	1284	2347
Phosphomevalonate kinase	PMEVK	TR8483	629	280	507	TR18131	1090	1945	1147
Diphosphomevalonate decarboxylase -1	MVD	TR8518	893	812	878	TR19178	0.000	0.000	0.000
Diphosphomevalonate decarboxylase -2	MVD	-	-	-	-	TR20824	1446	504	2748
Isopentenyl diphosphate isomerase -1	IDI	TR5135	1531	1295	1034	TR905	7731	11473	7648
Isopentenyl diphosphate isomerase -2	IDI	TR11187	3116	1721	1980				

Figure 2. Mevalonate biosynthesis of sterols in *S.marinoi* and *C. cryptica*. (A) Incorporation of ^{13}C -labelling (red spot and blue square) in dihydrobrassicasterol, fucosterol and clionasterol after feeding of $1\text{-}^{13}\text{C}$ -glucose to *C. cryptica* under heterotrophic conditions. SAMe = S-adenosyl methionine; (B) Transcript sequences of genes coding for enzymes of the mevalonate pathway in *S. marinoi* and *C. cryptica*.

of squalene also in these species. It is worth noting that AltSQE of *S. marinoi* and *C. cryptica*, as well those of other diatoms, showed high sequence similarities (above 60%) with plant $\Delta 5$ -sterol reductase (DWARF7) that is responsible for one of the latest steps of the phytosterol biosynthesis. The successive reaction of cyclization is proposed to give cycloartenol by cycloartenol synthase (CAS) that was significantly upregulated in both *S. marinoi* and *C. cryptica*. Orthologous CAS genes were also identified in the published genomes of *T. oceanica*, *T. pseudonana*, *P. tricornutum* and *F. cylindrus*. The phylogenetic analysis of these sequences showed clusterization of these proteins with plant CASs despite the diatom sequences form an independent group from plants (Fig. 3). Within the lineage, the putative enzymes were splitted in two sub-branches with separation of the centric species *S. marinoi*, *C. cryptica*, *T. pseudonana* and *T. oceanica* from the pennates *P. tricornutum*, *P. arenysensis*, *Fragilariopsis solaris* and *F. cylindrus*. Cyclase activity of *Arabidopsis* CAS1 is strictly dependent on the conserved domains KMQGYNGSQ (406 aa – 414 aa) and TADHGWPISDC (474 aa – 484 aa), with Tyr410, His477 and Ile481 that are key residues for the catalytic mechanism⁵¹. Putative CAS from *S. marinoi* and *C. cryptica* showed domains [(830 aa - 838 aa) and (910 aa - 920 aa) in *S. marinoi*; (353 aa - 361 aa) and (436 aa - 446 aa) in *C. cryptica*] and catalytic residues (Tyr834, His913, Ile917 in *S. marinoi*; Tyr357, His439 and Ile443 in *C. cryptica*) that were identical to those of *A. thaliana* (Supplementary Fig. S5).

C4-Demethylation of the triterpene precursor. C4-demethylation is a crucial step in sterol biosynthesis in plants, mammals and fungi. The biochemical transformation requires the consecutive action of three enzymes widely conserved across phyla, namely sterol-4-methyl oxidase (SMO), 3-hydroxysteroid dehydrogenase/C4-decarboxylase (C4D), and a sterone ketoreductase (SKR)⁵². The complex is tethered to the membrane by ergosterol biosynthetic protein 28 (ERG28), a fourth protein that has no catalytic function but plays a key role in preventing accumulation of biosynthetic 4-methyl sterol intermediates^{53,54}. Removal of both methyl groups occurs successively by a single SMO in mammals and yeast, whereas in plants the process proceeds in two steps under control of two different SMOs, named SMO1 and SMO2. The analysis of the transcripts of *S. marinoi* and *C. cryptica* did not reveal any presence of ERG28. In analogy with phytosterol biosynthesis, we found sequences that correlate with plant SMO1 (Sm-TR11964, Sm-TR30213) and SMO2 (Sm-TR319) in *S. marinoi*, whereas the transcriptome of *C. cryptica* showed only one sequence for a putative SMO1 (Cc-TR32604). It is interesting that the BLAST analysis of the putative SMO1 encoded by Sm-TR11964 and Cc-TR32604 indicated a good homology with the human C4 methyl sterol oxidase that participates in the C4-demethylation of cholesterol precursors (Supplementary Table S3). Both diatoms also revealed sequences related to fungal ERG26, a decarboxylating

Biosynthetic Function	Predicted enzymatic function by Blast2GO		<i>S. marinoi</i> transcripts	Expression (FPKM counts)			<i>C. cryptica</i> transcripts	Expression (FPKM counts)		
				Replicate 1	Replicate 2	Replicate 3		Replicate 1	Replicate 2	Replicate 3
Terpene elongation and cyclization	<i>Farnesyl pyrophosphate synthetase</i>	FPPS	TR5774	747	1469	2779	TR24671	1580	1363	1796
	<i>Squalene synthase</i>	SQS	TR5837	7653	9077	10903	TR905	7731	11473	7648
	<i>Alternative Squalene Epoxidase</i>	AltSQE	TR7561	7708	1005	1765	TR29442	13916	16726	10147
	<i>Cycloartenol synthase</i>	CAS	TR12960	4902	5597	5266	TR3126	2825	3011	3105
Construction of the sterol skeleton	<i>Cycloartenol-C-24-methyltransferase -1</i>	SMT1	TR612	399669	547431	495121	TR44522	4872	4562	10006
	<i>Cycloartenol-C-24-methyltransferase -2</i>	SMT1	—	—	—	—	TR45147	34639	25720	33901
	<i>Sterol Methyl transferase</i>	SMT2	TR11812	19023	13794	15818	TR29102	66731	62025	70507
	<i>Sterol Methyl transferase</i>	SMT2	TR21333	24344	13997	14326	TR17021	33805	42406	37980
	<i>C-4 Methylsterol oxidase</i>	3βHSD-D	TR28093	2287	2523	1960	—	—	—	—
	<i>Sterol-4-methyl oxidase 1</i>	SMO1	TR11964	40534	37090	35927	TR32604	20979	29827	25702
	<i>Sterol-4-methyl oxidase 1</i>	SMO1	TR30213	3061	3258	2926	TR28329	0.000	1731	2017
	<i>Sterol-4-methyl oxidase 2</i>	SMO2	TR319	17019	18736	18676	TR32604	20979	29827	25702
	<i>Sterol-4-α-carboxylate 3-dehydrogenase</i>	ERG26/D4C	TR28078	48329	29541	39087	TR14638	11001	7800	11499
	<i>Cycloeucaenol cycloisomerase</i>	CPI	TR11937	8673	8864	8943	TR30071	3081	3643	4002
	<i>Sterol 14 demethylase -1</i>	CYP	TR7602	5822	3161	3628	TR12199	32581	38398	40078
	<i>Sterol 14 demethylase -2</i>	CYP	—	—	—	—	TR23734	6396	4710	5582
	<i>Sterol 14 demethylase -3</i>	CYP	—	—	—	—	TR30096	70903	65490	74001
	<i>Delta 14 sterol reductase</i>	C14SR	TR13312	5630	5771	5383	TR19404	47799	46207	50117
	<i>Delta 5 Sterol desaturase</i>	Δ5 SD/DWARE7	TR28093	2287	2523	1960	TR48724	0.000	7168	2531
	<i>7 Dehydrocholesterol reductase</i>	Δ7SR/DWARE5	TR2636	7653	17187	13127	TR24068	10801	9765	9443
	<i>Delta(24(24(1)))sterol reductase</i>	24SR/ERG4	TR213	14750	18308	18373	TR29411	88345	78049	66050
	<i>24 dehydrocholesterol reductase -2</i>	24-DHCR	TR10998	45299	64582	64882	—	—	—	—
<i>22 sterol desaturase</i>	ERG5	—	—	—	—	TR3228	11535	12460	13684	

Table 2. Transcript sequences of genes coding for enzymes of sterol biosynthetic pathway in *S. marinoi* and *C. cryptica*.

enzyme of the C4D family that is responsible for the formation of 3 oxo-sterols from several 3β-4-carboxysterols in fungi, plants and mammals.

Biosynthesis of phytosterols. *S. marinoi* and *C. cryptica* transcriptomes exhibited respectively, 13 and 16 expressed genes related to the steps leading to C₂₈ and C₂₉ phytosterols from cycloartenol in agreement with a common plant pathway (Table 2). Expressed genes of key steps in phytosterol biosynthesis showed perfect matching among *S. marinoi* and *C. cryptica*, and high percentage of query coverage and identity with plant genes (Supplementary Table S3 and Supporting Material). Following the labelling results reported above, the analysis revealed orthologs of plant SMT-1 and SMT-2 that are implied in the methylation of cycloartenol and 24-methylenelophenol to give C24-methyl and C24-ethyl derivatives, respectively. On phylogenetic basis,

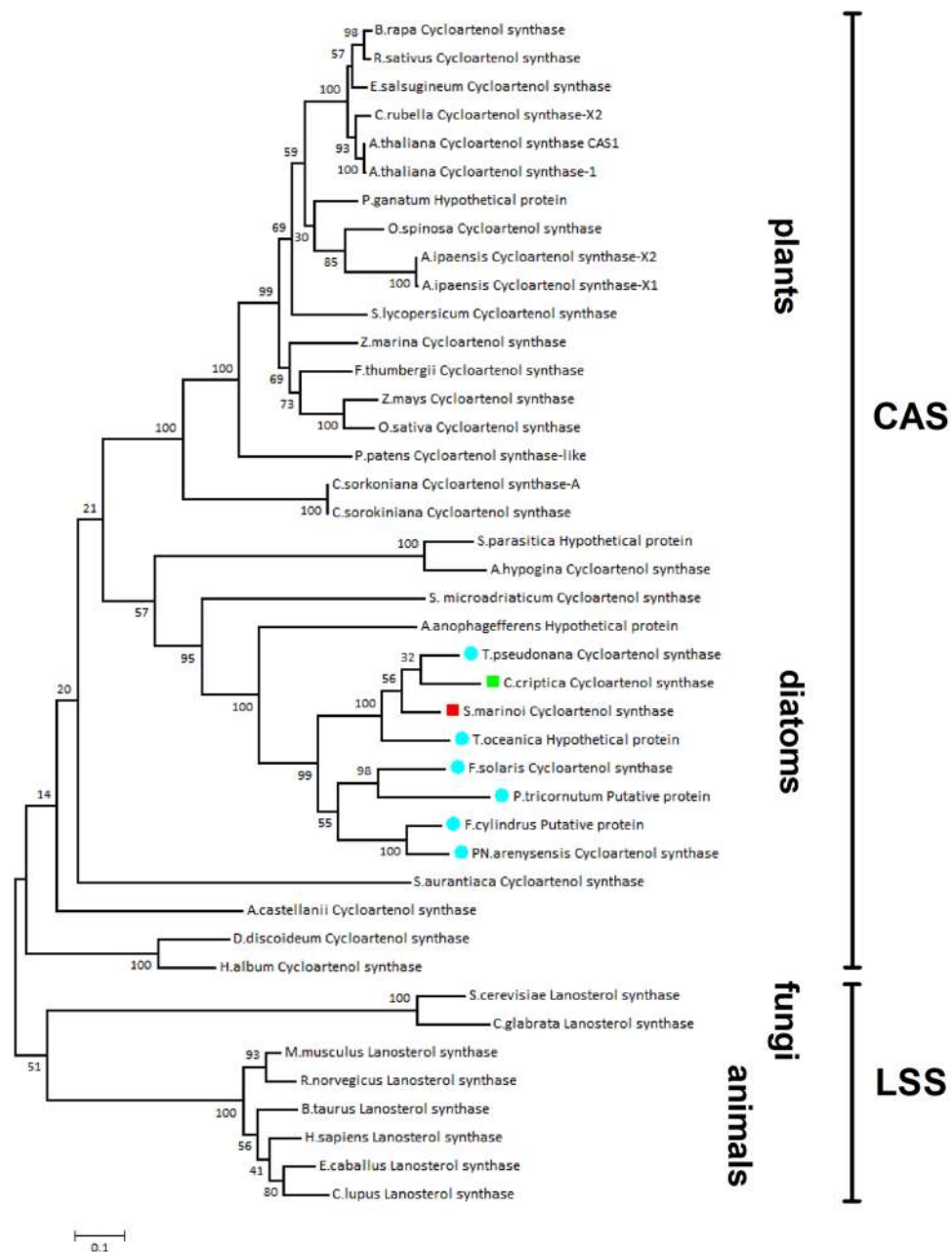


Figure 3. Un-rooted phylogenetic tree of cycloartenol synthase (CAS) and lanosterol synthase (LSS) in diatoms, plants, algae fungi and animals. Alignment used maximum likelihood method. The bootstrapping test (replicate = 100) is reported on each node. Red square = *S. marinoi*; Green square = *C. cryptica*; Light blue spot = other diatom.

Sm-TR612, Cc-TR44522 and Cc-TR45147 were identified as SMT1 orthologs and clustered with sequences of plants, fungi and other diatoms (Fig. 4). On the other hand, Sm-TR11812, Sm-TR21333, Cc-TR29102 and Cc-TR17021 formed an independent group of SMT that diverged from SMT2 of plants and the green alga *Chlamydomonas reinhardtii* but showed homology with sequences of diatoms and the red alga *Chondrus crispus*. Phylogenesis using the plant nicotinate methyltransferase (NANMT)⁵⁵ as outgroup revealed that differentiation of this group of genes from SMT of plants and green algae may have occurred very early and prior to specialization between SMT 1 and SMT2 (Supplementary Fig. S6). We have not proved the catalytic function of this separate class of putative SMT but it is intriguing to speculate that their occurrence is related to the *E* stereochemistry of fucosterol that is found in diatom and red algae in contrast with the *Z* configuration of isofucosterol that is common in plants.

In plants and microalgae, C5-sterol desaturase (DWARF7), 7-dehydrocholesterol reductase (DWARF5) and 24-sterol reductase (DWARF1) control the latest steps in the pathway for isomerization of the double bond in the B-ring and reduction of C24(C28)-double bond in the alkyl chain. Transcripts of *S. marinoi* and *C. cryptica*

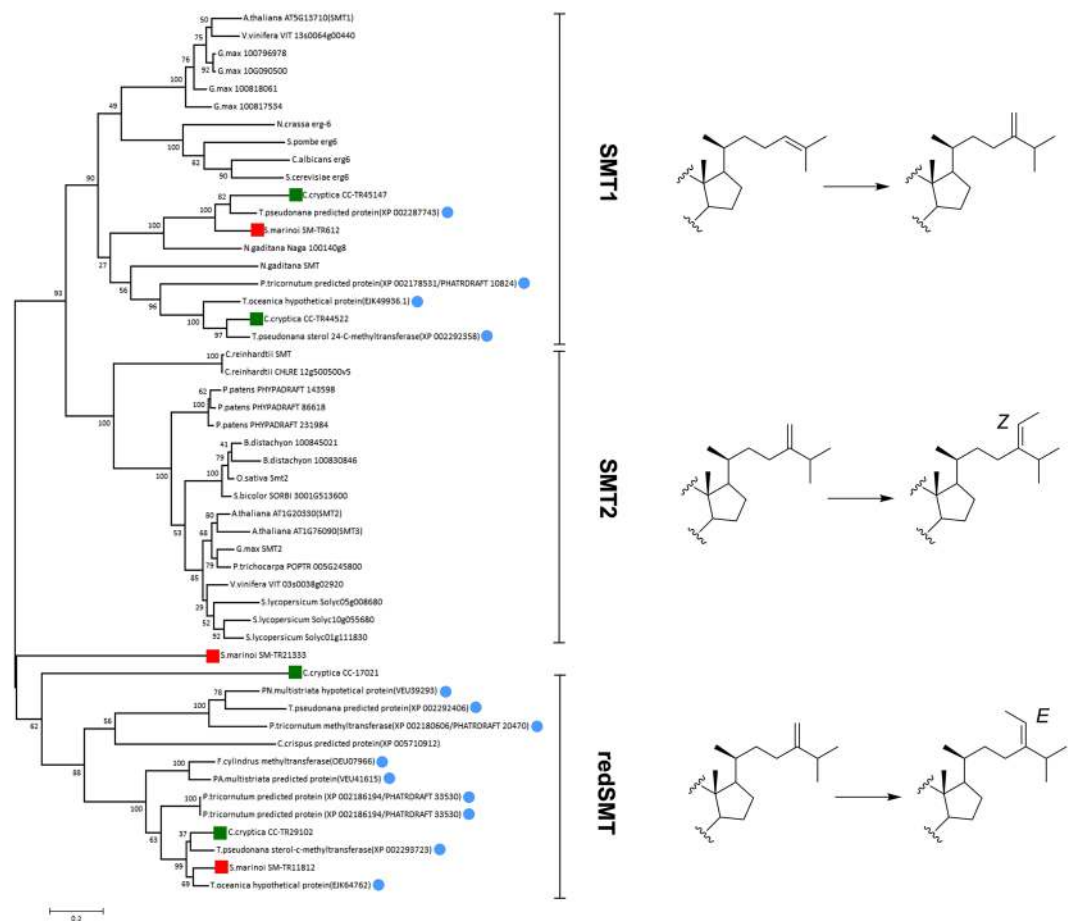


Figure 4. Un-rooted phylogenetic tree of sterol methyl transferase (SMT) of diatoms (Red square = *S. marinoi*; Green square = *C. cryptica*; Light blue spot = other diatom), plants and algae. Reactions catalyzed by SMT are shown next to each sub-family (SMT1, SMT2 and red-SMT). Alignment used maximum likelihood method. The bootstrapping test (replicate = 100) is reported on each node. SMT1 = sterol methyl transferase 1; SMT2 = sterol methyl transferase 2; redSMT = sterol methyl transferase of diatoms and red algae.

showed sequences homologous to DWARF7 and DWARF5 that grouped with putative proteins of other diatoms (similarities above 40%) and a query coverage ranged from 50% to 86% to proteins of *A. thaliana*. In particular, At-DWARF7 and At-DWARF5 are homologous to the putative proteins Sm-TR28093 and Sm-TR2636 of *S. marinoi*, as well as Cc-TR48724 and Cc-TR24068 of *C. cryptica*. In agreement with these results, un-rooted phylogenetic trees constructed using DWARF7/C5SR/ERG3 and DWARF5/7/DHCR from plants, animals, fungi and yeasts showed clusterization of diatom sequences in the plant branch (Supplementary Figs. S7 and S8).

Transcripts did not identify significant match with plant DWARF1 but indicated the presence of two putative sequences (Sm-TR213 and Cc-TR29411) with high similarity with Delta-(24)-sterol reductase (ERG4) from *Saccharomyces cerevisiae* and other fungi and yeasts (Fig. 5). Comparison versus different diatoms (*T. oceanica*, *T. pseudonana*, *P. tricornutum* and *F. cylindrus*) identified orthologous sequences of Sm-TR213 and Cc-TR29411, thus suggesting the wide occurrence and function of these genes in the lineage. Interestingly, the divergence of the diatom transcripts from the plant pathway for the reduction of the alkenyl group at C-24 is consistent with the β -stereochemistry of dihydrobrassicasterol and clionasterol in comparison to the α -orientation of campesterol and β -sitosterol in plants (Fig. 5). We have not been able to establish an unambiguous identification Δ_{22} sterols (trace components) in our study even if the molecular analysis of *C. cryptica* showed a transcript related to fungal 22-sterol reductase (ERG5) (Table 2).

Biosynthesis of cholesterol and desmosterol. In plants, cholesterol and related sterols are little common. Only recently, a biosynthetic pathway of cholesterol has been proposed on the basis of functional assays including gene silencing and reactivity of recombinant tomato proteins⁷. According to this study, plant cholesterologenesis follows the Kandutsch-Russell pathways based on enzymes derived from the phytosterol biosynthesis after methylation of cycloartenol by sterol side chain reductase (SSR2)⁵⁶. In addition to phytosterols, *S. marinoi* produces desmosterol and cholesterol, while these products are absent in *C. cryptica*. Desmosterol is directly converted to cholesterol by 24-DHCR in animals but it has been seldom reported in plants. For these reasons, *S. marinoi* and *C. cryptica* can serve as useful models to study the cholesterol biosynthesis in diatoms and, in

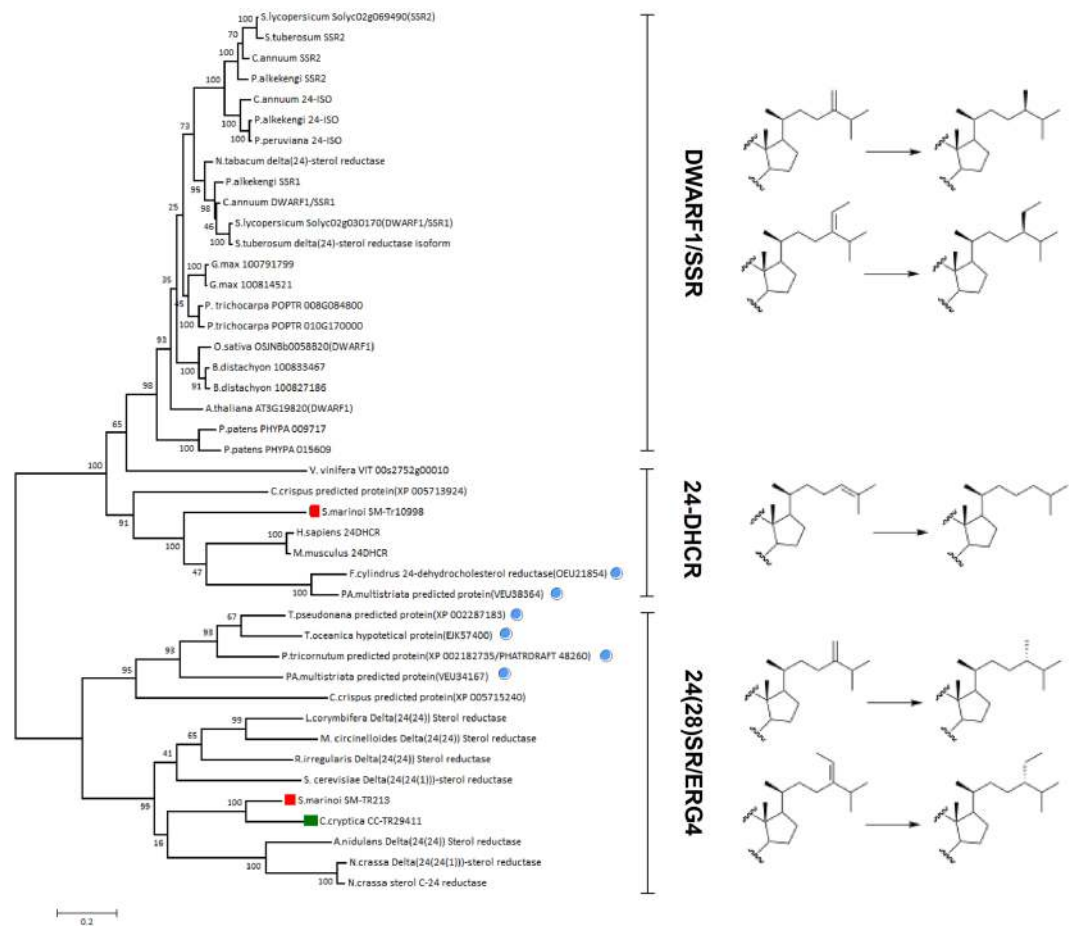


Figure 5. Un-rooted phylogenetic tree of 24-sterol reductases in diatoms, plants, algae, fungi and animals. Reactions catalyzed by each sub-family are shown next to the enzymatic groups. Alignment used maximum likelihood method. The bootstrapping test (replicate = 100) is reported on each node. Red square = *S. marinoi*; Green square = *C. cryptica*; Light blue spot = other diatom. DAWRF1 = plant 24-sterol reductase; SSR = plant sterol side chain reductase; 24-DHCR = animal 24-dehydrocholesterol reductase; 24(28)SR/ERG4 = fungal Delta(24(24)1)-sterol reductase.

analogy to plants, to investigate whether enzymes of the phytosterol pathway are possibly utilized for cholesterologenesis in this microalgal lineage.

Differential analysis of the two diatoms identified a sequence of a putative 24-DHCR (Sm-TR10998) that was present only in *S. marinoi*. Interestingly, Sm-TR10998 showed high similarity with the human 24-DHCR (Query coverage 94.7%; Positives 67.4%) and, at less extent, with SSR2 of tomato (*Solanum lycopersicum*, Query coverage 92%; Positives 53%) (Supplementary Fig. S7). No similar sequences were found in *C. cryptica*. As showed in the un-rooted phylogenetic tree of Fig. 5, Sm-TR10998 clustered in the 24-DHCR branch together with homologous sequences of the diatoms *F. cylindrus* (NCBI accession number: OEU21854) and *Pseudo-nitzschia multistriata* (NCBI accession number: VEU38364) that are both able to biosynthesize cholesterol³². In agreement with the evolutionary origin of diatoms, an ortholog of Sm-TR10998 was found in the red alga *C. crispus* (NCBI accession number: XP_005713924) that biosynthesizes sterol mixtures containing up to 94% of cholesterol⁵⁷.

Discussion

C. cryptica and *S. marinoi* are cosmopolitan diatom species that contribute significantly to phytoplankton blooms in temperate oceans. GC-MS profiling revealed a common mixture of sterols that is dominated by phytosterols with presence of cholesterol and desmosterol only in the latter species. These data are consistent with the literature that indicates major presence of C₂₇ sterols in radial centrics like *S. marinoi*³². As depicted in the biosynthetic proposal of Fig. 6, the route to the sterol skeleton of 24-methylene cholesterol, dihydrobrassicasterol, fucosterol and clionasterol from IPP and DMAPP can have high similarity with plants. 24-Methylene cholesterol is the most common sterol of diatoms³² and a key intermediate of the biosynthesis of phytosterols. Despite this, levels of 24-methylene cholesterol in plants are usually low while this metabolite together with fucosterol and clionasterol (also known as γ -sitosterol) has been often reported in algae and lower invertebrates^{6,9}.

In the literature, there are contrasting data about the biosynthesis of the isoprene units in diatoms. In fact, while *Rhizosolenia setigera*, *P. tricornutum*, and *Nitzschia ovalis* are reported to use mevalonate (MVA), *Haslea ostrearia* shows evidence for methylerythritol (MEP) as precursor of sterols^{58,59}. By feeding studies with labelled glucose,

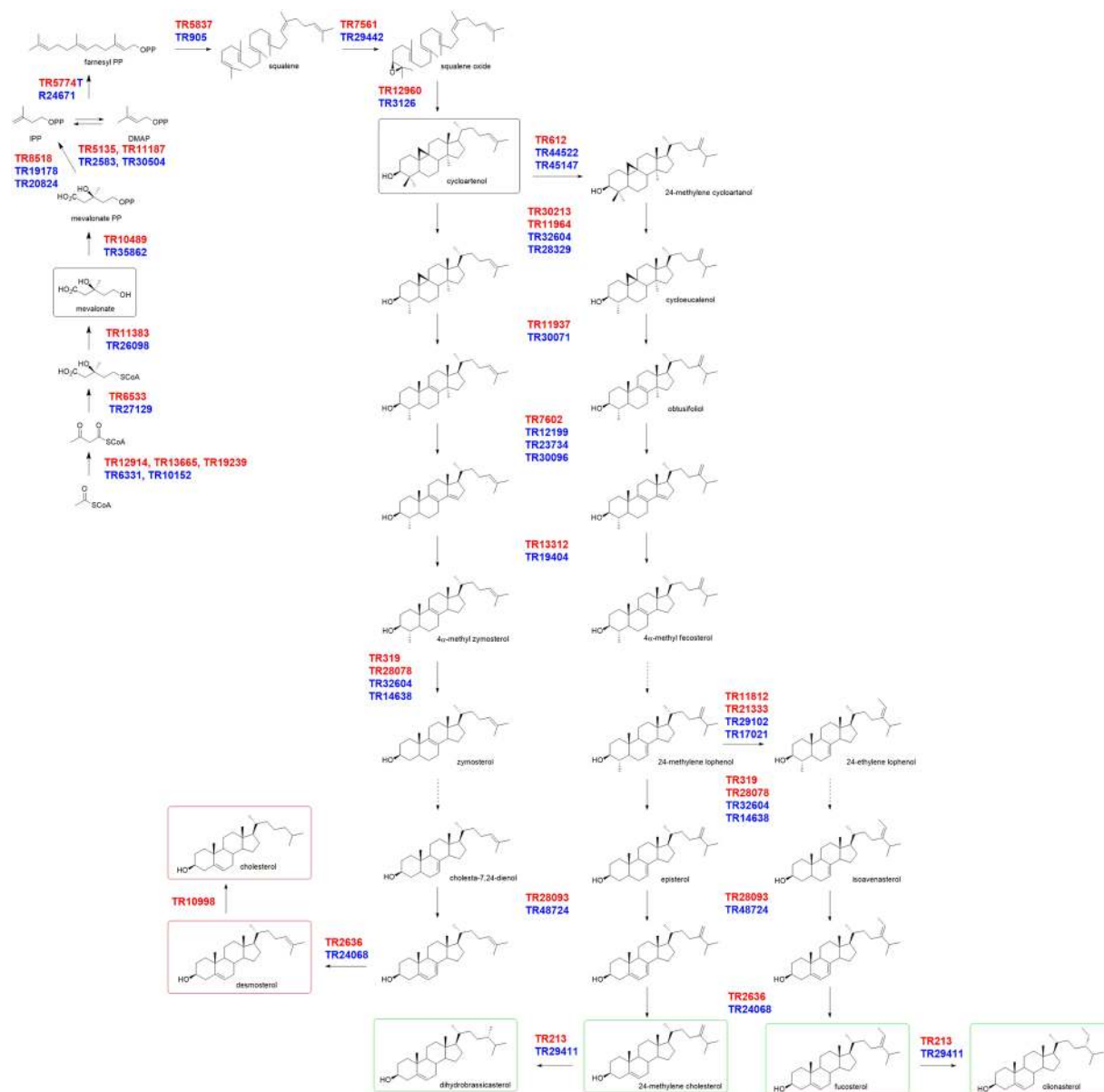


Figure 6. Proposed biosynthetic pathway of sterols in diatoms. Transcript sequences (Red = *S. marinoi*; Blue = *C. cryptica*) are in agreement with Table 1 and deposited under accession code SRP108217 and PRJNA561910. Red square = mammalian sterols; Green square = phytosterols.

we proved origin of the sterol skeleton only by the MVA pathway in *C. cryptica* thus corroborating the results on desmosterol of *R. setigera* by incorporation of $1\text{-}^{13}\text{C}$ acetate⁵⁸. The results of the labelling experiments found a match in the transcriptome sequences which showed a complete and expressed MVA pathway in *C. cryptica* (Fig. 2). Significantly, orthologs of these genes were well represented in the transcriptome of *S. marinoi*, as well as were found in the genome of *T. oceanica*, *T. pseudonana*, *P. tricorutum* and *F. cylindrus* (Supplementary Table S2).

After conversion of farnesyl pyrophosphate to squalene, the key step of the process is cyclization to cycloartenol by cycloartenol synthase (CAS) via squalene oxide. *C. cryptica* and *S. marinoi* showed a single copy of a plant-like CAS and the presence of a terbinafine-insensitive SQE, named AltSQE, that has been recently proposed to catalyze synthesis of squalene oxide in *S. marinoi* and other diatoms⁵⁰. The wide occurrence of 24-methylene cholesterol in diatoms³² is consistent with the suggestion that, like in plants, methylation of the sterol side chain occurs very early in the biosynthetic process. In agreement with this view, the transcriptome analyses indicated the presence of different orthologs of plant SMT1, which are presumably responsible for C-24 methylation of cycloartenol to oryzanol^{33,60}. On the other hand, our results showed major differences between plants and diatoms for SMT2 that presides methylation of 24-methylenelophenol to yield C_{28} phytosterols. It is noteworthy that this divergence corresponds to the difference in the stereochemistry of fucosterol and isofucosterol (Fig. 4) that are the end products of these enzymes in diatoms and plants respectively. SMTs embrace a class of proteins that have inherited a high degree of sequence similarity from a common ancestor. Recently, Nes and

coworkers have discussed the differences of SMTs in relation to substrate- and phylum-specificity in green algae⁶¹. Our comparative analysis suggests that diatoms and the red alga *C. crispus* contain SMT sequences that may have evolved independently from plants, green algae and fungi. It is possible that diatoms have acquired these genes by the red alga that was engulfed during the secondary endosymbiotic event that is common to all stramenophiles, as well as we cannot exclude that these “red SMTs” can accept substrates other than 24-methylenelophenol. Further studies are necessary to prove the catalytic function of these proteins and the substrate specificity. However, the absence of scrambling of methyl groups in fucosterol and clionasterol indicates a high affinity of these enzymes for the substrates.

The high and specific ¹³C-incorporation of phytosterols of *C. cryptica*, as well as the absence of isotopic scrambling, features the operation of a unique and straightforward route similar to brassinolide pathway. In support to this view, sequence alignment of the putative enzymes of *S. marinoi* and *C. cryptica* revealed high homology with the corresponding proteins of plants. We found only major divergences of the transcript sequences (Sm-TR213 and Cc-TR29411) related to the reduction steps leading from 24-methylene cholesterol and fucosterol to dihydrobrassicasterol and clionasterol, respectively. These latter compounds are common algal metabolites and, despite the structural similarities with campesterol and β -sitosterol of plants, are produced by stereochemical reduction that envisages a mechanism sharply different from that described for phytosterols⁶. Accordingly, Sm-TR213 and Cc-TR29411 show high homology with fungal ERG4 that catalyzes hydrogen attack on 24-*si* face of the Δ 24(28) double bond, which is consistent with the β -stereochemistry of the methyl and ethyl groups of dihydrobrassicasterol and clionasterol in *S. marinoi* and *C. cryptica* (Fig. 5). On the basis of the phylogenetic results, we suggest that reduction of 24-methylene cholesterol and fucosterol can be carried out by a single fungal-like Δ 24(28)-sterol reductase. A similar hypothesis has been also put forward for the reduction of 24-methylene cholesterol and isofucosterol by a single 24-sterol reductase in plants⁶. Differently from the report on *P. tricornutum*³³, *S. marinoi* and *C. cryptica* do not synthesise Δ ₂₂ sterols (e.g. diatomsterol). It is possible that the Δ ₂₂ unsaturation can be introduced in the late steps of the pathway by a 22-sterol desaturase similar to ERG5 that was detected in the transcripts of *C. cryptica*. However, the absence of typical fungal sterols (e.g. ergosterol) together with the lack of other molecular proofs argue against a fungus-like process for the sterol biosynthesis in these diatoms.

Our data clearly indicated no lanosterol synthase genes in *S. marinoi* and *C. cryptica* thus highlighting origin of cholesterol and desmosterol from cycloartenol. This process has been recently proved in tomato that can synthesize C₂₇, C₂₈ and C₂₉ sterols by a promiscuous pathway⁷. Interestingly, despite the phylogenetic distances between tomato and diatoms, a certain degree of similarities can be identified in the sequences of key enzymes of sterol biosynthesis (Supporting Material). In tomato the divergence between cholesterol and phytosterols stems from the conversion of cycloartenol to cycloartanol by sterol side chain reductase enzyme (SSR2)⁵⁶. However, since this reaction occurs very early in the sequence of enzymatic steps leading to cholesterol, the pathway is not compatible with the synthesis of desmosterol. On the other hand, desmosterol can accumulate as terminal product in diatoms and red algae (e.g. *Porphyridium purpureum*). Desmosterol is mostly known as the immediate precursor of cholesterol in the Block pathway in animals including humans, as well as it is a common intermediate in the conversion of C₂₈ and C₂₉ plant sterols to cholesterol by the tobacco hornworm and other insects^{11,62}. Transcriptome of *S. marinoi* indicated a sequence (Sm-TR10998) with a predicted function related to 24-DHCR from *Homo sapiens*. The enzyme reduces desmosterol to cholesterol with the highest affinity for sterol skeleton lacking methyl groups at C-4, like desmosterol and 7,24-cholestandiol⁶³. Differently from the proposal role of SSR2 in tomato, this specificity and the finding of desmosterol in *S. marinoi* suggest occurrence of 24-reduction in the late steps of the pathway (Fig. 6). Orthologs of Sm-TR10998 occur in all known genomes of diatoms producing cholesterol³², thus highlighting a general role of this enzyme in the lineage.

In conclusion, we suggest that the sterol pathway of *S. marinoi* and *C. cryptica* preserves the general architecture recently described in Solanaceae plants with cyclization of squalene oxide to cycloartenol rather than lanosterol that is characteristic of the fungal genealogy⁷. Major differences seem to be related to the putative enzymes that preside modification, namely methylation and reduction, of the alkyl side chain by enzymes encoded by genes probably derived from algae, fungi and animals. The acquisition of these genes explains well the composition and the stereochemistry of sterols in diatoms, as well as confirms the metabolic plasticity of these microalgae that are evolved through multiple events of symbiosis and horizontal gene transfer. Further studies are in progress to correlate the molecular data with biochemical tests.

Methods

General. HPLC analyses have been performed on a Jasco system (PU-2089 Plus-Quaternary gradient pump equipped with a Jasco MD-2018 Plus photodiode array detector). Cell pellet were lyophilized by a Savan Micro Modulyo freeze dryer (Thermo Scientific, Austin, TX, USA). GC-MS analysis was carried out by an Ion-Trap Polaris Q MS instrument (Thermo) coupled to a Focus gas chromatograph (Thermo). ¹H and ¹³C NMR spectra were recorded on Bruker DRX 600 spectrometer equipped with an inverse TCI CryoProbe. All chemicals and analytical grade solvents were purchased from Sigma Aldrich. Molecular and bioinformatics analyses were performed by Genomix 4Life s.r.l. (Baronissi, Salerno, Italy).

Microalgae cultures. Stock cultures of the diatom *Skeletonema marinoi* (CCMP 2092), *Skeletonema costatum*, *Pseudonitzschia arenysensis* (B758), *Thalassiosira weissflogii* (P09), *Cyclotella cryptica* (CCMP 331) and *Phaeodactylum tricornutum* were purchased from Bigelow Laboratories. Strains were maintained in f/2 medium⁶⁴ in a growth chamber at 20 ± 2 °C under 14:10 light:dark cycle with a photon flux density of 100 μ mol quanta m⁻² sec⁻¹. Cultures were harvested by centrifugation at 3750 rpm for 10 minutes at 12 °C using a swing-out rotor. Pellets were frozen in liquid nitrogen and stored at -80 °C until analysis.

Labelling studies. For biosynthetic experiments, *C. cryptica* was incubated in 1 L-sterile culture flasks (TPP25–300 cm²) in prefiltered sterile (0.22 µm) f/2 medium supplemented with antibiotics according to published protocols^{65,66}. Under these conditions, cells were grown heterotrophically on 2 g L⁻¹ of glucose or 1-¹³C-glucose in the dark at 20 ± 2 °C under constant gentle agitation. Cultures were harvested and stored at –80 °C until analysis, as described above.

Sterol extraction and purification. After lyophilization, microalgal pellets were extracted by the modified Folch method. The oily residues were dried, reconstituted in diethyl ether and methylated by diazomethane (0.4 mL of a saturated ether solution per 10 mg extract). After removing the organic solvent under N₂, sterols were purified on silica gel column (100 mg silica gel/ mg fraction). Homogeneous fractions of sterol mixtures were achieved in petroleum ether/diethyl ether 85:15 (v/v). Product elution was monitored by SiO₂-TLC on 0.2-mm aluminum-coated sheets (Merck, Germany) developed with petroleum ether/ethyl ether (1:1; v/v). Final purification of sterols was carried out by HPLC on a reverse phase column (C18-Luna, Phenomenex, 5 µm 100 Å 250 × 10 mm) by methanol/acetonitrile/water/iPrOH/acetone (33:33:4:5:25; v-v) according to our published protocols^{67,68}. Elution was accomplished by a flow of 1 mL min⁻¹ and monitored by UV absorbance at 210 nm.

Sterol analysis. For GC-MS analysis, sterols were acetylated by Ac₂O (100 µL/ 0.5 mg sample) in dry pyridine (500 µL/0.5 mg sample) under magnetic stirring overnight at room temperature. The excess of organic solvents was removed under N₂ stream. GC-MS runs of acetylated sterols were carried out in EI mode (70 eV) by isocratic elution at 300 °C for 30 min, with a 5% diphenyl-polysiloxane capillary column (OV-5 column) (15 m x 0.32 mm ID, 0.10 µm) and helium as gas carrier. Samples dissolved in n-hexane were directly injected (2 µL) in split mode (1:10), with a blink window of 3 minutes, inlet temperature of 300 °C, transfer line set at 310 °C and ion source temperature of 300 °C. Sterols were identified according to literature and for comparison with commercial standards^{35–40,69–72}. For NMR experiments (¹H, ¹³C, COSY, TOCSY, HSQC), natural and standard sterols were dissolved in 700 µL CDCl₃ and transferred to 5-mm NMR tube. For biosynthetic studies, ¹³C NMR spectra were acquired with a relaxation delay of 6 seconds to reduce the effect of longitudinal relaxation.

Transcriptome analysis and bioinformatics. RNA was extracted by Trizol method and RNA sequencing experiment was performed on three biological replicates for diatoms. Before use, RNA concentration in each sample was assayed with a ND-1000 spectrophotometer (NanoDrop) and quality was assessed with the Agilent 2100 Bioanalyzer with Agilent RNA 6000 nano kit (Agilent Technologies, Santa Clara, CA, USA). Indexed libraries were prepared from 4 µg/ea purified RNA with TruSeq Stranded mRNA Sample Prep Kit (Illumina) according to the manufacturer's instructions. Libraries were quantified using the Agilent 2100 Bioanalyzer (Agilent Technologies) and pooled such that each index-tagged sample was present in equimolar amounts, with a final concentration of the pooled samples of 2 nM. The pooled samples were subject to cluster generation and sequencing using an Illumina HiSeq. 2500 System (Illumina) in a 2 × 100 paired-end format at a final concentration of 8 pmol. Transcriptome was assembled using Trinity⁷³. The high-quality reads were selected as input to perform the transcriptome assembly and the resulting sequences were translated into proteins by using the longest complete ORF. The sequences of the assembled transcripts were translated into proteins with Transdecoder and the software Blast2GO was used to associate a function to identified transcripts. To compute abundance estimation, the high-quality reads were aligned to the Trinity transcriptome using bowtie2. Then, either RSEM is executed to estimate expression values based on the resulting alignments, generating the raw counts. In order to define the set of expressed genes, raw read counts were normalized using the TMM method (Trimmed mean). Phylogenetic analyses were refined by MEGA 6.0 software⁷⁴ and alignments were obtained by the MUSCLE algorithm. Phylogenetic trees were built by the maximum likelihood method with the LG + G substitution model (for CAS, 7DHCR/DWF5, SMT and 24-DHCR/DWF1) and LG + G + I (for C5SD/DWF7). Test of phylogeny was performed by the bootstrap method with a number of bootstrap replication equal to 100.

Data availability

The raw sequencing data from this study have been submitted to the NCBI SRA database (<http://www.ncbi.nlm.nih.gov>). The RNA sequencing data of *S. marinoi* and *C. cryptica* have been deposited under accession code SRP108217 and PRJNA561910. The sterol related sequences of interest are reported in the Supplementary Material. All other supporting data from this study are available within the article and its Supplementary Information Files, or from the corresponding authors upon request.

Received: 30 September 2019; Accepted: 10 February 2020;

Published online: 06 March 2020

References

- Schaller, H. New aspects of sterol biosynthesis in growth and development of higher plants. *Plant Physiology and Biochemistry* **42**, 465–476 (2004).
- Boutté, Y. & Grebe, M. Cellular processes relying on sterol function in plants. *Curr. Opin. Plant Biol.* **12**, 705–713 (2009).
- Vriet, C., Lemmens, K., Vandepoele, K., Reuzeau, C. & Russinova, E. Evolutionary trails of plant steroid genes. *Trends Plant Sci* **20**, 301–308 (2015).
- Lu, Y. *et al.* Regulation of the cholesterol biosynthetic pathway and its integration with fatty acid biosynthesis in the oleaginous microalga *Nannochloropsis oceanica*. *Biotechnol. Biofuels* **7**, 81 (2014).
- Rohmer, M. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nat. Prod. Rep.* **16**, 565–574 (1999).
- Nes, W. D. Biosynthesis of cholesterol and other sterols. *Chem. Rev.* **111**, 6423–6451 (2011).
- Sonawane, P. D. *et al.* Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism. *Nat. Plants* **3**, 16205 (2017).
- Giner, J. L. & Djerassi, C. Evidence for sterol side-chain dealkylation in *Chlamydomonas reinhardtii*. *Phytochemistry* **31**, 3865–3867 (1992).

9. Giner, J.-L. Biosynthesis of sterol side chains. *Chem. Rev.* **93**, 1735–1752 (1993).
10. Tomazic, M. L., Najle, S. R., Nusblat, A. D., Uttaro, A. D. & Nudel, C. B. A novel sterol desaturase-like protein promoting dealkylation of phytosterols in *Tetrahymena thermophila*. *Eukaryot. Cell* **10**, 423–434 (2011).
11. Ikekawa, N., Morisaki, M. & Fujimoto, Y. Sterol Metabolism in Insects: Dealkylation of Phytosterol to Cholesterol. *Acc. Chem. Res.* **26**, 139–146 (1993).
12. Garvey, M., Moriceau, B. & Passow, U. Applicability of the FDA assay to determine the viability of marine phytoplankton under different environmental conditions. *Mar. Ecol. Prog. Ser.* **352**, 17–26 (2007).
13. Armbrust, E. V. The life of diatoms in the world's oceans. *Nature* **459**, 185–192 (2009).
14. Falkowski, P. G. The role of phytoplankton photosynthesis in global biogeochemical cycles. *Photosynthesis Research* **39**, 235–258 (1994).
15. Verity, P. G. & Smetacek, V. Organism life cycles, predation, and the structure of marine pelagic ecosystems. *Mar. Ecol. Prog. Ser.* **130**, 277–293 (1996).
16. Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A. & Quéguiner, B. Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochem. Cycles* **9**, 359–372 (1995).
17. Allen, J. T. *et al.* Diatom carbon export enhanced by silicate upwelling in the northeast Atlantic. *Nature* **437**, 728–732 (2005).
18. Stonik, V. S. & Stonik, I. Low-molecular-weight metabolites from diatoms: structures, biological roles and biosynthesis. *Mar. Drugs* **13**, 3672–3709 (2015).
19. Martin-Jézéquel, V. *et al.* Effects of organic and inorganic nitrogen on the growth and production of domoic acid by *Pseudo-nitzschia multiseriata* and *P. australis* (Bacillariophyceae) in culture. *Mar. Drugs* **13**, 7067–7086 (2015).
20. Brunson, J. K. *et al.* Biosynthesis of the neurotoxin domoic acid in a bloom-forming diatom. *Science* **361**, 1356–1358 (2018).
21. d'Ippolito, G. *et al.* Lipoxygenases and Lipoxygenase Products in Marine Diatoms. In *Methods in Enzymology* 69–100, <https://doi.org/10.1016/bs.mie.2018.02.021> (2018).
22. Lamari, N. *et al.* Specificity of lipoxygenase pathways supports species delineation in the marine diatom genus *Pseudo-nitzschia*. *PLoS One* **8**, e73281 (2013).
23. Cutignano, A. *et al.* Lipoxygenase products in marine diatoms: A concise analytical method to explore the functional potential of oxylipins. *J. Phycol.* **47**, 233–243 (2011).
24. d'Ippolito, G. *et al.* Biosynthetic intermediates and stereochemical aspects of aldehyde biosynthesis in the marine diatom *Thalassiosira rotula*. *Phytochemistry* **67**, 314–322 (2006).
25. Perfeito, C., Ambrósio, M., Santos, R., Afonso, C. & Abranches, R. Increasing fucoxanthin production in *Phaeodactylum tricornerutum* using genetic engineering and optimization of culture conditions. *Front. Mar. Sci.* **5**, 3–5 (2018).
26. Dambek, M. *et al.* Biosynthesis of fucoxanthin and diadinoxanthin and function of initial pathway genes in *Phaeodactylum tricornerutum*. *J. Exp. Bot.* **63**, 5607–5612 (2012).
27. Hildebrand, M. Diatoms, biomineralization processes, and genomics. *Chem. Rev.* **108**, 4855–4874 (2008).
28. Ehrlich, H. & Witkowski, A. *Biomineralization in Diatoms: The Organic Templates*. In **6**, 39–58 (2015).
29. Brunner, E. *et al.* Chitin-based organic networks: An integral part of cell wall biosilica in the diatom *Thalassiosira pseudonana*. *Angew. Chemie - Int. Ed.* **48**, 9724–9727 (2009).
30. Ehrlich, H. *et al.* Multiphase Biomineralization: Enigmatic Invasive Siliceous Diatoms Produce Crystalline Calcite. *Adv. Funct. Mater.* **26**, 2503–2510 (2016).
31. Gallo, C., d'Ippolito, G., Nuzzo, G., Sardo, A. & Fontana, A. Autoinhibitory sterol sulfates mediate programmed cell death in a bloom-forming marine diatom. *Nat. Commun.* **8** (2017).
32. Rampen, S. W., Abbas, B., Schouten, S. & Sinninghe Damste, J. S. A comprehensive study of sterols in marine diatoms (Bacillariophyta): Implications for their use as tracers for diatom productivity. *Limnol. Oceanogr.* **55**, 91–105 (2010).
33. Fabris, M. *et al.* Tracking the sterol biosynthesis pathway of the diatom *Phaeodactylum tricornerutum*. *New Phytol.* **204**, 521–535 (2014).
34. Dupont, S. *et al.* Ergosterol biosynthesis: a fungal pathway for life on land? *Evolution (N. Y.)* **66**, 2961–2968 (2012).
35. Wright, J. L. C., McInnes, A. G., Shimizu, S. & Khalil, D. I. A. N. D. W. Identification of C-24 alkyl epimers of marine sterols by I3C nuclear magnetic resonance spectroscopy. *Can. J. Chem.* **56**, 1898–1903 (1978).
36. Popjak, G., Edmond, J., Anet, F. A. L. & Easton, N. R. Carbon-13 NMR studies on cholesterol biosynthesized from [¹³C]mevalonates. *J. Am. Chem. Soc.* **99**, 931–935 (1977).
37. Iida, T., Tamura, T. & Matsumoto, T. *Proton nuclear magnetic resonance identification and discrimination of side chain isomers of phytosterols using a lanthanide shift reagent* **21**, 326–338 (1980).
38. McInnes, A. G., Walter, J. A. & Wright, J. L. C. C NMR spectra of delta 24(28) phytosterols. *Org. Magn. Reson.* **13**, 302–303 (1980).
39. Hwang, S. H., Jang, J. M. & Lim, S. S. Isolation of fucosterol from *Pelvetia siliquosa* by high-speed countercurrent chromatography. *Fish. Aquat. Sci.* **15**, 191–195 (2012).
40. Dzeha, T., Jaspars, M. & Tabudravu, J. Clionasterol, a Triterpenoid from the Kenyan Marine Green Macroalga *Halimeda macroloba*. *West. Indian Ocean. J. Mar. Sci.* **2**, 157–161 (2004).
41. Véron, B., Billard, C., Dauguet, J.-C. & Hartmann, M.-A. Sterol composition of *Phaeodactylum tricornerutum* as influenced by growth temperature and light spectral quality. *Lipids* **31**, 989–994 (1996).
42. Gladu, P. K., Patterson, G. W., Wikfors, G. H., Chitwood, D. J. & Lusby, W. R. Sterols of some diatoms. *Phytochemistry* **30**, 2301–2303 (1991).
43. Barrett, S. M., Volkman, J. K., Dunstan, G. A. & LeRoi, J. M. Sterols of 14 Species of Marine Diatoms (Bacillariophyta). *J. Phycol.* **31**, 360–369 (1995).
44. Volkman, J. K. & Hallegraeff, G. M. Lipids in marine diatoms of the genus *Thalassiosira*: Predominance of 24-methylenecholesterol. *Phytochemistry* **27**, 1389–1394 (1988).
45. Goad, L. J., Lenton, J. R., Knapp, F. F. & Goodwin, T. W. Phytosterol side chain biosynthesis. *Lipids* **9**, 582–595 (1974).
46. Nes, W. R., Krevitz, K. & Behzadan, S. Configuration at C-24 of 24-Methyl and 24-Ethylcholesterol in Tracheophyte. *Lipids* **11**, 118–126 (1976).
47. Hellebust, J. A. Glucose Uptake by *Cyclotella cryptica*: Dark Induction and Light Inactivation of Transport System 2. *J. Phycol.* **7**, 345–349 (1971).
48. White, A. W. Growth of Two Facultatively Heterotrophic Marine Centric Diatoms 2. *J. Phycol.* **10**, 292–300 (1974).
49. Rohmer, M. *et al.* Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. *Biochem. J.* **295**, 8 (1993).
50. Pollier, J. *et al.* A widespread alternative squalene epoxidase participates in eukaryote steroid biosynthesis. *Nat. Microbiol.* **4**, 226–233 (2019).
51. Kolesnikova, M. D., Xiong, Q., Lodeiro, S., Hua, L. & Matsuda, S. P. T. Lanosterol biosynthesis in plants. *Arch. Biochem. Biophys.* **447**, 87–95 (2006).
52. Darnet, S. & Schaller, H. Metabolism and Biological Activities of 4-Methyl-Sterols. *Molecules* **24**, 451 (2019).
53. Mo, C. & Bard, M. Erg28p is a key protein in the yeast sterol biosynthetic enzyme complex. *J. Lipid Res.* **46**, 1991–1998 (2005).
54. Mialoundama, A. S. *et al.* *Arabidopsis* ERG28 Tethers the Sterol C4-Demethylation Complex to Prevent Accumulation of a Biosynthetic Intermediate That Interferes with Polar Auxin Transport. *Plant Cell* **25**, 4879–4893 (2013).

55. Li, W. *et al.* A novel N-methyltransferase in *Arabidopsis* appears to feed a conserved pathway for nicotinate detoxification among land plants and is associated with lignin biosynthesis. *Plant Physiol.* **174**, 1492–1504 (2017).
56. Sawai, S. *et al.* Sterol Side Chain Reductase 2 Is a Key Enzyme in the Biosynthesis of Cholesterol, the Common Precursor of Toxic Steroidal Glycoalkaloids in Potato. *Plant Cell* **26**, 3763–3774 (2014).
57. Tasende, M. G. Fatty acid and sterol composition of gametophytes and sporophytes of *Chondrus crispus* (Gigartinales, Rhodophyta). *Sci. Mar.* **64**, 421–426 (2000).
58. Massé, G., Belt, S. T., Rowland, S. J. & Rohmer, M. Isoprenoid biosynthesis in the diatoms *Rhizosolenia setigera* (Brightwell) and *Haslea ostrearia* (Simonsen). *Proc. Natl. Acad. Sci.* **101**, 4413–4418 (2004).
59. Cvejić, J. H. & Rohmer, M. CO₂ as main carbon source for isoprenoid biosynthesis via the mevalonate-independent methylerythritol 4-phosphate route in the marine diatoms *Phaeodactylum tricornutum* and *Nitzschia ovalis*. *Phytochemistry* **53**, 21–28 (2000).
60. Desmond, E. & Gribaldo, S. Phylogenomics of sterol synthesis: insights into the origin, evolution, and diversity of a key eukaryotic feature. *Genome Biol. Evol.* **1**, 364–81 (2009).
61. Nes, W. Sterol methyl transferase: enzymology and inhibition. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1529**, 63–88 (2000).
62. Ciuffo, L. F., Murray, P. A., Thompson, A., Rigden, D. J. & Rees, H. H. Characterisation of a desmosterol reductase involved in phytosterol dealkylation in the silkworm, *Bombyx mori*. *PLoS One* **6**, (2011).
63. Pedretti, A., Bocci, E., Maggi, R. & Vistoli, G. Homology modelling of human DHCR24 (seladin-1) and analysis of its binding properties through molecular docking and dynamics simulations. *Steroids* **73**, 708–719 (2008).
64. Guillard, R. R. L. & Ryther, J. H. Studies of Marine Planktonic Diatoms: I. *Cyclotella nana* (Hustedt), and *Detonula confervacea* (Cleve) Gran. *Can. J. Microbiol.* **8**, 229–239 (1962).
65. Kobayashi, K., Kobiyama, A., Kotaki, Y. & Kodama, M. Possible occurrence of intracellular bacteria in *Pseudonitzschia multiseries*, a causative diatom of amnesic shellfish poisoning. *Fish. Sci.* **69**, 974–978 (2003).
66. Bruckner, C. G. & Kroth, P. G. Protocols for the Removal of Bacteria from Freshwater Benthic Diatom Cultures 1. *J. Phycol.* **45**, 981–986 (2009).
67. Fontana, A. *et al.* High-Resolution NMR Detection of Cholesterol Oxides in Spray-Dried Egg Yolk. *J. Food Sci.* **57**, 869–872 (1992).
68. Fontana, A., Antoniazzi, F., Ciavatta, M. L., Trivellone, E. & Cimino, G. ¹H-NMR Study of Cholesterol Autooxidation in Egg Powder and Cookies Exposed to Adverse Storage. *J. Food Sci.* **58**, 1286–1290 (1993).
69. Seo, S. *et al.* Biosynthesis of the 24-methylcholesterols dihydrobrassicasterol and campesterol in cultured cells of *Amsonia elliptica*: incorporation of [1,2-¹³C₂] acetate and [2-¹³C, ²H₃] acetate. *J. Chem. Soc. Perkin Trans 1*, 569, <https://doi.org/10.1039/p19920000569> (1992).
70. Seo, S. *et al.* Biosynthesis of isofucoesterol from [2-¹³C ²H₃] acetate and [1,2-¹³C₂] acetate in tissue cultures of *Physalis peruviana* — the stereochemistry of the hydride shift from C-24 to C-25. *J. Chem. Soc. Perkin Trans. 1*, 105–109, <https://doi.org/10.1039/P19900000105> (1990).
71. Uomori, A. *et al.* Biosynthesis of campesterol and dihydrobrassicasterol in cultured cells of *Amsonia elliptica*. *Phytochemistry* **31**, 1569–1572 (1992).
72. Fujimoto, Y., Sato, N., Okuzumi, T., Yamada, J. & Morisaki, M. Stereochemistry of the reduction of 24-ethyl-desmosterol to sitosterol in tissue cultures of *Oryza sativa*. *Bioorganic Med. Chem. Lett.* **8**, 205–208 (1998).
73. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
74. Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).

Acknowledgements

This study was supported by the project “Integrated Exploitation of Algal Biomass for Production of Energy” (PON01_02740) funded by the National Operational Program for Research and Competitiveness 2007–2013. The authors acknowledge the technical support of Mr Lucio Caso and Dr. Salvatore Morra. AF thanks the Doctorate School of Biology of the University of Naples “Federico II”.

Author contributions

C.G. carried out the experiments and supported bioinformatic analysis; S.L. performed bioinformatic analysis; G.N. and E.M. carried out chemical analysis, structure elucidation, sterol semisynthesis and sterol derivatization; A.S. gave support to development of culture conditions and diatom cultivation; G.d.I. conceived the original idea and supervised the experiments; A.F. designed the study, planned and supervised the experiments, analyzed the data; A.F. wrote the manuscript in consultation with C.G., S.L., G.d.I. and G.N.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-60993-5>.

Correspondence and requests for materials should be addressed to G.d. or A.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020