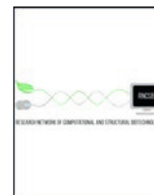
Computational and Structural Biotechnology Journal 20 (2022) 1901–1913

COMPUTATIONAL ANDSTRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj



Insights into the biosynthesis pathway of phenolic compounds in microalgae

Angelo Del Mondo [a,*], Clementina Sansone [a,b], Christophe Brunet [a]

[a] Stazione zoologica Anton Dohrn, sede Molosiglio Marina Acton, via ammiraglio F. Acton, 55., 80133 Napoli, Italy
[b] Institute of Biomolecular Chemistry, CNR, via Campi Flegrei 34, Pozzuoli 80078, Na, Italy

ARTICLE INFO

Article history:
Received 2 February 2022
Received in revised form 4 April 2022
Accepted 14 April 2022
Available online 20 April 2022

Keywords:
Polyphenols
Flavonoids
Microalgae
Blue biotechnology
In silico analysis

ABSTRACT

Among the most relevant bioactive molecules family, phenolic compounds (PCs) are well known in higher plants, while their knowledge in microalgae is still scarce. Microalgae represent a novel and promising source of human health benefit compounds to be involved, for instance, in nutraceutical composition. This study aims to investigate the PCs biosynthetic pathway in the microalgal realm, exploring its potential variability over the microalgal biodiversity axis. A multistep *in silico* analysis was carried out using a selection of core enzymes from the pathway described in land plants. This study explores their presence in ten groups of prokaryotic and eukaryotic microalgae.. Analyses were carried out taking into account a wide selection of algal protein homologs, functional annotation of conserved domains and motifs, and maximum-likelihood tree construction. Results showed that a conserved core of the pathway for PCs biosynthesis is shared horizontally in all microalgae. Conversely, the ability to synthesize some subclasses of phenolics may be restricted to only some microalgal groups (i.e., Chlorophyta) depending on featured enzymes, such as the flavanone naringenin and other related chalcone isomerase dependent compounds.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Phenolic compounds (PCs) represent the largest group of secondary metabolites in plants, ranging from simple aromatic rings to more complex molecules and comprising flavonoids, phenolic acids, tannins, lignans or coumarins. They cover a broad range of functions, and are involved in response against environmental stresses, such as heavy metals (e.g., chromium, aluminium), salinity, temperature increase (33–38 °C), pesticides, drought or UV radiations [1–6]. They also improve nutrient uptake thanks to metallic ions chelation capacity, with the synthesis and exudation of PCs supporting iron acquisition and root-ferric reductase activity [7]. Moreover, it has been observed that plant PCs (e.g., 5-O-caffeoylquinic acid, cinnamic acid, 4-methoxy-cinnamic acid hexoside, K-3-O rutinoside, Q-3-O-rutinoside, Q-3-O-glucoside and Q-3-O-glucuronide) might positively affect mycorrhizal fungi and bacterial root colonizers [8]. PCs do also act as allelochemicals against invading organisms as weeds or pathogens e.g., herbivores, nematodes, phytophagous insects, fungal and bacterial pathogens [9–13].

In plants, PCs derive from the biosynthetic pathway of phenylpropanoids, starting with phenylalanine and tyrosine, these two amino acids providing from the shikimic acid and malonic acid pathways [14]. The upstream pathway of phenylpropanoids starts with the transformation of the aminoacid L-phenylalanine by the phenylalanine ammonia lyase (PAL) enzyme, providing a large panel of simple molecules which are then involved in the formation of a plethora of products [15]. The latter includes the hydroxybenzoates (C6-C1) and hydroxycinnamates (C6-C3), involved in the production of the salicylic acid and in the biosynthesis of coumarins ([15], dark-green route in Fig. 1), gallic acid and its derived tannins in the prephenate pathway ([15], blue route in Fig. 1). The main building blocks for lignin biosynthesis are monolignols i.e., hydroxy-cinnamyl alcohols, coniferyl alcohol, and sinapyl alcohol, with typically minor amounts of p-coumaroyl alcohol ([16], green and red routes in the Fig. 1). The downstream flux in the biosynthesis of lignins is essentially regulated by the enzymes CCR1/2 and CAD in plants [17]. Flavonoids synthesis (orange, light-green and purple routes in Fig. 1) starts from naringenin synthetized thanks to the enzymes chalcone synthase and chalcone isomerase (CHS and CHI, [18]). Flavonoids are the largest class of phenylpropanoids in plants, with a basic structure made of two aromatic rings (one from phenylalanine and the other one from the condensation of

* Corresponding author.
E-mail address: angelo.delmondo@szn.it (A. Del Mondo).

https://doi.org/10.1016/j.csbj.2022.04.019
2001-0370/© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Fig. 1. Major routes in phenylpropanoid/flavonoid pathway: red line: lignins; blue line: gallotannins and elagitannins; dark-green line: coumarins; orange line: stillbenoids; purple line: isoflavonoids; light-green line: flavonoids (flavanols, 3-flavonols); pink line: catechins, anthocyanins. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

three malonic acids) linked by three carbons [18]. The main flavonoid classes are the flavones, flavonols, and isoflavones. Flavones (e.g., apigenin, luteolin) and flavonols (e.g., kaempferol, quercetin) are flavonoids absorbing light in the UV region [19,20]. Anthocyanins are water-soluble pigments belonging to the flavonoids' family. They are normally stored in the vacuole where they react to the presence of acid residues with variations in colour. In plants, the synthesis of proanthocyanidins branches off the flavonoid pathway from either 2,3-cis or -trans flavanols and is initiated by the dihydroflavonol 4-reductase (DFR), producing leucoanthocyanidin [21]. The key enzyme anthocyanidin synthase (ANS) can later form the intermediate anthocyanidins. Anthocyanidin reductase (ANR) or leucoanthocyanidin reductase (LAR) can operate the conversion into (+)-catechin and (−)-epicatechin, respectively [22,23]. The latter are then converted into anthocyanins and tannins ([21], pink route in Fig. 1). It is still not clarified whether their polymerization proceeds enzymatically or not [22,24]. Although studies reported the PCs biosynthetic pathway in higher plants [15], or in other organisms, such as lichens [25], bryophytes [26], ferns [27–28] or knowledge in algae is almost missing [29–31]. Algae, from the tiny single cell (microalga) to multicellular seaweeds (macroalga), include prokaryotic cyanobacteria and eukaryotes belonging to diverse evolutionary lineages. Scientific and applied interests in algal biology are growing in the last decade pulled by the "blue biotechnology" development. Indeed, microalgae are attractive being small organisms with fast growth rate, low double time, low nutrient requirement, and able to synthetize secondary metabolites under stress condtions for instance. Microalgae represent a relevant source of bioactive compounds and can become a biofactory for the production of compounds suitable for the market, for instance, related to human health maintenance or protection. In this panorama, PCs are of interest [31], such as in diverse industry sectors such as in nutraceutical product formulation as therapeutic agents for diabetes or cancer, in food as additives and preservatives, in cosmetics as UV-protection and antioxidant agents, and in the textile industry [32]. However, more investigations are required in microalgae being still understudied, conversely to microalgal carotenoids for instance which already fill the industrial market (e.g, astaxanthin, [33]).

This study aims to decipher the biosynthetic pathway of PCs in microalgae advancing an *in silico* analysis on a selection of core enzymes retrieved from the key steps of the biosynthesis of PCs in plants (as presented above, Fig. 1) and exploring the synthetic routes of PCs in microalgae. Yet, this investigation undertakes a comparative analysis between different microalgal divisions. This analysis consisted in i) retrieving putative algal protein homologs of selected core enzymes in the phenylpropanoid pathway, ii) annotating functional domains in each enzyme class, iii) highlight-

Table 1
Enzyme list. Full description of the 29 core enzymes of phenylpropanoid/flavonoid pathway selected for this study. Plant queries and EC annotation were retrieved from Kyoto Encyclopedia of Genes and Genomes - KEGG database.

Phenylpropanoid/ flavonoid core enzymes	Query ID	Plant species	Precursor	Product	EC number	Class
PAL phenylalanine ammonia-lyase	NP_181241; NP_187645; NP_190894; NP_196043.2	*Arabidopsis thaliana*	phenylalalnine	cinnamic acid	EC:4.3.1.24	Lyases
TAL tyrosine ammonia-lyase	XP_015625120	*Oryza sativa*	tyrosine	p-coumaric acid	EC:4.3.1.25	Lyases
C4H *trans*-cinnamate 4-monooxygenase	XP_024156591; NP_180607	*Rosa chinensis; Arabidopsis thaliana*	cynnamoyl-CoA	various	EC:1.14.14.91	Oxidoreductases
COMT caffeic acid 3-O-methyltransferase	XP_024166330; NP_001319057.1	*Rosa chinensis; Arabidopsis thaliana*	caffeic acid	ferulic acid	EC:2.1.1.68	Transferases
C3H coumarate 3-hydroxylase	AAY54293.1	*Ginkgo biloba*	p-coumaric acid	caffeic acid	EC:1.14.14.1	Oxidoreductases
C3′H C3-convertase	O22203.1	*Arabidopsis thaliana*	p-coumaroyl-CoA	caffeoyl-CoA	EC 1.14.14.96	Oxidoreductases
F5H ferulate 5-hydroxylase	AAF78943.1	*Arabidopsis thaliana*	ferulic acid	5-hydroxyferulic acid	EC:1.14.14. B13	Oxidoreductases
4CL 4-coumarate–CoA ligase	NP_175579; NP_176482.1; NP_188760.3; NP_188761.1; NP_192425.1	*Arabidopsis thaliana*	p-coumaric acid	p-cumaroyl-CoA	EC:6.2.1.12	Ligases
CCR1 cynnamoyl coA reductase	AAG46037.1	*Arabidopsis thaliana*	cinnamaldehyde	cinnamoyl-CoA	EC:1.2.1.44	Oxidoreductases
CCR2 cynnamoyl coA reductase	AAG53687.1	*Arabidopsis thaliana*	cinnamaldehyde	cinnamoyl-CoA	EC:1.2.1.44	Oxidoreductases
CAD cinnamyl-alcohol dehydrogenase	NP_177412.1; NP_179765.1; NP_179780.1	*Arabidopsis thaliana*	various	various	EC:1.1.1.195	Oxidoreductases
CHS chalcone synthase	NP_001268064; XP_002264989.1; XP_002272129.2; XP_003634064.1; XP_010655919.1; XP_010662500.1	*Vitis vinifera*	p-cumaroyl-CoA	naringenin-chalcone	EC:2.3.1.74	Transferases
STS stilbene synthase	NP_001268046.1	*Vitis vinifera*	p-coumaroyl-CoA	resveratrol	EC:2.3.1.95	Transferases
ROMT *trans*-resveratrol di-O-methyltransferase	B6VJS4.2	*Vitis vinifera*	resveratrol	pterostilbene	EC:2.1.1.240	Transferases
CHI chalcone isomerase	NP_001268033.1; XP_002280158.1	*Vitis vinifera*	naringenin-chalcone	naringenin	EC:5.5.1.6	Isomerases
FNS flavone synthase I	Q7XZQ8	*Vitis vinifera*	various	apigenin, luteolin	EC:1.14.20.5	Oxidoreductases
FNS flavone synthase II	XP_024155936; XP_024155939.1; XP_024155941.1	*Rosa chinensis*	various	various	EC:1.14.19.76	Oxidoreductases
IFS isoflavone synthase	Q9SXS3.1	*Glycyrrhiza echinata*	various	various	EC:1.14.14.87	Oxidoreductases
F3H flavanone3-hydroxylase	NP_190692	*Arabidopsis thaliana*	various	various	EC:1.14.11.9	Oxidoreductases
DFR dihydroflavonol 4-reductase	P51102.2	*Arabidopsis thaliana*	various	various	EC:1.1.1.219	Oxidoreductases
F2H 2′-hydroxyisoflavone reductase	P52575.1	*Medicago sativa*	various	various	EC:1.3.1.45	Oxidoreductases
UFGT (UDP flavonoid glycosyltransferase)	P16165.1	*Zea mays*	cyanidin, delphinidin, etc	various	EC:2.4.1.115	Transferases
FS flavonol synthase	NP_001190266.1; XP_015624815.1	*Arabidopsis thaliana; Oryza sativa*	dihydrokaempferol	kaempferol	EC:1.14.20.6	Oxidoreductases
F3′H flavanone3′-hydroxylase	NP_196416	*Arabidopsis thaliana*	kaempferol	quercetin	EC:1.14.14.82	Oxidoreductases
LAR leucoanthocyanidin reductase	XP_015630916	*Oryza sativa*	leucocyanidin	(+)-catechin	EC:1.17.1.3	Oxidoreductases
ANS anthocyanidin synthase	NP_001031700.1	*Arabidopsis thaliana*	various	various	EC:1.14.20.4	Oxidoreductases
ANR anthocyanidin reductase	XP_015635209.1; XP_015637098; XP_015637099.1	*Oryza sativa*	cyanidin	(−)-epicatechin	EC:1.3.1.77	Oxidoreductases
3-dehydroquinate dehydratase / shikimate dehydrogenase	XP_002884573	*Arabidopsis lyrata*	3-dihydroshikimate	shikimate	EC:4.2.1.10	Lyases
polyphenol oxidase	NP_001268045.1; XP_010647098.2	*Vitis vinifera*	various	various	EC:1.10.3.1	Oxidoreductases

ing conserved aminoacidic motifs into the selected domains of interest, and iv) constructing maximum-likelihood phylogenies.

2. Materials and methods

2.1. Dataset of protein sequences and screening for homologs

Twenty-nine core enzymes of phenylpropanoid/flavonoid pathway were selected from the Kyoto Encyclopedia of Genes and Genomes (KEGG) reference pathway in plants (Table 1).

A BLASTp search (v. 2.7.1.) [34] using the enzymatic classes as queries against the entire UniProt Swiss-Prot protein database [35] was performed to detect orthologs of phenylpropanoid and flavonoid pathway enzymes within different algal divisions. Analytical outputs regarded the manually annotated and reviewed protein belonging to Cyanobacteria, Bacillariophyta, Chlorophyta, Cryptophyta, Rhodophyta, Dinophyta, Euglenozoa, Haptophyta, Eustigmatophyta and Phaeophyta-Xanthophyta clades. A BLASTp search against the UniProt TrEMBL protein database [36] was also done, including the computationally analyzed records – not yet annotated - belonging to the same clades. The best hit matches for each species were retained for down-stream analyses. Using a subset of representative taxa in each microalgal division (Table 1), a heatmap from the matrix of normalized BLAST scores was built using Graphpad Prism 8.0 software. A principal component analysis (PCA, using a variance–covariance matrix) was carried out on the same matrix using PAST 4.04 software [37].

2.2. Multiple alignments and Maximum-Likelihood (ML) phylogenetic analysis

Sequences retrieved by BLASTp search in the UniProt TrEMBL protein database were first aligned with MAFFT software v7.397 [38] applying default parameters to generate phylogenetic trees for the core enzymes. Alignments were then cleaned with TrimAl v1.4 [39] using the values 0.25 for gap threshold, 0.25 for residue overlap threshold, and 0.9 for sequence overlap. The best-fit model of molecular evolution for each dataset was selected with ModelFinder, implemented in IQ-TREE v1.6.11 [40] and operating the corrected Akaike Information Criterion (AICc). Maximum-likelihood trees were constructed with FastTree v2.1.11 [41] applying default parameters. All trees were visualized and edited using iTOL v.5 [42].

2.3. Functional annotation and selection of Domains of Interest (DOIs)

Functional annotation was performed using the software InterProScan (version 5.33) [43] to retrieve the InterPro Domains associated to the queries together with their coordinates along the investigated algal sequences collected in the UniProt database. A selection of the domains of interest (DOIs) was manually curated, discarding non-characteristic or ambiguous domains, onlyretaining the functional ones. For each best hit protein satisfying quality criteria and DOIs selection, only the sequence region corresponding to the coordinates of the detected InterPro Domain was retained. Followingly, the retained domain sequences detected for each enzymatic class were used to identify the enrichment motifs among classes by MEME and STREME tools of MEME suite software (v. 5.3.3.) [44]. Higher plants sequences collected by the previous BLAST search and subjected to same data treatment were then employed as reference sequences for STREME consensus analysis. Relevant MEME motifs in selected DOIs were added as visual data to ML-phylogeny, whereas positive algal DOIs in STREME analysis were then parsed on p-value score and percent of positive algal sequences (Table A.1).

3. Results and discussion

The BLAST scores contribution of the twenty-nine core enzymes (Table 1) among the different microalgal groups showed a great variability (Fig. 2), confirmed by the PCA (axis 1: 46.4%; axis 2: 18.2% of the variability, Fig. 3). Cyanobacteria was heterogeneous and dispersed radiantly among quadrants (Fig. 3**)**. While ancestral prokaryotes, Cyanobacteria display the greatest variability in the pathway, revealing their evolution with time [45,46]. Primary (Chlorophyta, Rhodophyta, Eustigmatophyta) and secondary (Bacillariophyta, Cryptophyta, Charophyta, Haptophyta, Dinophyta) endosymbiotic algae displaced toward opposite direction, suggesting that the endosymbiotic gene transfer (EGT) played a fundamental role as driver in the evolution of phenylpropanoids in different phyla. The only tertiary endosymbiotic alga in the dataset, *Euglena*, encountered massive gene loss, strongly decreasing its weight in the analysis (Fig. 3). Following this first descriptive overall approach, the analysis of key enzymes building algae-based maximum likelihood phylogenetic trees was then performed to compare the biosynthetic pathway of the main PCs amongst microalgae.

3.1. The routes for lignins, coumarins and gallic acid derived-tannins

3.1.1. Phenylalanine ammonia lyase (PAL)

In plants, the enzymes PAL (phenylalanine ammonia lyase) or TAL (tyrosine ammonia lyase) catalyse the non-oxidative deamination of phenylalanine (or tyrosine) to *trans*-cinnamate and direct the carbon flow from the shikimate pathway to the various branches of the general phenylpropanoid metabolism (Fig. 1). PAL activity produces *trans*-cinnamic acid, whose derivatives have been detected in several algal divisions [31]. While the activity of PAL enzyme has been demonstrated *in vitro* in cyanobacteria [47], to the best of our knowledge no experimental demonstration has been provided in eukaryotic algae. Structural data of PALs are available for the cyanobacteria *Anabaena* and *Nostoc* [48] and for the streptophyte algae *Klebsormidium* and *Nitella* [30]. In Cyanobacteria, PALs are smaller in size than their eukaryotic counterparts; however, they show similar substrate selectivity and kinetic activity toward L-phenylalanine over L-tyrosine [48]. Structure elucidation by protein X-ray crystallography confirmed that the two cyanobacterial PALs are similar in tertiary and quaternary structure to plant and yeast PALs [48].

In our study, the occurrence of PAL (or TAL) homologs in algae showed disparate hits: its presence was recorded in the brown seaweed *Ectocarpus* and in the haptophyte *Emiliania huxleyi,* and to a lesser extent within Cyanobacteria and Chlorophyta, as shown in Fig. 2 and Fig. 4. This unexpected feature addresses the question on the initiation of the biosynthetic phenylpropanoid pathway in the remaining considered algal divisions, e.g. in Bacillariophyta or Rhodophyta, in which phenolic compounds production is however well documented. With the exception of *Ectocarpus,* PAL deficiency appears to be common to all secondary endosymbiotic algae, while it could have been partly loss in Cyanobacteria and Chlorophyta.

PAL sequences from *Arabidopsis thaliana* and a TAL sequence from *Oryza sativa* (Table 1) were employed to build a ML phylogeny tree using both enzymes which own an aromatic amino acid lyase domain (IPR001106) (Table A.1). The Maximum-likelihood (ML) phylogenetic analysis of PAL in the diverse microalgal groups (Fig. 4) agrees with the previously reported phylogenetic observations [49,50]. The brown alga *Ectocarpus* and the two Chlorophytes *Micractinium* and *Gonium* shared a very robust node (100) and cluster with the archaeal *Halobacterium salinarum* (77.5). Notably, conserved motifs of PAL-domain in *Gonium* were reversed, lone case

Fig. 2. Heatmap build by normalized BLASTp scores obtained from plant entries of 29 selected core enzyme in phenylpropanoid pathway using. A subset of 47 algal taxa.

Fig. 3. PCA of normalized BLASTp scores obtained from plant entries of 29 selected core enzymes in phenylpropanoid pathway. Cyanobacteria (dots), primary endosymbiotic event (triangles) and secondary endosymbiotic event (squares) eukaryotic algae are plotted following the colour legend in Fig. 2. Tertiary endosymbiotic event alga *Euglena* is marked by a void circle.

among the MEME analyzed sequences, while *Ectocarpus* did not show reliable motif conservation. The haptophyte *Emiliania* appeared to be closely related to the Euglenozoa clade (98.5) and both shared a robust node with *Candidatus Marinamargulis bacterium* (98.3). For this Carbon-Nitrogen lyase, cyanobacterial species belonged to diverse clusters composed by various classes of Eubacteria, probably accounting for duplication and loss among prokaryotic algae and hinting for a hidden paralogy.

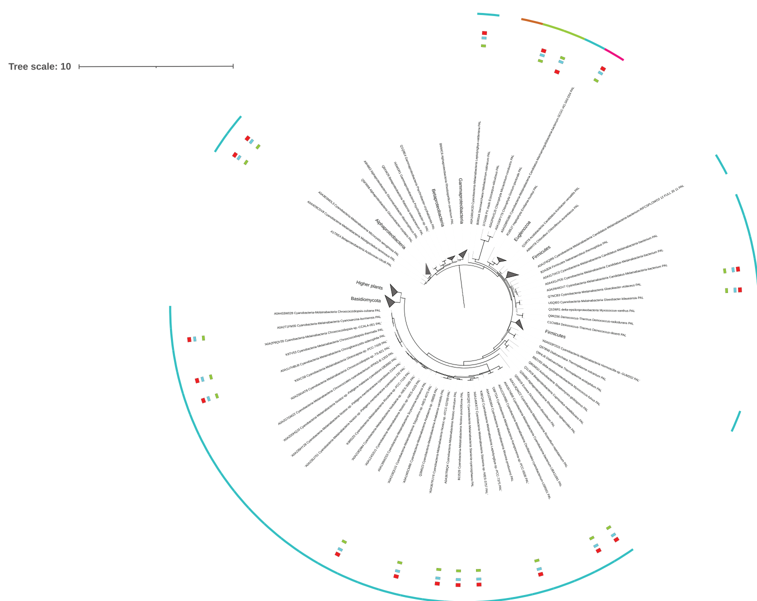


Fig. 4. ML-tree for PAL enzyme. Phylogeny was obtained starting from 236 protein sequences. Thicker clades indicate node support >75%. Green, cyan and magenta boxes represent motifs retrieved by MEME analysis of algal sequences within the domain IPR001106. Algal divisions are indicated as color arches, following the legend presented in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.1.2. Shikimate dehydrogenase (SDH) and 3-dehydroquinate dehydratase (3-DHQ)

Plant SDH enzymes are fused to dehydroquinate dehydratases to form a bifunctional SDH/DHQ. This enzyme catalyzes the dehydration of 3-dehydroquinate to 3-dehydroshikimate (3-DHS) and the reversible reduction of the latter into shikimate [51]. The 3-dehydroquinate can be converted to quinic acid, an important precursor for aromatic amino-acids and for other molecules, some with notable herbicidal or pharmacological activities [52]. It is also noteworthy that dehydroshikimate is the precursor for abscisic acid (ABA) hormone and protocatechuic acid, both reported in microalgae [31,53]. Protocatechuic acid has been identified in different microalgal groups, i.e., the diatom *Phaeodactylum tricornutum*, the Eustigmatophyte *Nannochloropsis gaditana* and the green alga *Nannochloris* sp. and *Tetraselmis suecica* [54]. Interestingly, protocatechuic acid was reported as growth promoter for the microalga *Euglena gracilis* [55]. Use of protocatechuic acid might open new cost-effective strategy for improving the microalgal productivity in large-scale cultivations. The plant bifunctional SDH/DHQ enzyme presented the highest similarity score within Chlorophyta and Rhodophyta groups (Fig. 2). A weaker although substantial similarity score was reported for diatoms, *Chrysochromulina*, *Emiliania*, and *Symbiodinium* (Fig. 2). No SDH/DHQ enzyme hit was retrieved in Cyanobacteria and *Euglena.* Interestingly, microalgae which do not possess PAL enzyme hold 3-DHQ enzyme with the higher scores (Fig. 2), opening a question on the putative role of SDH/DHQ enzymes in the onset of phenylpropanoid pathway. The 3-DHQ algal homologs contain a domain associated with prokaryotic shikimate-5-dehydrogenase, a Carbon-Oxigen lyase in the SDH superfamily (IPR011342) (Table A.1). This domain, involved in NADP-dependent reduction of 3-dehydroshikimate to shikimate is present in several algal phyla, even though patchy distributed among three clades (Fig. A.1): a) *Ectocarpus*, Eustigmatophyta, Bacillariophyta; b) Cyanobacteria; c) Chlorophyta, Rhodophyta, Cryptophyta, Haptophyta, and a candidate cyanobacterial division.

3.1.3. 4-Coumarate ligase (4CL)

The co-A ligase 4CL and the hydroxylase C4H are upstream enzymes for the synthesis of flavonoids, stillbenoids, coumarins, lignins, catechins, and aurones through the intermediate p-coumaroyl-CoA.

In plants, 4-coumarate ligase 4CL catalyzes the stepwise conversion of hydroxycinnamic acids into the corresponding CoA esters via the intermediate AMP derivatives. A sequence similarity between prokaryotic and eukaryotic 4CL has been reported [56] and both does function via an ATP-dependent covalent binding of AMP to their substrate.

The precursor p-coumaroyl-CoA is necessary for the synthesis of several flavonoid derivatives (i.e., 5-deoxyflavonoids or dihydroflavonols) present in the major algal divisions of cyanobacteria, diatoms and green algae [31]. The BLAST search for 4CL revealed substantial similarity score in cyanobacteria, diatoms, and haptophytes (Fig. 2). Although the haptophytes *Emiliania* and *Chrysochromulina* as well as diatoms underwent a secondary endosymbiosis event, only diatoms showed a true protein orthology for 4CL in our study (Fig. 5). Haptophytes and diatoms showed the highest similarity scores for 4CL with respect to higher plants (Fig. 2). The common node between these two groups and green algae in the 4CL-based phylogenetic tree might reveal a green algal ancestor in the evolution of plant 4CL enzyme (Fig. 5). Remarkably, the characteristic domain IPR000873 (owned by all sequences in our phylogeny) highlighted a Ser/Thr/Gly-rich trait followed by a conserved Pro-Lys-Gly triplet (see red box in Fig. 5) that remains highly and stably conserved within all the analysed algal taxa.

3.1.4. Cinnamoyl Co-A reductase (CCR) and cinnamoyl alcohol dehydrogenase (CAD)

The enzymes CCR1/2 and CAD involved in lignin biosynthesis in plants [17] appear conserved in algae, with the only exception of the tertiary endosymbiotic alga *Euglena* (Fig. 2).

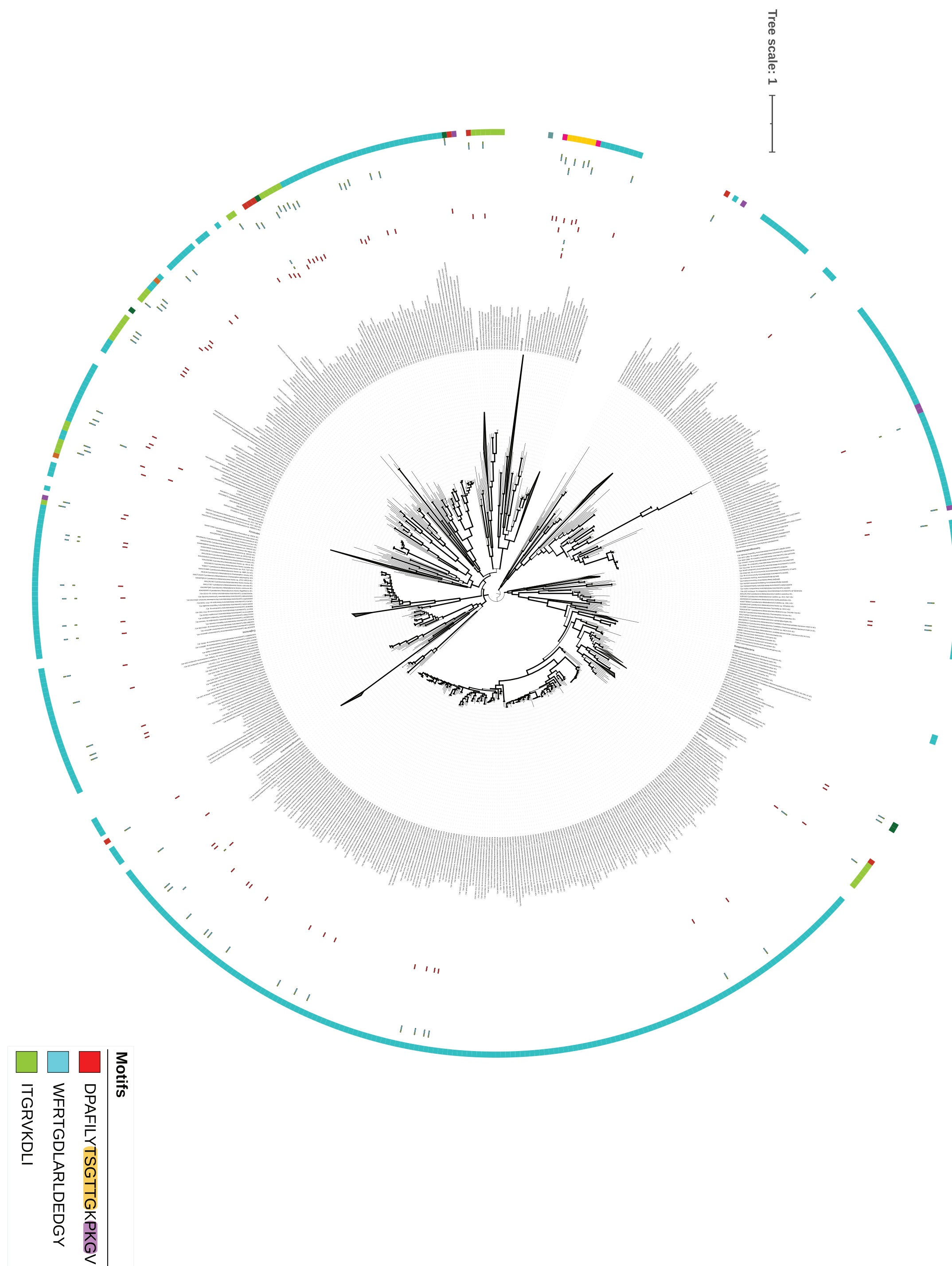


Fig. 5. ML-tree for 4CL enzyme. Phylogeny was obtained starting from 882 protein sequences. Thicker clades indicate node support >75%. Green, cyan and magenta boxes represent motifs retrieved by MEME analysis of algal sequences within the domain IPR000873. The Ser/Thr/Gly-rich trait and a conserved Pro-Lys-Gly triplet are highlighted in the red box. Algal divisions are indicated as color arches, following the legend presented in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Plant CCR 1 and 2 convert hydroxycinnamoyl CoA esters to their corresponding cinnamyl aldehydes in monolignol biosynthesis and differ by substrate affinity towards feruloyl-CoA and caffeoyl-CoA, respectively [57]. Algal homologs for CCR displayed a cyanobacterial ancestry (Fig. A.2) and subsequent gene duplication would be supported by paralogy of Bacillariophytes, Haptophytes and Chlorophytes.CADs are zinc-dependent dehydrogenases and catalyze the last step of monolignol biosynthesis by reducing cinnamaldehyde into cinnamyl alcohols prior to their transport through the membranes and polymerization in the cell wall [58].

In higher plants three classes of CADs are present, following a scheme of duplication and extinction with complex spatial and temporal regulation [59]. Lignin, lignin-like compounds, and sporopollenin were found in the cell wall of seaweeds, such as in *Coleochaete* (Chlorophyta) [60], *Cystoseira* (Phaeophyta) [61] and *Sargassum* (Phaeophyta) [62] as well as in microalgae [63–65]. It has been recently elucidated that sporopollenin precursors originate in the phenylpropanoid pathway in plants, following a scheme that is superimposed to canonical lignin biosynthesis [66]. The discovery of lignified cell wall in the seaweed *Calliarthron* enlightened the possibility of convergent evolutionary history of this trait, given the early separation among green and red algal lineages [67]. Recently, a hypothesis of lateral gene transfer from bacteria to algae has been proposed to explain the presence of lignin-related compounds in seaweeds [68].

In our study, beside red and green algae, also diatoms (possessing a silica cell wall), cyanobacteria and *Euglena* showed remarkable similarity scores for CAD (Fig. 2). Very few studies investigated the presence of lignins in microalgae [69]. Lignin has been reported in *Staurastrum* sp. and in a mixture of cyanobacteria and eukaryotic microalgae [69]. This surprising feature might indicate that in microalgae, lignin precursors and CAD activity may play another role that the deposition of lignin in the cell wall. Hydrolized lignin and lignin-derived products from plants are known for their capacity to inhibit algal growth [70,71]. It is noteworthy that lignin might have important biotechnological potential e.g., for biopolymers, heavy metal removal or energy [72], requiring more investigations in microalgae. The phylogeny analysis revealed that a large group of OTUs including also plant and bacteria shares the domain IPR002328 (Zinc-containing 'long-chain' alcohol dehydrogenases) (red star clade in Fig. 6). Within the red star clade, a great group of eubacteria clusters together with cyanobacteria and eukaryotic algae, suggesting a common prokaryotic origin for this enzyme.

Dihydroflavonol 4-reductase– DFR is an oxidoreductase involved in the diversification of flavonoids starting from dihydrokaempferol or dihydroquercetin, both flavonols reported in algae [31]. Both CAD and DFR (Table 1) are oxidoreductases acting on CH-OH group of donors with NAD+ or NADP+ as acceptor; although they catalyze different reactions, the two enzymes own a very similar asset. Previous reports indicated significant CAD homologies with dihydroflavonol-4-reductases from various plant species [73]. It is known that minor changes in DFR sequence may determine alteration in substrate specificity [74]. DFR sequences included in our phylogeny do also present the NADP-binding domain (IPR001509) required for the enzymatic activity (Table A.1). The CAD sequences interspersed in the DFR clades were not featured by neither Zinc-containing 'long-chain' alcohol dehydrogenases domain, nor by a NADP-binding domain. Both CAD and DFR enzymes displayed a horizontal distribution among algal divisions (Fig. 2), and the tree topology (Fig. 6**)** hinted that ancestral algal CADs may have diversified into DFRs, establishing separate lignin and flavonoid routes.

3.2. Routes for stillbenoids, flavonoids and isoflavones

3.2.1. Chalcone synthase (CHS) and Stillbene synthase (STS)

Chalcone synthase and Stillbene synthase are two transferases involved in the onset of the flavonoid pathway and stillbenoids biosynthesis (i.e., resveratrol), respectively. Stillbenoids display a heterogeneous distribution in the plant Kingdom while not yet revealed in algae [75].

CHS catalysis serves as initial step for flavonoid biosynthesis, associated with the production of chalcones [76]. This enzyme, ubiquitous in higher plants, is known as type III PKS (Polyketide synthase type) in the family of polyketide synthase enzymes (PKS). CHS followed a patchy distribution among algal divisions, displaying the highest similarity scores in the brown macro-alga *Ectocarpus,* although lacking in other groups e.g., diatoms (Fig. 2). The presence of several flavonoid compounds reported in diatoms [77,29] might indicate that flavanol reduction could be deputed to other oxidoreductase(s) with larger substrate specificity. Yet, it was hypothesized that styrylchromone toxins isolated from the marine cryptophyte *Chrysophaeum taylori* are formed by a type III PKS enzyme [78], while our study did not identify CHS homologs in another marine cryptophyte, *Guillardia theta* (Fig. 2 and Fig. 7).

In the algal CHS-based ML-phylogeny tree (Fig. 7), generally cyanobacteria ranked at a basal position, although a subsequent clusterization did exist in higher clades. This happened as a sister clade in the red algae division (orange star in Fig. 7) as well as in the apical branch containing higher plants as a sister clade of brown algae (pink star in Fig. 7). This distribution would suggest the rise of CHS enzyme more than once in algal divisions, converging with the differences among the widespread flavonoids present in green algae and the more complex flavonoids of red and brown seaweeds [21].

Consensus retrieval for the identified domains indicated the presence of Polyketide synthase type III domain (IPR011141), and a 3-oxoacyl-[acyl-carrier-protein]-synthase3 domain (IPR004655) (Table A.1). The first, deputed to perform three sequential condensation steps with acetate units from malonyl-CoA to form a tetraketide intermediate starting from a CoA-ester, showed horizontal conservation among all microalgal divisions. The second (responsible of the elongation in plant type II fatty acid synthase systems) was only present in a restricted group of algae sharing one common node and including Cyanobacteria, Rhodophyta, Chlorophyta, *Nannochloropsis*, and *Symbiodinium* (green star in Fig. 7).

3.2.2. Chalcone isomerase – CHI

Chalcone isomerase – CHI drives the fundamental step to generate flavanones. Isomerasic activity is relevant for the diversification of flavonoids with e.g., the naringenin-chalcone conversion into naringenin. This colourless flavanone has been reported in the haptophyte *Diacronema lutheri* and in the chlorophyte *Haematococcus pluvialis* [29]. Its distribution appeared heterogeneous among algal divisions (Fig. 2), being absent in cyanobacteria, red algae, and diatoms. The function of CHI, and thereby the ability to synthetize flavanones, could be originated from the green algal lineage, and followingly encountered gene duplication and selective loss, persisting in Chlorophyta, Phaeophyta, Cryptophyta and in the class of Eustigmatophyceae, as well as in higher plants (Fig. A.3). In STREME analysis of retrieved sequences, the DOI of chalcone isomerase domain (IPR016087) showed 100% positive algal sequences in a motif consensus comparison with higher plants (Table A.1).

Flavanones can be further converted to flavones or to dihydroflavonols thanks to the flavone synthase II (FNS) or to the flavanone3-hydroxylase (F3H) enzyme, respectively [18].

These products can be then converted to flavonols (through flavonol synthase (FLS) enzyme activity), or to leucocyanidin (through flavanone3′-hydroxylase (F3′H) or dihydroflavonol 4-reductase (DFR) enzyme activities), entering into the anthocyanidin pathway [18]. A large subset of enzymes deputed to the modification and interconversion of flavonoids appear to be evenly present among microalgal divisions, with the only exception of cyanobacteria. However, several flavonoids were found in some cyanobacterial species (e.g., quercetin, rutin, apigenin, daidzein, catechin or epicatechin [31]) addressing the hypothesis of the non-enzymatic production of these molecules or on the presence of other enzymes in cyanobacteria. Isoflavones, such as daidzein or genistein were retrieved in the green microalga *Haematococcus pluvialis,* together with flavones or flavonols [29]. Also, they have

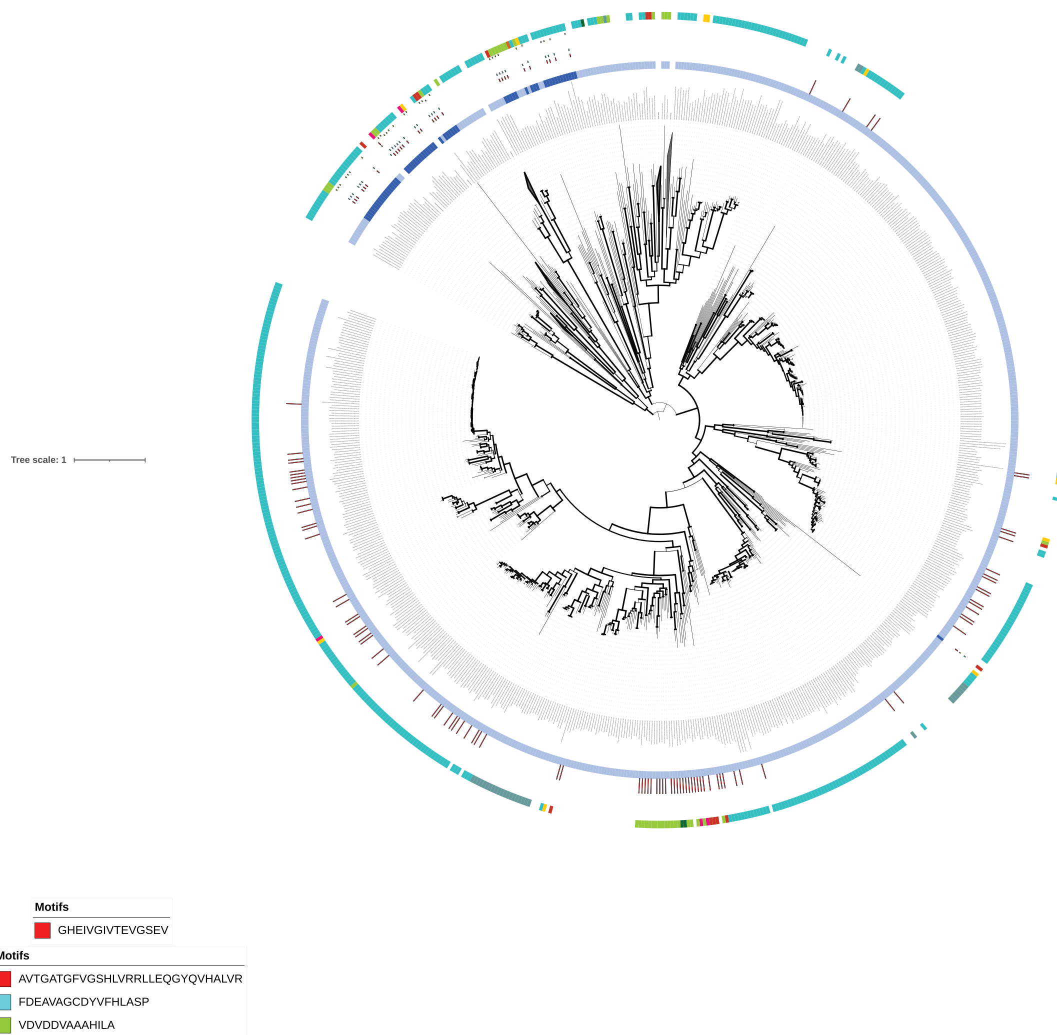


Fig. 6. ML-tree for CAD-DFR enzymes. Phylogeny was obtained starting from 722 CAD and 109 DFR protein sequences. Thicker clades indicate node support >75%. At the end of leaves, dark and light blue colors differentiate DFR and CAD sequences, respectively. Red bars and green, cyan and magenta boxes represent motifs retrieved by MEME analysis of algal sequences within the respective domains IPR002328 and IPR001509. Algal divisions are indicated as color arches, following the legend presented in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

been found in the diatom *Phaeodactylum tricornutum* and in the red microaga *Porphyridium* [29]. The lack of information on the distribution in other species of these two isoflavones prevent any taxonomical-derived hypothesis.

3.3. Route for anthocyanins

3.3.1. Anthocyanidin reductase (ANR) and leucoanthocyanidin reductase (LAR)

ANR uses cyanidin as a precursor to form (−)-epicatechin, while LAR synthetizes (+)-catechin by transforming leucocyanidin. Both catechin and epicatechin are found in red and brown algae [79,80], in the diatom *Phaeodactylum tricornutum* [77], in *Euglena gracilis* [81], and in some green algae [80,81] and cyanobacteria [80–82]. Catechin and epicatechin present biological interests, being reported as antioxidant and anti-inflammatory [83].

The BLASTp search for DFR, ANS (anthocyanidin synthase) and ANR (anthocyanidin reductase), enzymes necessary for the synthesis of catechin and epicatechin (Fig. 2), is in agreement with previous reported observations in algae; difference regards *Euglena*, in which no substantial hits were retrieved despite experimental reports for both (−)-epicatechin and (+)-catechin [81].

Recent studies hypothesized an algal origin for anthocyanins in plants although its presence in algae remains uncertain [84]. Cyanidin presence in algae has never been revealed, although free and glycosylated catechins were detected in several algal divisions [31]. The analysis of conserved DOIs in ANR and LAR putative algal homologs, reveals the presence of a NAD-dependent epimerase/dehydratase domain (IPR001509) for ANR, and the presence of a NmrA-like domain (IPR008030) for LAR [85] (Table A.1).

3.3.2. UPD-glucose:flavonol 3-O-D glucosyltransferase (UFGT)

Another transferase which proceeds downstream the flavonoid pathway (light-green and pink routes in Fig. 1) is the UPD-glucose: flavonol 3-O-D glucosyltransferase (UFGT). In plants this glucosyltransferase catalyzes the transfer of glucose from UDP-glucose to a flavanol, a step that is necessary to proceed toward anthocyanin pigment biosynthesis. A study reports that the UGT family expanded during the transition from algae (e.g., the chlorophyte *Chlamydomonas reinhardtii*) to vascular plants [86]. UFGT operates the conversion of intermediate anthocyanidins into a variety of

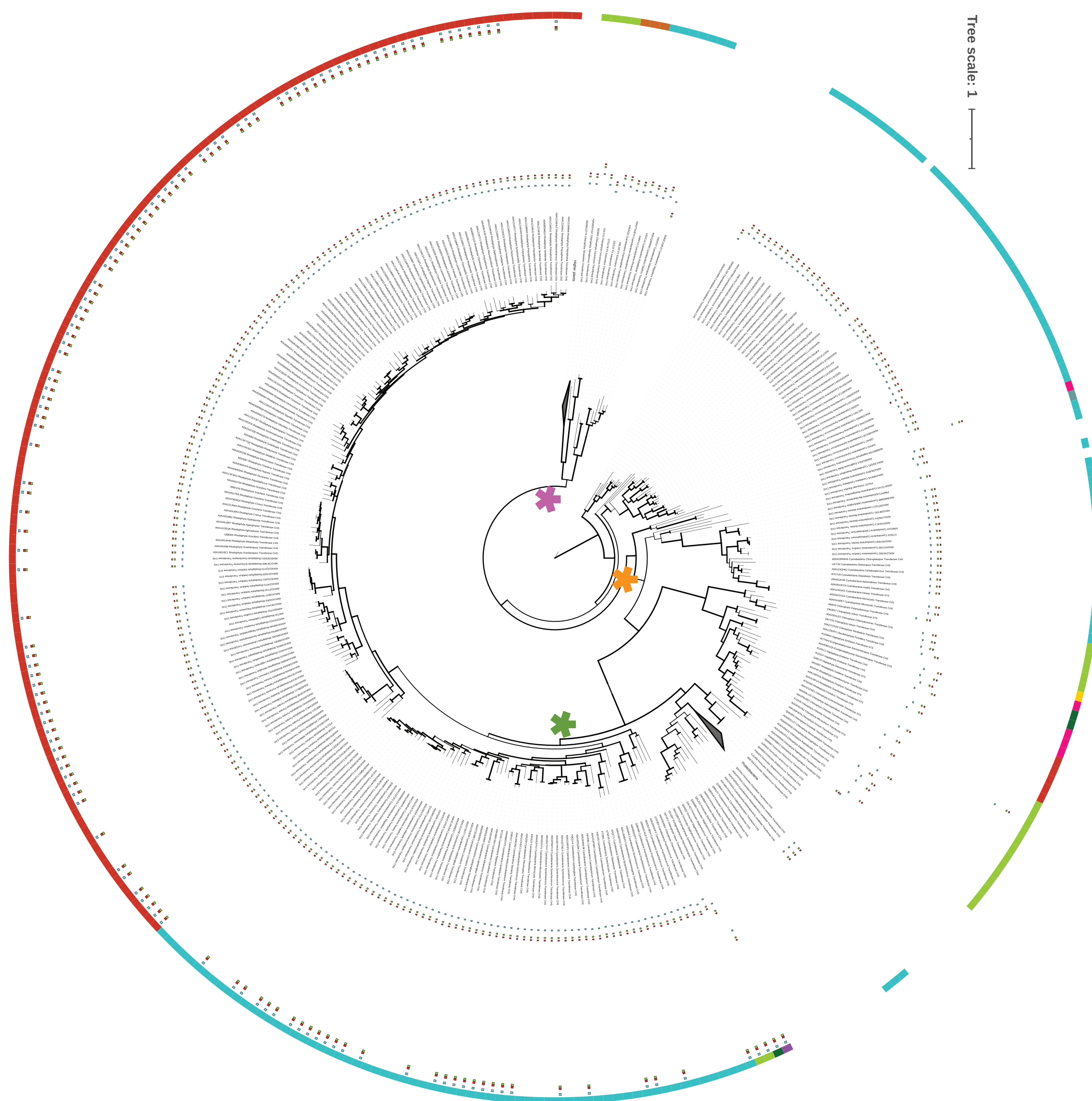


Fig. 7. ML-tree for CHS enzyme. Phylogeny was obtained starting from 506 protein sequences. Thicker clades indicate node support >75%. Green, cyan and magenta boxes represent motifs retrieved by MEME analysis of algal sequences within the domains IPR011141 (inner) and IPR004655 (outer). Algal divisions are indicated as color arches, following the legend presented in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

anthocyanins, which to date, have not been reported in algae [87]. Only one study reported the presence of cyanidin and malvidin derivatives in the green seaweed *Caulerpa racemosa* [88]. Recent studies point out an ancestral origin of anthocyanins and tannins in algae [89]. However, uncommon purple pigments with a chemical structure closely related to anthocyanins have been rarely found in algae, for instance in the cyanobacterium *Tolypothrix tenuis* [90], or in the Charophyta *Zygogonium ericetorum* [91,92]. Also, the blue pigment marennine to date retrieved only from the diatom *Haslea ostrearia* might belong to this group [93,94]. This pigment presents many commercial or biological interests, due to its antioxidant, antibacterial or antiviral activities [94]. The BLAST search for UFGT homologs in algae revealed low identity scores, with substantial hit in *Ectocarpus*, *Chrysochromulina* and partly in diatoms and green algae (Fig. 2). No hits were retrieved in Rhodophyta, Cryptophyta, Eustigmatophyceae and Dinophyceae.

ML phylogeny for *bona fide* algal UFGT (Fig. A.4) did not show significant relation with prokaryotic or eukaryotic non-photosynthetic OTUs. A first subgroup of unicellular green algae clustered intimately with higher plants, while another subgroup revealed more ancestral traits, being distributednear to other microalgae, including cyanobacteria. To date, there is no report of UFGT enzyme activity in algae, even though its functional domain (IPR002213) displays a conserved region among the considered algal sequences (Table A.1).

3.4. Branching off the pathway

The chemodiversity of PCs in microalgae is further enhanced by possible further diversifications, such as additional methylation or multiple hydroxylation. These modifications can change the role or the bioactive capacity of PCs without dramatical change on the carbon skeleton.

3.4.1. O-methyltransferases (OMT)

The enzymatic O-methyltransferasic activity is crucial to direct the biosynthesis of lignins with the caffeic acid 3-O-methyltransferase (COMT) or the diversification of stillbenoids, with *trans*-resveratrol di-O-methyltransferase (ROMT). Since stillbenoids are not retrieved in algae, the putative OMT activity of

the retrieved homologs relates with the interconversion of phenolic acids starting from caffeic acid (or caffeoyl-CoA) intermediate to generate ferulic acid (or feruloyl-CoA) before further modification by CCR and CAD to synthetize lignin and lignin like compounds and diversify phenolic acids (red route in Fig. 1). Putative COMT homologs were found in a subgroup of green algae positioned as a basal clade in our phylogeny (Fig. A.5). Land plant sequences were located as a sister clade of cyanobacteria and diatoms, both sharing a common node with a wide group of eukaryotic and prokaryotic microalgae. The only two bacterial putative COMT present in the phylogeny, belonging to the genus of photosynthesizing nonsulfur purple bacteria *Rhodobacter*, intimately cluster within the basal chlorophyte clade with a significant node support. This relation could reveal an Horizontal Gene Transferevent for the evolution of OMT in the microalgal world.

3.4.2. Oxidoreductases

Oxidoreductases can extensively increase the degree of diversification of phenolic compounds. The selected oxidoreductases can be grouped in four macro-classes on the basis of EC number, namely: a) *EC 1.1* – Acting on the CH-OH group of donors (i.e., CAD, DFR); b) *EC 1.2* – Acting on the aldehyde or oxo group of donors (i.e., CCRs); c) *EC 1.3* – acting on the CH-CH group of donors (ANR, F2H); 4) *EC 1.14* – paired donors with incorporation of molecular oxygen (i.e., C4H, FS, ANS).

The role of C4H, deputed to the conversion of cinnamic acid into p-coumaric acid (or cinnamoyl-CoA into p-coumaroyl-CoA), may be bypassed in case Tyr is initiator of the pathway, due to the preliminary and necessary conversion of Phe into cinnamic acid [15]. The oxidoreductases C3H, C3′H and F5H further increase the diversity of phenolic compounds acting on a number of substrates [15]. In the Fig. 1, the enzymes C3H, C4H, F5H (red route), IFS, F2H, F3H, FLS (light-green route), FNS (purple route), and ANR, ANS (pink route) are represented.

The algal homologs of the oxidoreductases listed in Table 1 involved in the branching off of the pathway were included in a phylogeny to evaluate possible interrelation (Fig. A.6). The considered oxydoreductases displayed a high degree of inter-dispersion among the different clades, in agreement with results reported for C4H or F5H which show ambiguous orthologs in land plant phylogenies [50].

3.5. Extracellular polyphenol oxidases and secretion of phenolic compounds

It is reported that a number of cyanobacteria and green algae are able to degrade phenolic compounds outside the cell wall [95–100]. In Chlorophyta cultivation, the release of polyphenol oxidases with laccase-like activity in the medium reaches its maximum during the stationary growth phase and seems to increase with copper sulphate administration [101]. This might be of interest for biotechnological remediation of polluted waters [101].

In our study, extracellular polyphenol-oxidase enzymes resulted in spotted hits in several groups of microalgae, including cyanobacteria and green algae, as well as the microalgae *Fragilariopsis* (Bacillariophyceae), *Galdieria* (Rhodophyta), *Nannochloropsis* (Eustigmatophycea), *Emiliania* (Haptophyta), *Symbiodinium* (Dynophyta), and the seaweeds *Porphyra* (Rhodophyta) and *Ectocarpus* (Phaeophyta) (Fig. 2). These findings suggest that the ability to secrete enzymes able to degrade phenolic compounds in the external environment may be present in diverse algal divisions and have profound ecological implications in determining the efficacy of extracellular signalling [102].

For example, it is known that aquatic plants may exert an algicidal allelopathy especially towards cyanobacteria and green microalgae through the secretion of various phenolic compounds [103].

This feature depends on the considered species and/or compounds. For instance, Wang and co-workers [104] observed that submerged macrophytes inhibited the growth of the microalgae *Microcystis aeruginosa* (Cyanobacteria) and *Pseudokirchneriella subcapitata* (Chlorophyta). This growth inhibition has been attributed to the synthesis of two phenolics, (+)-catechin and pyrogallic acid, while other phenolics (e.g., gallic, ellagic, protocatechuis or caffeic acid) did not affect growth [104].

In another study, the addition of the two flavones apigenin and luteolin to the medium of the cyanobacterium *Mycrocystis aeruginosa* lowered the growth rate in concentration-dependent manner [105]. Also, the addition of the flavanone naringenin to different cyanobacteria cultures induced growth inhibition in six species and no significant effect in two cultures [106]. In all the species naringenin was internalized into the cyanobacteria cell wall structures, while to date the compound naringenin has not been detected in Cyanobacteria [106].

4. Conclusions

Enzymatic machinery determining key intermediates in the phenolic compounds synthetic pathway described in algae is almost conserved in all major algal divisions, suggesting a common cyanobacterial origin and an intra-phylum descent, with orthology into mono-phyletic groups. One exception concerns the initiating PAL enzyme that is not recurrent in all microalgal taxa, potentially suggesting alternative routes to initiate the PCs biosynthesis pathway. However, the pathway branches off to multiple directions thanks to a series of multifunctional enzymes as oxidoreductases, that can repeatedly and greatly modify compounds with a common backbone, although different algal groups may differ for the end-point products.

Endosymbiotic events and species radiation may have later diversified the PCs pattern, diverging toward more complex chemical structures and more specifical functions, e.g., diatom flavonoids, halogenated PCs in seaweeds, algal anthocyanins or algal lignins. Since the little information available on phenolic compounds determination in microalgae, it is not yet possible to draw some hypotheses or conclusion on potential relationship between specific compounds and taxonomical division. This study contributes to improve the knowledge of PCs in microalgae and lays the foundation for future synthetic biology or bioengineering approach to make microalgae a cell factory for the synthesis of bioactive PCs.

CRediT authorship contribution statement

Del Mondo Angelo: Conceptualization, Methodology, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Sansone Clementina:** Conceptualization, Writing – review & editing, Project administration. **Brunet Christophe:** Conceptualization, Writing – review & editing, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by Stazione Zoologica Anton Dohrn and by “Antitumor Drugs and Vaccines from the Sea (ADViSE)” project

(PG/2018/0494374). Angelo del Mondo is supported by a post-doctoral fellowship funded by "Antitumor Drugs and Vaccines from the Sea (ADViSE)" project (PG/2018/0494374).

The authors acknowledge Drs. Ambrosino Luca and Miralto Marco (from the Service Bioinforma of the SZN) for their help and advices during the *in silico* analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.04.019.

References

[1] Ancillotti C, Bogani P, Biricolti S, Calistri E, Checchini L, et al. Changes in polyphenol and sugar concentrations in wild type and genetically modified *Nicotiana langsdorfi* Weinmann in response to water and heat stress. Plant Physiol Biochem 2015;97:52–61.
[2] Smirnov OE, Kosyan AM, Kosyk OI, Taran NY. Response of phenolic metabolism induced by aluminium toxicity in *Fagopyrum esculentum* moench. plants. Ukr Biochem J 2015;87:129–35.
[3] Sharma A, Thakur S, Kumar V, Kanwar MK, Kesavan AK, et al. Pre-sowing seed treatment with 24-epibrassinolide ameliorates pesticide stress in *Brassica juncea* L. through the modulation of stress markers. Front Plant Sci, 2016, doi:10.3389/fpls.2016.01569.
[4] Handa N, Kohli SK, Sharma A, Thukral AK, Bhardwaj R, et al. Selenium modulates dynamics of antioxidative defence expression, photosynthetic attributes and secondary metabolites to mitigate chromium toxicity in *Brassica juncea* L. plants. Environ Exp Bot 2019;161:180–92.
[5] Wang J, Yuan B, Huang B. Differential heat-induced changes in phenolic acids associated with genotypic variations in heat tolerance for hard fescue. Crop Sci 2019;59:667–74.
[6] Naikoo MI, Dar MI, Raghib F, Jaleel H, Ahmad B, et al. Role and regulation of plants phenolics in abiotic stress tolerance: an overview. In: Khan IM, Reddy PS, Ferrante A, Khan N, editors. Plant signaling molecules. Amsterdam, The Netherlands: Elsevier; 2019. p. 157–68.
[7] Nair RM, Hanumantha Rao B, Vemula A, Rathore A, Brumbarova T, et al. Availability of soil iron determines the distribution strategy and seed iron content in mungbean (*Vigna radiata*) plants. Plant Soil 2020;446:413–23.
[8] Jiménez-Gómez A, García-Estévez I, García-Fraile P, Escribano-Bailón MT, Rivas R. Increase in phenolic compounds of *Coriandrum sativum* L. after the application of a *Bacillus halotolerans* biofertilizer. J Sci Food Agr 2020;100 (6):2742–9.
[9] Ravin H, Andary C, Kovacs G, Molgaard P. Caffeic acid esters as in vitro inhibitors of plant pathogenic bacteria and fungi. Biochem Syst Ecol 1989;17:175–84.
[10] Dakora FD, Phillips DA. Diverse functions of isoflavonoids in legumes transcend anti-microbial definitions of phytoalexins. Physiol Mol Plant Pathol 1996;49(1):1–20.
[11] Lattanzio V, Lattanzio VMT, Cardinali A. Role of phenolics in the resistance mechanisms of plants against fungal pathogens and insects. In: Imperato F, editor. Phytochemistry: advances in research. Kerala, India: Research Signpost; 2006. p. 23–67.
[12] Weir TL, Park SW, Vivanco JM. Biochemical and physiological mechanisms mediated by allelochemicals. Curr Opin Plant Biol 2004;7:472–9.
[13] Xuan TD, Shinkichi T, Khanh TD, Min CI. Biological control of weeds and plant pathogens in paddy rice by exploiting plant allelopathy: an overview. Crop Prot 2005;24:197–206.
[14] Boudet AM. Evolution and current status of research in phenolic compounds. Phytochem 2007;68(22–24):2722–35.
[15] Fraser CM, Chapple C. The phenylpropanoid pathway in *Arabidopsis*. Arabidopsis Book 2011. https://doi.org/10.1199/tab.0152.
[16] Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. Lignin biosynthesis and structure. Plant Physiol 2010;153(3):895–905.
[17] Davin LB, Jourdes M, Patten AM, Kim KW, Vassão DG, et al. Dissection of lignin macromolecular configuration and assembly: Comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis. Nat Prod Rep 2008;25(6):1015–90.
[18] Liu W, Feng Y, Yu S, Fan Z, Li X, et al. The flavonoid biosynthesis network in plants. Int J Mol Sci 2021;22(23):12824.
[19] Kakani VG, Reddy KR, Zhao D, Mohammed AR. Effects of ultraviolet-B radiation on cotton (*Gossypium hirsutum* L.) morphology and anatomy. Ann Bot 2003;91:817–26.
[20] Caldwell MM, Bornman JF, Ballaré CL, Flint SD, Kulandaivelu G. Terrestrial ecosystems, increased solar ultraviolet radiation, and interactions with other climate change factors. Photochem Photobiol Sci 2007;6:252–66.
[21] Fernando IPS, Lee WW, Ahn G. Marine algal flavonoids and phlorotannins; an intriguing frontier of biofunctional secondary metabolites. Crit Rev Biotechnol 2021;42(1):23–45.
[22] Dixon RA, Xie DY, Sharma SB. Proanthocyanidins: a final fronter in flavonoid research? New Phytol 2005;165:9–28.
[23] Tian L, Pang Y, Dixon R. Biosynthesis and genetic engineering of proanthocyanidins and (iso)flavonoids. Phytochem Rev 2008;7:445–65.
[24] Pourcel L, Routaboul JM, Kerhoas L, Caboche M, Lepiniec L, et al. TRANSPARENT TESTA10 encodes a laccase-like enzyme involved in oxidative polymerisation of flavonoids in *Arabidopsis* seed coat. Plant Cell 2005;17:2966–80.
[25] Molina MC, Crespo A, Vicente C, Elix JA. Differences in the composition of phenolics and fatty acids of cultured mycobiont and thallus of *Physconia distorta*. Plant Physiol Biochem 2003;41(2):175–80.
[26] Soriano G, Del Castillo-Alonso MÁ, Monforte L, Núñez-Olivera E, Martínez-Abaigar J. Phenolic compounds from different bryophyte species and cell compartments respond specifically to ultraviolet radiation, but not particularly quickly. Plant Physiol Biochem 2019;134:137–44.
[27] Pathania S, Kumar P, Singh S, Khatoon S, Rawat A, et al. Detection of ptaquiloside and quercetin in certain Indian ferns. Curr Sci 2012;102:1683–91.
[28] Baskaran X, Geo Vigila A, Zhang S, Feng S, Liao W. A review of the use of pteridophytes for treating human ailments. J Zhejiang Univ Sci 2018;19 (2):1–35.
[29] Goiris K, Muylaert K, Voorspoels S, Noten B, De Paepe D, et al. Detection of flavonoids in microalgae from different evolutionary lineages. J Phycol 2014;50(3):483–92.
[30] De Vries J, De Vries S, Slamovits CH, Rose LE, Archibald JM. How embryophytic is the biosynthesis of phenylpropanoids and their derivatives in streptophyte algae? Plant Cell Physiol 2017;58(5):934–45.
[31] Del Mondo A, Smerilli A, Ambrosino L, Albini A, Noonan DM, et al. Insights into phenolic compounds from microalgae: structural variety and complex beneficial activities from health to nutraceutics. Crit Rev Biotechnol 2021;41 (2):155–71.
[32] Albuquerque BR, Heleno SA, Oliveira B, Barros L, Ferreira ICFR. Phenolic compounds: current industrial applications, limitations and future challenges. Food Funct 2020;12:14–29.
[33] Guerin M, Huntley ME, Olaizola M. *Haematococcus* astaxanthin: applications for human health and nutrition. Trends Biotechnol 2003;21(5):210–6.
[34] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, et al. BLAST+: architecture and applications. BMC Bioinform 2009;10:421.
[35] Boutet E, Lieberherr D, Tognolli M, Schneider MM, Bansal P, et al. UniProtKB/Swiss-Prot, the manually annotated section of the UniProt Knowledge Base: how to use the entry view methods in molecular biology. Methods Mol Biol 2016;1374:23–54.
[36] Uniprot_consortium. UniProt: a hub for protein information. Nucl Acids Res 2015;43:204–12.
[37] Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4(1):9. Available: http://palaeo-electronica.org/2001_1/past/issue1_01.htm. Accessed 2021 Nov 30.
[38] Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol Biol Evoln 2013;30 (4):772–80.
[39] Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinform 2009;25(15):1972–3.
[40] Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2014;32(1):268–74.
[41] Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS ONE 2010. https://doi.org/10.1371/journal.pone.0009490.
[42] Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucl Acids Res 2021;49(1):293–6.
[43] Jones P, Binns D, Chang HY, Fraser M, Li W, et al. InterProScan 5: genome-scale protein function classification. Bioinform 2014;30(9):1236–40.
[44] Bailey TL. STREME: accurate and versatile sequence motif discovery. Bioinform 2021;37(18):2834–40.
[45] Tomitani A, Knoll AH, Cavanaugh CM, Ohno T. The evolutionary diversification of cyanobacteria: Molecular-phylogenetic and paleontological perspectives. Proc Natl Acad Sci 2006;103(14):5442–7.
[46] Garcia-Pichel F, Lombard J, Soule T, Dunaj S, Wu SH, et al. Timing the evolutionary advent of cyanobacteria and the later great oxidation event using gene phylogenies of a sunscreen. mBio, 2019, doi:10.1128/mbio.00561-19.
[47] Aydaş BS, Ozturk S, Aslim B. Phenylalanine ammonia lyase (PAL) enzyme activity and antioxidant properties of some cyanobacteria isolates. Food Chem 2013;136(1):164–9.
[48] Moffitt MC, Louie GV, Bowman ME, Pence J, Noel JP, et al. Discovery of two cyanobacterial phenylalanine ammonia lyases: kinetic and structural characterization. Biochem 2007;46(4):1004–12.
[49] Emiliani G, Fondi M, Fani R, Gribaldo S. A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land. Biol Direct 2009. https://doi.org/10.1186/1745-6150-4-7.
[50] De Vries S, Fürst-Jansen JMR, Irisarri I, Dhabalia Ashok A, Ischebeck T, et al. The evolution of the phenylpropanoid pathway entailed pronounced radiations and divergences of enzyme families. Plant J 2021;107 (4):975–1002.
[51] Peek J, Christendat D. The shikimate dehydrogenase family: functional diversity within a conserved structural and mechanistic framework. Arch Biochem Biophys 2015;566:85–99.

[52] Krämer M, Bongaerts J, Bovenberg R, Kremer S, Müller U, et al. Metabolic engineering for microbial production of shikimic acid. Metab Eng 2003;5:277–83.
[53] Lu Y, Xu J. Phytohormones in microalgae: a new opportunity for microalgal biotechnology? Trends Plant Sci 2015;20(5):273–82.
[54] Haoujar I, Cacciola F, Abrini J, Mangraviti D, Giuffrida D, et al. The contribution of carotenoids, phenolic compounds, and flavonoids to the antioxidative properties of marine microalgae isolated from Mediterranean Morocco. Molecules 2019;24(22):4037.
[55] Tan X, Zhu J, Wakisaka M. Effect of protocatechuic acid on *Euglena gracilis* growth and accumulation of metabolites. Sustainability 2020;12:9158.
[56] Schröder J, Protein sequence homology between plant 4-coumarate: CoA ligase and firefly luciferase. Nucl Acids Res, 1989;17(1):460-460.
[57] Zhou R, Jackson L, Shadle G, Nakashima J, Temple S, et al. Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. Proc Natl Acad Sci 2010;107(41):17803–8.
[58] Wang Y, Chantreau M, Sibout R, Hawkins S. Plant cell wall lignification and monolignol metabolism. Front Plant Sci 2013;9(4):220.
[59] Guo DM, Ran JH, Wang XQ. Evolution of the Cinnamyl/Sinapyl Alcohol Dehydrogenase (CAD/SAD) gene family: the emergence of real lignin is associated with the origin of Bona Fide CAD. J Mol Evol 2010;71(3):202–18.
[60] Delwiche CF, Graham LE, Thomson N. Lignin-like compounds and sporopollenin coleochaete, an algal model for land plant ancestry. Science 1989;245(4916):399–401.
[61] Dovgan I, Medvedeva E. Change in the structural elements of the lignin of the brown alga *Cystoseira barbata* at different ages. Chem Nat Compd 1983;9:85–8.
[62] Alzate-Gaviria L, Domínguez-Maldonado J, Chablé-Villacís R, Olguin-Maciel E, Leal-Bautista RM, et al. Presence of polyphenols complex aromatic "Lignin" in *Sargassum* spp. from Mexican Caribbean. J Mar Sci Eng 2021;9(1):6.
[63] Atkinson Jr AW, Gunning BE, John PC. Sporopollenin in the cell wall of *Chlorella* and other algae: Ultrastructure, chemistry, and incorporation of (14) C-acetate, studied in synchronous cultures. Planta 1972;107(1):1–32.
[64] Komaristaya VP, Gorbulin O. Sporopollenin in the composition of cell walls of *Dunaliella salina* Teod. (Chlorophyta) zygotes. Int J Algae 2006;8:43–52.
[65] He X, Dai J, Wu Q. Identification of sporopollenin as the outer layer of cell wall in microalga *Chlorella protothecoides*. Front Microbiol 2016;7:1047.
[66] Xue J-S, Zhang B, Zhan HD, Lv Y-L, Jia X-L, et al. Phenylpropanoid derivatives are essential components of sporopollenin in vascular plants. Mol Plant 2020;13(11):1644–53.
[67] Martone PT, Estevez JM, Lu F, Ruel K, Denny MW, et al. Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. Curr Biol 2009;19(2):169–75.
[68] Labeeuw L, Martone PT, Boucher Y, Case RJ. Ancient origin of the biosynthesis of lignin precursors. Biol Direct 2015. https://doi.org/10.1186/s13062-015-0052-y.
[69] Zanchetta E, Damergi E, Patel B, Borgmeyer T, Pick H, et al. Algal cellulose, production and potential use in plastics: Challenges and opportunities. Algal Res 2021;56:102288.
[70] Libralato G, Avezzù F, Volpi Ghirardini A. Lignin and tannin toxicity to *Phaeodactylum tricornutum* (Bohlin). J Hazard Mater 2011;194:435–9.
[71] Monlau F, Sambusiti C, Barakat A, Quéméneur M, Trably E, et al. Do furanic and phenolic compounds of lignocellulosic and algae biomass hydrolyzate inhibit anaerobic mixed cultures? A comprehensive review. Biotechnol Adv 2014;32(5):934–51.
[72] Mariana M, Alfatah T, Khalil AHPS, Yahya EB, Olaiya NG, et al. A current advancement on the role of lignin as sustainable reinforcement material in biopolymeric blends. J Mat Res Technol 2021;15:2287–316.
[73] Goffner D, Van Doorsselaere J, Yahiaoui N, Samaj J, Grima-Pettenati J, et al. A novel aromatic alcohol dehydrogenase in higher plants: molecular cloning and expression. Plant Mol Biol 1998;36(5):755–65.
[74] Johnson ET, Ryu S, Yi H, Shin B, Cheong H, et al. Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. Plant J 2001;25(3):325–33.
[75] Rivière C, Pawlus AD, Mérillon JM. Natural stilbenoids: distribution in the plant kingdom and chemotaxonomic interest in Vitaceae. Nat Prod Rep 2012;29(11):1317–33.
[76] Tohge T, Yonekura-Sakakibara K, Niida R, Watanabe-Takahashi A, Saito K. Phytochemical genomics in *Arabidopsis thaliana*: A case study for functional identification of flavonoid biosynthesis genes. Pure Appl Chemy 2007;79(4):811–23.
[77] Rico M, López A, Santana-Casiano JM, Gonzàlez AG, Gonzàlez-Dàvila M. Variability of the phenolic profile in the diatom *Phaeodactylum tricornutum* growing under copper and iron stress. Limnol Oceanogr 2012;58(1):144–52.
[78] Sasso S, Pohnert G, Lohr M, Mittag M, Hertweck C. Microalgae in the postgenomic era: a blooming reservoir for new natural products. FEMS Microbiol Rev 2012;36(4):761–85.
[79] López A, Rico M, Rivero A, Suárez de Tangil M. The effects of solvents on the phenolic contents and antioxidant activity of *Stypocaulon scoparium* algae extracts. Food Chem 2011;125:1104–9.
[80] Machu L, Misurcova L, Ambrozova JV, Orsavova J, Mlcek J, et al. Phenolic content and antioxidant capacity in algal food products. Molecules 2015;20(1):1118–33.
[81] Jerez-Martel I, García-Poza S, Rodríguez-Martel G, Rico M, Afonso-Olivares C, et al. Phenolic profile and antioxidant activity of crude extracts from microalgae and Cyanobacteria strains. J Food Qual 2017. https://doi.org/10.1155/2017/2924508.
[82] Blagojević D, Babić O, Rašeta M, Šibul F, Janjušević L, et al. Antioxidant activity and phenolic profile in filamentous cyanobacteria: the impact of nitrogen. J Appl Phycol 2018;30(4):2337–46.
[83] Santangelo R, Silvestrini A, Mancuso C. Ginsenosides, catechins, quercetin and gut microbiota: current evidence of challenging interactions. Food Chem Toxicol 2019;123:42–9.
[84] Campanella JJ, Smalley JV, Dempsey ME. A phylogenetic examination of the primary anthocyanin production pathway of the Plantae. Bot Stud 2014. https://doi.org/10.1186/1999-3110-55-10.
[85] Min T, Kasahara H, Bedgar DL, Youn B, Lawrence PK, et al. Crystal structures of pinoresinol-lariciresinol and phenylcoumaran benzylic ether reductases and their relationship to isoflavone reductases. J Biol Chem 2003;278:50714–23.
[86] Caputi L, Malnoy M, Goremykin V, Nikiforova S, Martens S. A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. Plant J 2011;69(6):1030–42.
[87] Piatkowski BT, Imwattana K, Tripp EA, Weston DJ, Healey A, et al. Phylogenomics reveals convergent evolution of red-violet coloration in land plants and the origins of the anthocyanin biosynthetic pathway. Mol Phylogenet Evol 2020. https://doi.org/10.1016/j.ympev.2020.106904.
[88] Tanna B, Choudhary B, Mishra A. Metabolite profiling, antioxidant, scavenging and anti-proliferative activities of selected tropical green seaweeds reveal the nutraceutical potential of *Caulerpa* spp. Algal Res 2018;36:96–105.
[89] Wen W, Alseekh S, Fernie AR. Conservation and diversification of flavonoid metabolism in the plant kingdom. Curr Opin Plant Biol 2020;55:100–8.
[90] Fujita Y, Hattori A. Occurrence of a purple bile pigment in phycoerythrin-rich cells of the blue-green alga, *Tolypothrix tenuis*. J Gen Appl Microbiol 1963;9(2):253–6.
[91] Alston RE. An investigation of the purple vacuolar pigment of *Zygogonium ericetorum* and the status of "algal anthocyanins" and "phycoporphyrins". Am J Bot 1958;45(9):688–92.
[92] Aigner S, Remias D, Karsten U, Holzinger A. Unusual phenolic compounds contribute to ecophysiological performance in the purple-colored green alga *Zygogonium ericetorum* (Zygnematophyceae, Streptophyta) from a high-alpine habitat. J Phycol 2013;49(4):648–60.
[93] Pouvreau JB, Morançais M, Fleury F, Rosa P, Thion L, et al. Preliminary characterisation of the blue-green pigment "marennine" from the marine tychopelagic diatom *Haslea ostrearia* (Gaillon/Bory) Simonsen. J Appl Phycol 2006;18(6):757–67.
[94] Gastineau R, Turcotte F, Pouvreau JB, Morançais M, Fleurence J, et al. Marennine, promising blue pigments from a widespread *Haslea* diatom species complex. Mar Drugs 2014;12(6):3161–89.
[95] Lima SAC, Castro PML, Morais R. Biodegradation of p-nitrophenol by microalgae. J Appl Phycol 2003;15(2/3):137–42.
[96] Kleckner V, Kosaric N. Degradation of phenols by algae. Environ Technol 1992;13:493–501.
[97] Palanisami S, Saha SK, Lakshmanan U. Laccase and polyphenol oxidase activities of marine cyanobacteria: a study with Poly R-478 decolourization. World J Microbiol Biotechnol 2009;26(1):63–9.
[98] Afreen S, Shamsi TN, Baig MA, Ahmad N, Fatima S, et al. A novel multicopper oxidase (laccase) from cyanobacteria: Purification, characterization with potential in the decolorization of anthraquinonic dye. PLoS ONE 2017. https://doi.org/10.1371/journal.pone.0175144.
[99] Asadgol Z, Forootanfar H, Rezaei S, Mahvi AH, Faramarzi MA. Removal of phenol and bisphenol—A catalyzed by laccase in aqueous solution. J Environ Health Sci Eng 2014. https://doi.org/10.1186/2052-336X-12-93.
[100] Otto B, Schlosser D. First laccase in green algae: purification and characterization of an extracellular phenol oxidase from *Tetracystis aeria*. Planta 2013;240(6):1225–36.
[101] Otto B, Beuchel C, Liers C, Reisser W, Harms H, et al. Laccase-like enzyme activities from chlorophycean green algae with potential for bioconversion of phenolic pollutants. FEMS Microbiol Lett 2015;362(11):fnv072.
[102] Diaz JM, Plummer S, Tomas C, Alves-de-Souza C. Production of extracellular superoxide and hydrogen peroxide by five marine species of harmful bloom-forming algae. J Plankt Res 2018;40(6):667–77.
[103] Gross EM. Allelopathy of aquatic autotrophs. Crit Rev Plant Sci 2003;22(3–4):313–39.
[104] Wang J, Zhu J, Liu S, Liu B, Gao Y, et al. Generation of reactive oxygen species in cyanobacteria and green algae induced by allelochemicals of submerged macrophytes. Chemosphere 2011;85(6):977–82.
[105] Huang H, Xiao X, Ghadouani A, Wu J, Nie Z, et al. Effects of natural flavonoids on photosynthetic activity and cell integrity in *Microcystis aeruginosa*. Toxins 2015;7(1):66–80.
[106] Żyszka B, Anioł M, Lipok J. Modulation of the growth and metabolic response of cyanobacteria by the multifaceted activity of naringenin. PLoS ONE 2017. https://doi.org/10.1371/journal.pone.0177631.