

RESEARCH

Open Access



A key gene, *violaxanthin de-epoxidase-like 1*, enhances fucoxanthin accumulation in *Phaeodactylum tricornutum*

Chenjie Li^{1,2}, Yufang Pan¹, Wenxiu Yin¹, Jin Liu³ and Hanhua Hu^{1*}

Abstract

Background Fucoxanthin has been widely investigated owing to its beneficial biological properties, and the model diatom *Phaeodactylum tricornutum*, possessing fucoxanthin (Fux) chlorophyll proteins as light-harvesting systems, is considered to have the potential to become a commercial cell factory for the pigment production.

Results Here, we compared the pigment contents in 10 different *P. tricornutum* strains from the globe, and found that strain CCMP631 (Pt6) exhibited the highest Fux content but with a low biomass. Comparison of mRNA levels revealed that higher Fux content in Pt6 was related with the higher expression of gene *violaxanthin de-epoxidase-like (VDL) protein 1 (VDL1)*, which encodes the enzyme catalyzing the tautomerization of violaxanthin to neoxanthin in Fux biosynthesis pathway. Single nucleotide variants of *VDL1* gene and allele-specific expression in strains Pt1 (the whole genome sequenced strain CCMP632) and Pt6 were analyzed, and overexpressing of each of the 4 *VDL1* alleles, two from Pt1 and two from Pt6, in strain Pt1 leads to an increase in downstream product diadinoxanthin and channels the pigments towards Fux biosynthesis. All the 8 *VDL1* overexpression (OE) lines showed significant increases by 8.2 to 41.7% in Fux content without compromising growth, and *VDL1* Allele 2 OE lines even exhibited the higher cell density on day 8, with an increase by 24.2–28.7% in two *Pt1VDL1-allele 2* OE lines and 7.1–11.1% in two *Pt6VDL1-allele 2* OE lines, respectively.

Conclusions The results reveal *VDL1*, localized in the plastid stroma, plays a key role in Fux over-accumulation in *P. tricornutum*. Overexpressing *VDL1*, especially *allele 2*, improved both the Fux content and growth rate, which provides a new strategy for the manipulation of Fux production in the future.

Keywords Diadinoxanthin, Fucoxanthin, Metabolic engineering, *Phaeodactylum tricornutum*, Strain screening, Violaxanthin de-epoxidase-like 1

Background

Diatoms account for approximately 40% of the annual marine biomass production [1, 2]. Diatoms possess fucoxanthin chlorophyll proteins (FCP) as light-harvesting systems, and fucoxanthin (Fux) allows the organisms to capture blue-green light under deep-sea environments, thus enabling diatoms to thrive in the modern oceans. In addition, Fux has been widely investigated owing to its beneficial biological properties such as anti-oxidant, anti-obesity, anti-diabetic, anticancer, and anti-microbial activities [3–8]. *Phaeodactylum tricornutum*,

*Correspondence:

Hanhua Hu
hanhuahu@ihb.ac.cn

¹ Key Laboratory of Algal Biology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Key Laboratory of Poyang Lake Environment and Resource Utilization, Ministry of Education, and Center for Algae Innovation & Engineering Research, School of Resources and Environment, Nanchang University, Nanchang 330031, China



© The Author(s) 2024. **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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

a model species of diatoms with the well-characterized genome and abundant gene editing tools, is considered to have potential to become a commercial cell factory for Fux production [9].

Although the biosynthesis pathway of Fux has not been fully elucidated yet, several enzymes involved in steps from violaxanthin to Fux have been identified in *P. tricornutum*. In the first step, violaxanthin de-epoxidase-like (VDL) protein 1 (VDL1) catalyzes the tautomerization of violaxanthin to neoxanthin in tobacco leaves and in vitro [10]. VDL2 catalyzes the conversion of diadinoxanthin (Ddx) to allenoxanthin, and zeaxanthin epoxidase 1 (ZEP1) performs the epoxidation of haptoxanthin to phaneroxanthin [11]. The final step of the pathway was catalyzed by carotenoid isomerase-like protein 5 (CRTISO5) [12]. In addition, several enzymes involved in the early pathway steps of carotenoid biosynthesis have also been identified in *P. tricornutum*. For example, only ZEP3 functions as a zeaxanthin epoxidase which catalyzes the conversion of zeaxanthin to violaxanthin, and ZEP2 most likely converts diatoxanthin (Dtx) to Ddx in the xanthophyll cycle of *P. tricornutum* [13]. Violaxanthin de-epoxidase (VDE) catalyzes the conversion of Ddx into Dtx [14]. Furthermore, roles of genes encoding 1-deoxy-D-xylulose 5-phosphate synthase (DXS), phytoene synthase (PSY), phytoene desaturase (PDS), ζ -carotene desaturase (ZDS), and lycopene β -cyclase (LCYB) in carotenoid/Fux accumulation have been revealed [15–18].

P. tricornutum is generally considered to have the limited ecological relevance, and 10 accessions belonging to four genotypes were isolated at several locations worldwide [19, 20]. These *P. tricornutum* strains showed different growth rates, cell morphs and sizes, triacylglycerol contents, and temperature adaptability [19, 21–23], and their non-photochemical quenching (NPQ) capacity was also different more or less [24]. NPQ is associated with the xanthophyll cycle which involved the de-epoxidation of Ddx to absorb excess excitation energy generated during photosynthesis and dissipate it as heat in diatoms [25]. We compared the pigment contents in the 10 *P. tricornutum* strains, and strain CCMP631 (Pt6) exhibited the highest Fux content but with a low biomass. To interpret the over-accumulation of Fux in Pt6, a comparison of the transcript levels of genes involved in the Fux biosynthesis were performed between strains Pt1 (the whole genome sequenced strain CCMP632) and Pt6, and *VDL1* was found to keep a higher expression level till the stationary growth phase. In this study, single nucleotide variants (SNVs) of *VDL1* gene and allele-specific expression (ASE) in Pt1 and Pt6 were analyzed, the subcellular localization of *VDL1* was determined,

and four *VDL1* alleles from Pt1 and Pt6 were over-expressed, respectively, in strain Pt1. We obtained an engineered strain that possessed a Fux content comparable with that of strain Pt6 and higher biomass, which has potential application in Fux production.

Results

Screening of the Fux-enriched *P. tricornutum* strain

Since 1897, ten *P. tricornutum* ecotypes have been isolated from nine different geographic locations around the world with different habitats including sea shores, estuaries, rock pools, and tidal creeks [19]. We compared their growth and pigment content under the nitrogen-replete condition. Obviously, strains Pt6, Pt7 and Pt8, which belong to the same genetic clade [20], possessed relatively higher Fux content among the ten *P. tricornutum* strains (Additional file 1: Fig. S1). Strain Pt6, in particular, had the highest Fux content (Fig. 1a). After the initial screening, dry biomass and Fux content in Pt6 were further compared with those of strain Pt1, the whole genome sequenced strain [26]. The dry biomass on day 4, 6 and

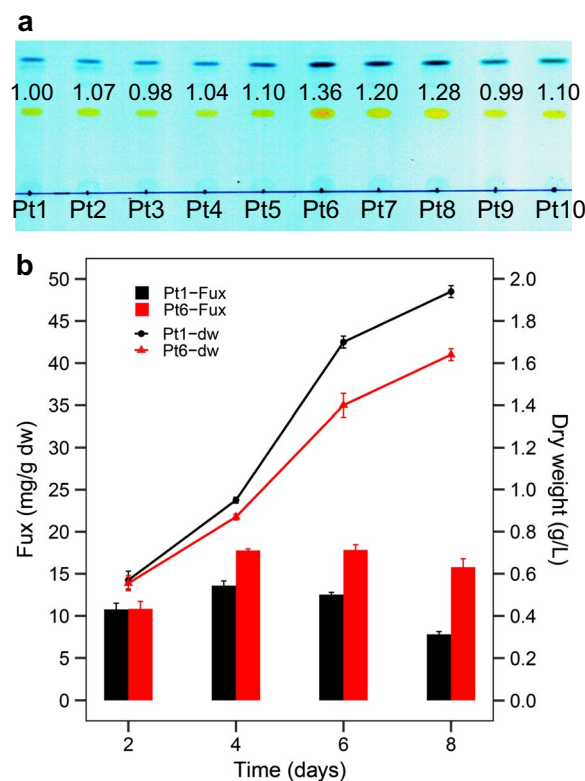


Fig. 1 Fucoxanthin content in different accessions of *P. tricornutum*. **a** Comparison of fucoxanthin (Fux) content in *P. tricornutum* strains Pt1 ~Pt10 by thin-layer chromatography. **b** Comparison of dry weight (dw) and Fux content in *P. tricornutum* strains CCMP2561 (Pt1) and CCMP631 (Pt6) cultivated in enriched nitrogen and phosphate medium

8 was 9%, 21%, and 18% higher in Pt1, respectively, but the corresponding Fux content was 31%, 42% and 102% higher in Pt6 (Fig. 1b).

To reveal why Fux content was higher in Pt6, transcript levels of genes involved in carotenoid biosynthesis in wild-type *P. tricornutum* strains Pt1 and Pt6 were compared (Fig. 2a). Most of the genes showed higher mRNA levels in Pt6 throughout the whole growth phase (Fig. 2b), especially gene *VDL1*, whose mRNA level was 1.1 ~ 8.54-fold higher from 48 to 156 h (Fig. 2c).

Sequence comparison of VDL1 in strains Pt1 and Pt6

Single nucleotide polymorphisms (SNPs) had been found in ten wild-type *P. tricornutum* strains Pt1–Pt10 [20], and allelic variation and allele-specific expression had also been observed in this diatom. According to the SNPs and INDELS of *VDL1*, the 10 strains were divided into four clades (Additional file 1: Fig. S2), consistent with the four genetic clades based on the genome resequencing data [20]. The coding region of Pt6 *VDL1* alleles (allele 1 and allele 2) boasted 11 SNVs (Additional file 1: Fig. S3), resulting in 6 amino acid differences between the encoded proteins (Additional file 1: Fig. S4). Pt1 *VDL1* allele 1 had an insertion of four nucleotides at position 138 before the translation initiation site, which might result in a shift of the translation initiation site, and the presence of additional 68 amino acids in its protein sequence (Additional file 1: Fig. S4) as annotated in JGI v2.0 (<https://mycocosm.jgi.doe.gov/Phatr2>) and Ensembl (https://protists.ensembl.org/Phaeodactylum_tricornutum). However, the sequence from -131 to

-203 bp in the upstream of *VDL1* allele 1 was not transcribed according to the analysis of published RNA-seq data (Additional file 1: Fig. S5; Additional file 2: Table S1). In addition, 15 SNVs was detected in the coding region of Pt1 *VDL1* alleles 1 and 2 (Additional file 1: Fig. S3). Although the amino acid differences were found in the two alleles encoding *VDL1* sequences of strains Pt1 and Pt6, predicted structures of the four *VDL1* were very similar (Additional file 1: Fig. S6). To detect if the transcript of *VDL1* was affected by ASE, the published RNA-seq data of *P. tricornutum* strain Pt1 were analyzed. The results showed that transcript of *VDL1* increased gradually from 1 to 228 h during the whole cultivation period, but transcripts of alleles 1 and 2 showed no significant difference in response to nitrogen stress and culture time (Additional file 1: Fig. S7; Additional file 2: Table S2), and the average ratio of RPKM values of allele 1 versus those of allele 2 was more or less than 1. Seemingly, the higher Fux content in Pt6 might not be attributed to the variation of sequence and ASE in gene *VDL1*.

Phylogeny and localization of *P. tricornutum* VDL1

To reveal the phylogenetic relationship of VDL proteins, 26 protein sequences were obtained by BLAST search against the NCBI protein database and a maximum likelihood tree was constructed. *VDL1* sequences from the class Bacillariophyta (diatoms) formed a large clade, which was most closely related to the class Phaeophyta (brown algae). Diatom *VDL1* was clearly distinct from those of classes Chromista (golden algae), Pyrrophyta (dinoflagellates), and Haptophyta (haptophytes) (Fig. 3).

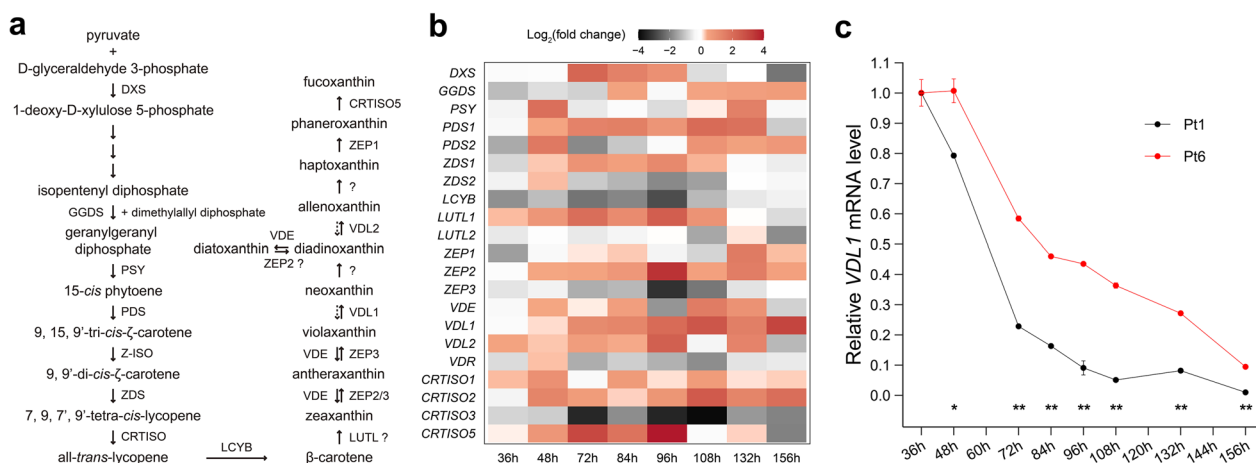


Fig. 2 Transcript levels of genes encoding components involved in carotenoid biosynthesis in wild-type *P. tricornutum*. **a** Genes involved in fucoxanthin biosynthesis. **b** Transcriptional fold changes in strain CCMP631 (Pt6) relative to strain CCMP2561 (Pt1). **c** Relative *VDL1* mRNA levels in strains Pt1 and Pt6. Abbreviations: DXS, 1-deoxy-D-xylulose 5-phosphate synthase; PSY, phytoene synthase; GGDS, geranylgeranyl diphosphate synthase; PDS, phytoene desaturase; ZDS, ζ-carotene desaturase; LCYB, lycopene β-cyclase; LUTL, lutein deficient-like; ZEP, zeaxanthin epoxidase; VDE, violaxanthin de-epoxidase; VDL, violaxanthin de-epoxidase-like; VDR, VDE-related; CRTISO5, carotenoid isomerase 5. * and ** indicate statistically significant differences of $P < 0.05$ and $P < 0.01$, respectively

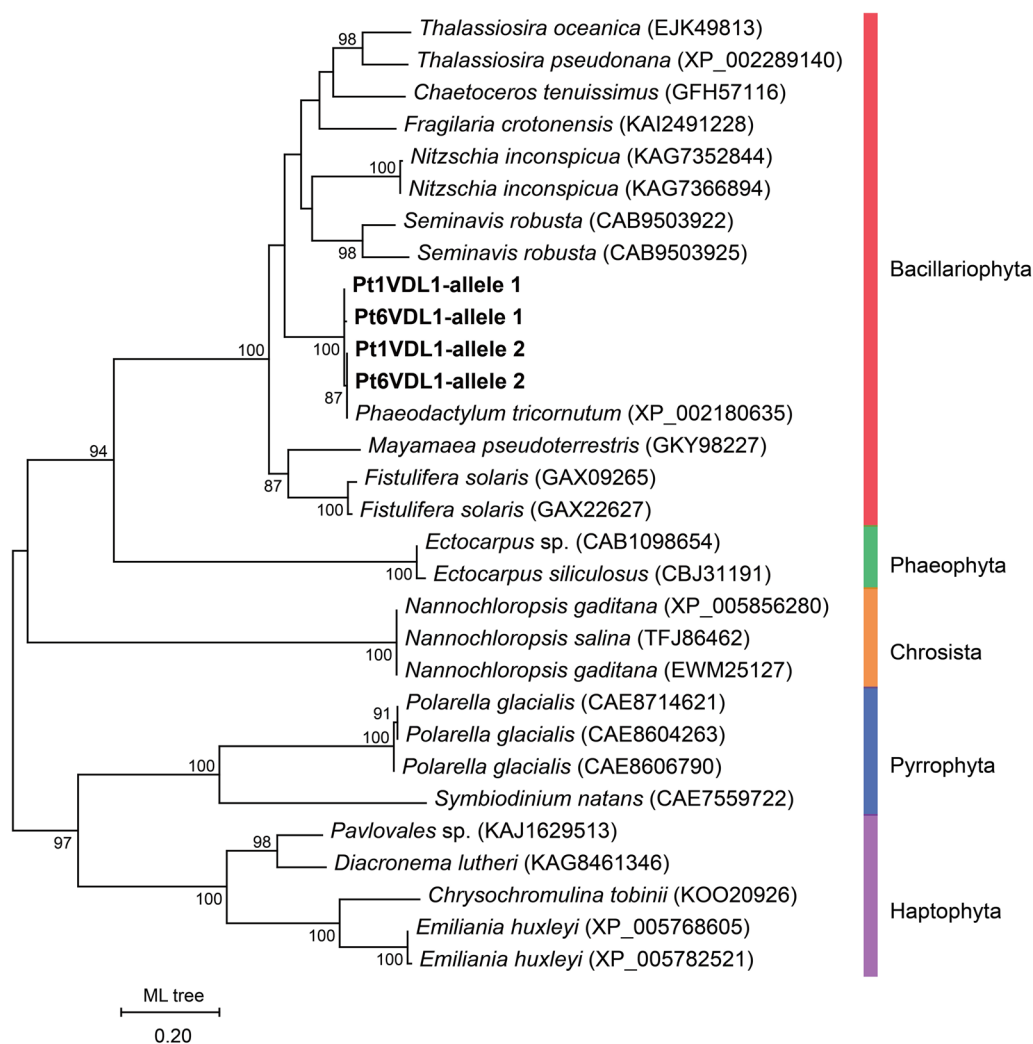


Fig. 3 Phylogenetic relationships of VDL1 proteins from different species

The similarity between *P. tricornutum* VDL1 and its homologue from other diatom species was above 60%, while the similarity of *P. tricornutum* VDL1 with those of species from the other four classes was below 50% (Additional file 2: Table S3).

SignalP 6.0 and iPSORT predicted that Pt1 *VDL1* allele 1 encoded sequence contained no signal peptide or plastid targeting sequence. TargetP and HECTAR also predicted a low probability of a signal peptide in it (0.0011 and 0.1373, respectively), but Pt1 *VDL1* allele 2 and two Pt6 *VDL1* alleles encoded sequences were predicted to have a signal peptide (Additional file 2: Table S4). However, when the four *VDL1* alleles were fused with eGFP, respectively, and expressed in wild-type Pt1 cells, the fluorescence signal of eGFP (green) overlapped with the chlorophyll fluorescence (red) for all four fusion constructs, indicating the plastid localization of VDL1. In

contrast, only chlorophyll fluorescence was observed in the wild-type control cells (Fig. 4).

***VDL1* overexpression increased Fux accumulation without compromising growth**

To overexpress *VDL1* in Pt1, four constructs containing sequence of *Pt1VDL1-allele 1*, *Pt1VDL1-allele 2*, *Pt6VDL1-allele 1*, and *Pt6VDL1-allele 2* were generated, respectively, and then were transferred into Pt1. Two overexpression (OE) lines were selected for each construct, and the relative mRNA expression levels of *VDL1* in all selected OE lines were significantly upregulated (3.2~25.0 fold) compared with that of Pt1, comparable to or even higher than that of Pt6 (Fig. 5a). However, the transcripts of *VDL2* and *VDE*, two genes with potential similarity in functions, changed hardly or even downregulated compared with wild-type Pt1 in OE lines (Fig. 5a). During the 8-day batch cultivation, the cell density of all

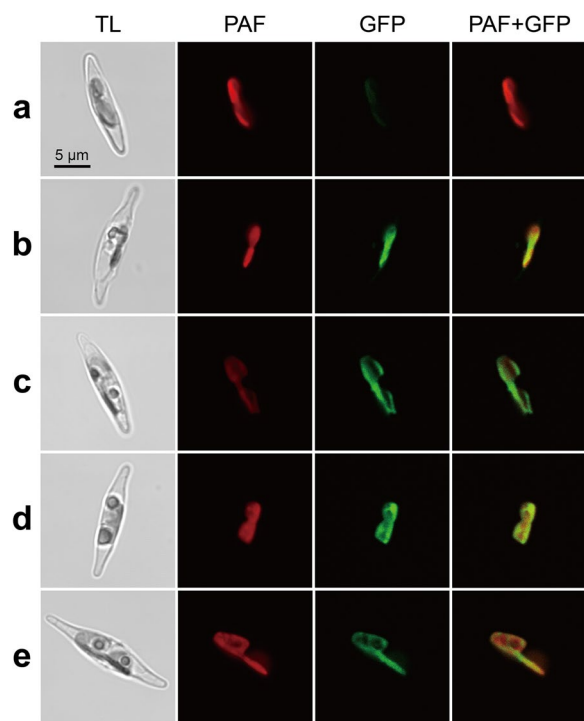


Fig. 4 Fluorescent microscope images of cells expressing VDL1-eGFP. Panels show microscopical images of transmitted light (TL), chlorophyll autofluorescence (red, PAF), and GFP fluorescence (green), together with merged images (PAF + GFP) of chlorophyll and GFP fluorescence from left to right. **a** Wild-type Pt1. **b** Pt1VDL1-allele 1::eGFP. **c** Pt1VDL1-allele 2::eGFP. **d** Pt1VDL6-allele 1::eGFP. **e** Pt1VDL6-allele 2::eGFP

OE lines was close to or higher than that of Pt1, especially on day 8, when Pt1VDL1-allele 2-OE1 had the highest cell density with a 28.7% increase relative to Pt1.

In contrast, Pt6 had a 24.7% lower cell density than Pt1 on day 8 (Fig. 5b). There was no significant difference in the maximum photosynthetic efficiency between wild-type Pt1 and all OE lines (Fig. 5c).

Under optimized culture conditions, Fux contents of all VDL1 OE lines were higher than that of Pt1, but were slightly lower than that of Pt6 by thin-layer chromatography (TLC) analysis (Fig. 6a). High-performance liquid chromatography (HPLC) analysis showed that the relative content of Fux, β -carotene, and Ddx in all VDL1 OE lines significantly increased by 8.2% to 41.7%, 6.6% to 32.6%, and 4.8% to 35.0%, respectively. In all OE lines, Pt1VDL1-allele 2-OE1 had the highest Fux content of 10.15 mg/g dry biomass, 42% higher than that of Pt1. All OE lines except two (Pt1VDL1-allele 2-OE3 and Pt6VDL1-allele 1-OE15) showed a significant increase of relative chlorophyll *a* content from 13.6% to 36.2%. However, strain Pt6 possessed the highest Fux content, reaching 11.57 mg/g dry biomass, which was 61.6% higher compared to Pt1 and 14% higher than that of Pt1VDL1-allele 2-OE1 line (an OE line with the highest Fux). Similarly, the highest relative content of chlorophyll *a* and β -carotene was also observed in Pt6 (Fig. 6b; Additional file 2: Table S5). Since the ratio of fucoxanthin to chlorophyll in *P. tricornutum* fucoxanthin chlorophyll protein (FCP) complex remains constant, it is reasonable that the chlorophyll content increases with the increasing fucoxanthin.

There were no significant differences in Dtx content between Pt1 and all OE lines, and the de-epoxidation state (DES) determined as $Dtx/(Dtx + Ddx)$ was about 0.08 ~ 0.11, with significant ($P < 0.01$) decreases observed only in Pt1VDL1-allele1-OE5 line (Fig. 7a). In addition, relative level of the pool of xanthophyll (Dtx + Ddx) normalized to the Fux significant decreased in all OE lines

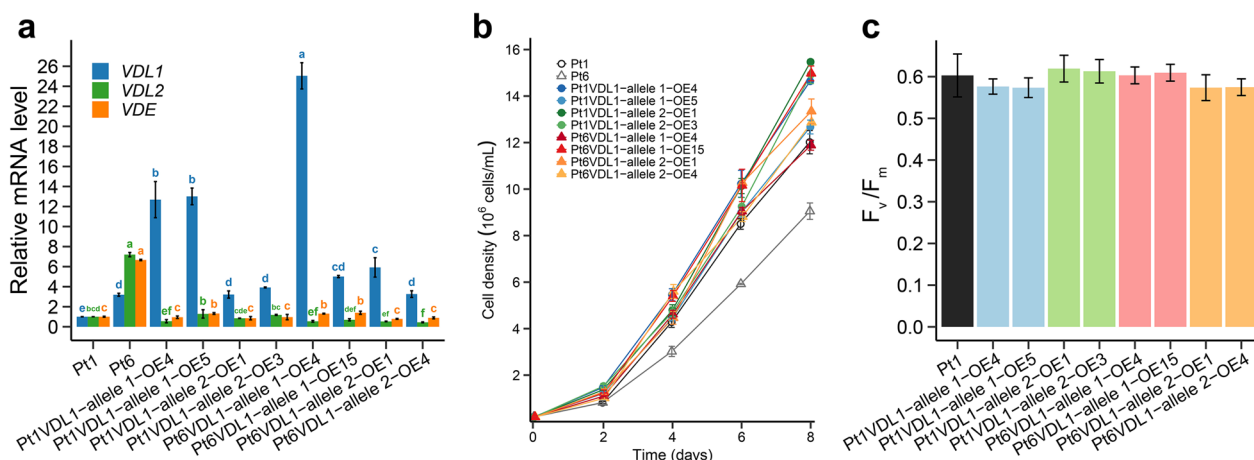


Fig. 5 Relative mRNA levels of VDL1, VDL2 and VDE detected by real-time quantitative PCR **a**, growth **b**, and photosynthetic efficiency (F_v/F_m) **c** of wild-type (Pt1 and/or Pt6) and VDL1 overexpression (OE) strains. Allele 1 and allele 2 represent the VDL1 alleles 1 and 2, respectively. Different letters above the bars indicate statistically significant differences ($P < 0.05$)

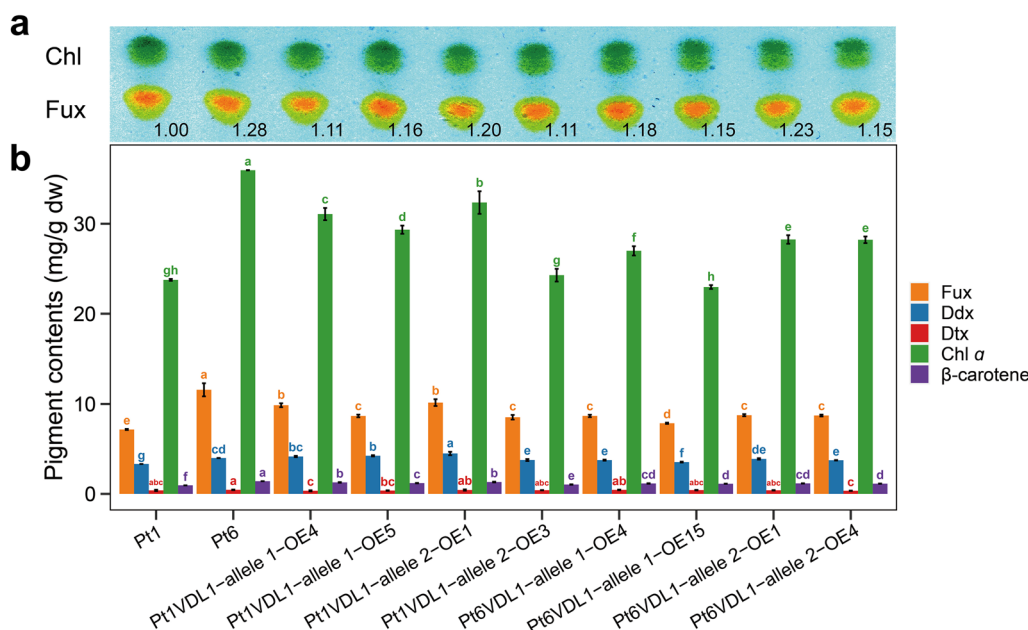


Fig. 6 Pigment contents of wild-types Pt1 and Pt6, and eight *VDL1* overexpression (OE) strains. Allele 1 and allele 2 represent the *VDL1* alleles 1 and 2, respectively. **a** Thin-layer chromatography (TLC) analysis. The values above the TLC panel indicate the relative pigment content normalized to Pt1, which was set as 1. **b** HPLC analysis. Fux: fucoxanthin, Ddx: diadinoxanthin, Dtx: diatoxanthin, Chl a: chlorophyll a, β-carotene: β-carotene. Different letters above the bars indicate statistically significant differences ($P < 0.05$)

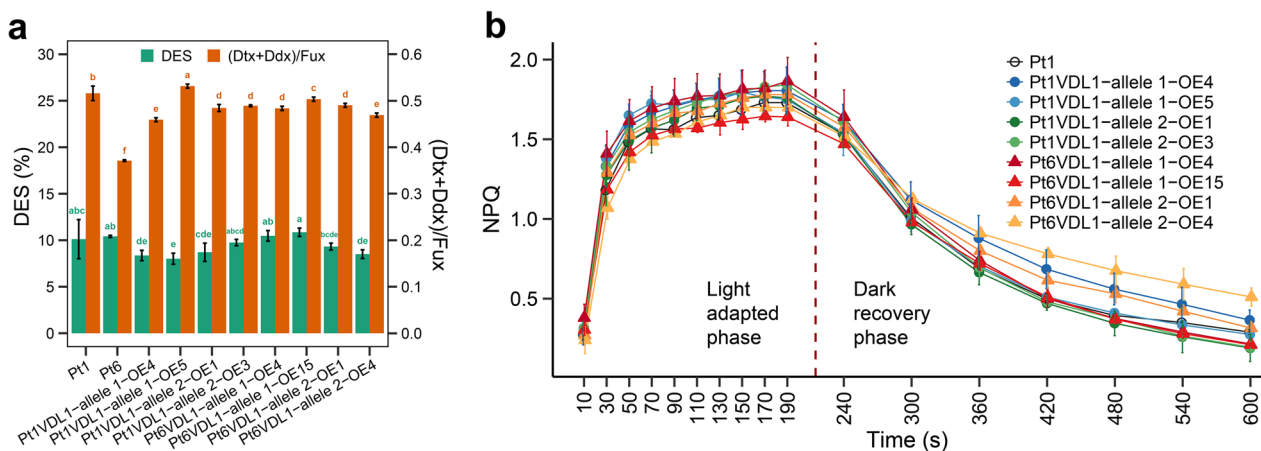


Fig. 7 De-epoxidation state (DES) determined as $Dtx/(Dtx + Ddx)$, relative level of the pool of xanthophyll ($Dtx + Ddx$) normalized to the fucoxanthin (Fux) content **a**, non-photochemical quenching (NPQ) kinetics **b** of wild-type (Pt1 and/or Pt6) and eight *VDL1* overexpression (OE) strains grown under low light ($60 \mu\text{mol photons m}^{-2}\text{s}^{-1}$). Allele 1 and allele 2 represent the *VDL1* alleles 1 and 2, respectively. Ddx: diadinoxanthin, Dtx: diatoxanthin. Different letters above the bars indicate statistically significant differences ($P < 0.01$)

except Pt1VDL1-allele 1-OE5 line. These results indicated that *VDL1* overexpression enhanced the synthesis of Ddx, the precursor of Fux and Dtx biosynthesis, and then Ddx flux mainly went to the Fux biosynthesis. Non-photochemical quenching kinetic profiles, mainly believed to be related with Dtx content, showed no significant changes in all *VDL1* OE strains and Pt1 (Fig. 7b) at low light, which also suggested *VDL1* overexpression

had no effect on the synthesis of Dtx. However, the relaxation of NPQ got slow in strains Pt1VDL1-allele 1-OE4, Pt6VDL1-allele 2-OE4 and Pt6VDL1-allele 2-OE15, while the relaxation of NPQ in the other strains showed no significant difference to that of WT.

To further interpret the physiological roles of *VDL1* overexpression in *P. tricornutum*, growth, Fux content and NPQ at extremely high light was investigated.

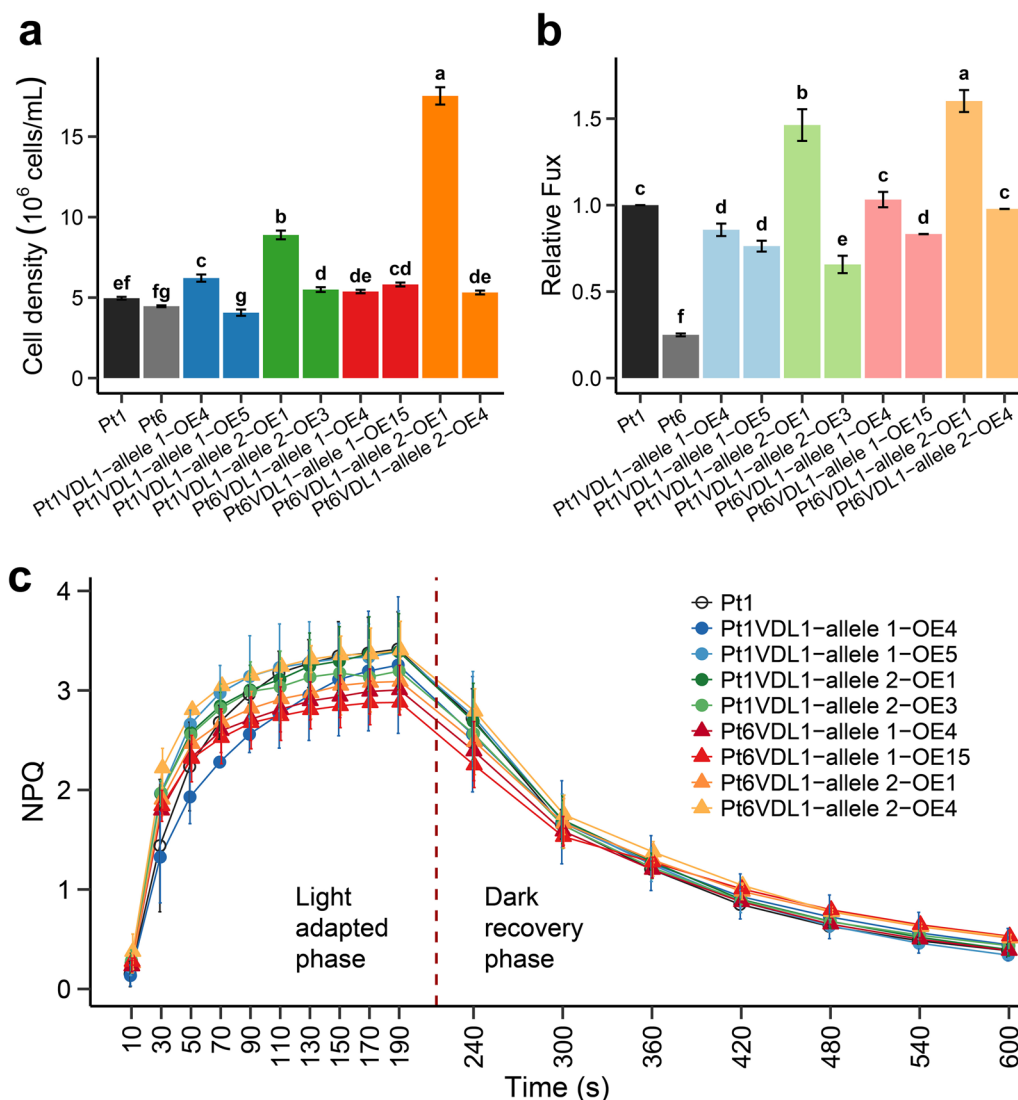


Fig. 8 Cell density **a**, relative fucoxanthin (Fux) content at day 6 **b**, non-photochemical quenching (NPQ) kinetics **c** of wild-type (Pt1 and/or Pt6) and eight *VDL1* overexpression (OE) strains grown under high light ($1000 \mu\text{mol photons m}^{-2} \text{s}^{-1}$). Allele 1 and allele 2 represent the *VDL1* alleles 1 and 2, respectively. Different letters above the bars indicate statistically significant differences ($P < 0.05$)

Under light irradiation conditions of $1000 \mu\text{mol photons m}^{-2} \text{s}^{-1}$, the occurrence of higher cell density (Fig. 8a) and fucoxanthin content (Fig. 8b) was only observed in transformants Pt1VDL1-allele 2-OE1 and Pt6VDL1-allele 2-OE1. However, the NPQ of the transformants did not show significant differences compared to the wild-type (Fig. 8c), though the NPQ of both the wild-type and transformants were significantly higher than that at low light (Fig. 7b). The results indicated that the increased fucoxanthin in *P. tricornutum* was favorable to the cell growth under high light condition, but a sole increase of fucoxanthin content would not affect NPQ.

Discussion

According to the sequence analysis of *VDL1* allele genes from *P. tricornutum* accessions Pt1 and Pt6, an insertion of four bases was found in the sequence of *Pt1VDL1-allele 1* (Additional file 1: Fig. S3). Such an insertion was also observed in accessions Pt2, Pt3, Pt4, and Pt9 (Additional file 2: Table S6), suggesting a common origin of the mutation at this site. The phylogenetic tree constructed based on the *VDL1* gene polymorphism represented 10 *P. tricornutum* accessions arranged in four different clades, consistent with the clade classification based on whole-genome variations in Pt1-10. It seemed that *VDL1* was not a conserved gene and its mutation frequency

was comparable to the overall mutation frequency of the species [20]. Distinct difference of Fux content was observed in Clades A to D, and Clade D (Pt6, Pt7, and Pt8) possessed the higher Fux content. Apparently, *VDL1* gene is the key gene for the regulation of Fux biosynthesis. Indeed, *VDL1* relative mRNA level was higher in Pt6 compared with Pt1, especially after exponential growth phase, which might contribute to the higher Fux content in Pt6. Most *cis*-regulatory elements located in the two *Pt6VDL1* alleles promoters were identical to those in the two *Pt1VDL1* alleles promoters (Additional file 2: Table S7), except an extra MYBHv1 binding site (CCAAT box) and an enhancer-like element involved in anoxic specific inducibility (GC-motif) present in *Pt6VDL1-allele 1* promoter (Additional file 1: Fig. S8). The differences might result in the higher expression level of *VDL1* in strain Pt6.

P. tricornutum *VDL1* catalyzes the tautomerization of violaxanthin to neoxanthin in tobacco leaves and in vitro [10]. Neoxanthin is the precursor not only of the major light-harvesting carotenoid Fux but also of Ddx involved in photoprotection. The Ddx content in Pt6, a natural strain with higher *VDL1* mRNA level, was higher than that in Pt1. Overexpressing of each of the 4 *VDL1* alleles, two from Pt1 and two from Pt6, in strain Pt1 virtually leads to an increase in downstream product Ddx. However, there were no significant differences in the relative content of Dtx, which is directly related to photoprotection, among all OE strains. These results indicate that overexpression of *VDL1* only increases the total amount of pigments in the Ddx cycle pool, which is likely to enhance the photoprotection ability of the OE strains under strong light without affecting the transient NPQ capacity [27, 28]. Under unstressed conditions (F_v/F_m ratio was close to 0.60), there was no significant difference in NPQ between the OE strains and Pt1, which further confirms that Dtx and DES were hardly affected by *VDL1* overexpression.

Although the relative content of Ddx+Dtx and that of Fux increased, the ratio of (Ddx+Dtx)/Fux decreased in 7 out of the 8 *VDL1* OE lines, which suggests that *VDL1* overexpression channels the pigments in the Ddx cycle towards Fux biosynthesis. It should be noted that an excellent characteristic to promote the total pool of pigments in the Ddx cycle qualifies *VDL1* OE strains as cell factories, for a stronger photoprotection ability may endow the lines with a greater advantage in Fux accumulation under high light intensity and prolonged illumination [29].

Proteomes revealed that carotenoid biosynthesis occurred in the plastid of *P. tricornutum*, and several key enzymes involved in Fux biosynthesis, including CRTISO5, ZEP1, ZEP2, ZEP3, VDE, and LUTL, were

present in the purified plastid [30]. Although *P. tricornutum* *VDL1* was not a predicted plastid protein by ASAFind, it was an assigned plastid protein by blasting against Arabidopsis proteome database in PPDB [30]. GFP fusion experiments in our study showed that *VDL1* was localized in *P. tricornutum* plastid stroma. Therefore it is safe to conclude that the biosynthesis of Fux from β -carotene takes place in the plastid. Furthermore, targeting of PtV*VDL1* to the thylakoid lumen rather than plastid stroma led to higher activity of the enzyme in tobacco leaves [10].

Although Pt6 possesses higher Fux content, its growth rate and cell density were lower compared with Pt1 under standard *f/2* nutrient condition. Overexpressing *VDL1* from Pt6 or Pt1 in Pt1 enhanced the Fux accumulation, but did not decrease the growth rate and cell density. In fact, on day 8 the cell density increased by 24.2–28.7% in two *Pt1VDL1-allele 2* OE lines, and 7.1–11.1% in two *Pt6VDL1-allele 2* OE lines, respectively. These results showed that elevated Fux would not affect the growth, and the low biomass in Pt6 should not be attributed to the higher *VDL1* transcript levels. In addition, most of the *VDL1* OE lines showed a significant increase in chlorophyll *a* and all OE lines showed an increase in β -carotene compared to wild-type Pt1. Chlorophyll *a* is involved in the formation of FCP and β -carotene is an important component of the photosystems. Increased relative content of FCP components including Fux, Ddx, and chlorophyll *a*, gave rise to the coordinated increase of other FCP components, thereby increasing the overall content of FCP without improving the photosynthetic efficiency (Fig. 5c). However, changes in photosystem components could ultimately stimulate the growth of the *VDL1* OE strains [29, 31].

After an extensive screening and comparison of different Pt strains from the globe, strain Pt6 was found to have the highest Fux content but slower growth, while Pt1 exhibited the opposite characteristics. Analysis of the transcript levels of genes involved in carotenoid biosynthesis pathway revealed that higher expression level of *VDL1* in Pt6 resulted in the higher Fux content. Overexpressing *VDL1* in Pt1 improved both Fux content and growth rate, and the genetically engineered strain has the potential cell factory to yield Fux.

Methods

Strains and growth conditions

Axenic cultures of *P. tricornutum* strains CCMP632 (Pt1) and CCMP631 (Pt6) were gifts from Chris Bowler and originated from the Provasoli-Guillard National Center for Marine Algae and Microbiota at Bigelow Laboratory for Ocean Sciences, USA. In batch static culture, *P. tricornutum* wild-types and the transformants were cultured in

artificial seawater containing $f/2$ nutrient additions [32] with an initial cell density of 2×10^5 cells/ml. The cultures were maintained under continuous light of $60 \mu\text{mol photons}/(\text{m}^2 \cdot \text{s})$ at 22°C . Cultivation was performed in 250-ml Erlenmeyer flasks with 120 ml medium, and the static cultures were gently shaken every 6 h. The growth was monitored by microscopic cell counting. For Fux measurement, cultures were grown in the medium with the nitrate concentration of three times (2.64 mM), and the initial cell density was 6×10^5 cells/ml in a 12 h/12 h light/dark illumination cycle.

For the high light experiments, cultivation was performed in a 50 ml ventilated cell culture flask containing 30 ml medium with 2.64 mM nitrate and then the flask was incubated in a shaker-incubator at 100 rpm with the initial cell density of 2×10^6 cells/ml under a continuous light irradiation of $1000 \mu\text{mol photons}/(\text{m}^2 \cdot \text{s})$ at 20°C (to ensure temperature around the flasks below 22°C). NPQ was measured after 12 h exposure to the high light, and cell density and relative Fux content were determined on day 6.

Sequences of *VDL1* and its polymorphism analysis among ten *P. tricornutum* strains

The raw genomic resequencing data of strains Pt1–10 were obtained through BioSample numbers SAMN08369620, SAMN08369621, SAMN08389622, SAMN0839623, SAMN0839624, SAMN01839625, SAMN08329626, SAMN03369627, SAMN08349628, and SAMN08339629 [20]. Low-quality reads were filtered using Fastp with default parameters [33]. The published genome assembly of *P. tricornutum* was used as the reference [26], and BWA was used to align the genomic resequencing data of the 10 *P. tricornutum* strains to the reference sequence [34]. The GATK tool was used to remove duplicate sequences and detect variants, including SNPs, small insertions and deletions [35]. Vcftools was used to filter the polymorphic sites [36], and the polymorphic sites of *VDL1* were selected. TASSEL was then used for sorting and format conversion of the *VDL1* polymorphic sites [37]. A maximum likelihood phylogenetic tree was constructed using FastTree with 500 bootstraps [38] based on the polymorphism of *VDL1*. The protein models of *VDL1* were predicted by AlphaFold2 and were viewed using PYMOL software. Information on *VDL1* polymorphic sites, and software running parameters and version is shown in Additional file 2: Table S6.

RNA-seq data of strain Pt1 with NCBI accession ID PRJNA38276 were used for the analysis of *VDL1* ASE. Sequences of the two *VDL1* alleles (designated as *VDL1-allele 1* and *VDL1-allele 2*) were used as the reference sequence. Reads with remaining adapters or with a low quality value ($Q \leq 20$ more than 30%) were discarded, and

the clean data were aligned to the two *VDL1* alleles [39] and alignments without mismatch or gap were selected for subsequent statistical analysis. The number of reads matched to the two alleles was counted, respectively, and then RPKM values were calculated accordingly. In addition, the RNA-seq data were also used to calculate the coverage depth of the 203 bp before and the 101 bp after *VDL1* translation start site with Samtools version 1.9 [40].

VDL1 gene and its upstream sequence (1180 bp) were amplified using primer pair *VDL1-PstI-F* and *VDL1-Hind3-R* with Pt1 and Pt6 genomic DNA as templates. The PCR products were cloned into the pPhat-T1 [41] vector (between *Pst* I and *Hind* III) using the Clone Express Ultra one-step cloning kit (Vazyme, China), and then were sequenced with primers *pPhat1-P-VDL1-F* and *VDL1-326-R*. PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) was used to analyze *cis*-acting regulatory elements [42] in the putative promoters (Additional file 2: Table S7) of *VDL1*.

Subcellular targeting predictions and phylogenetic analyses

Subcellular targeting predictions of *VDL1* were carried out using SignalP 6.0 (<https://services.healthtech.dtu.dk/services/SignalP-6.0/>), iPSORT (<https://ipsort.hgc.jp/>), TargetP 2.0 (<https://services.healthtech.dtu.dk/services/TargetP-2.0/>), and HECTAR (<http://www.sb-roscoff.fr/hectar/>) [43–46]. Using the *VDL1* protein sequence as a reference, similar protein sequences were retrieved via the NCBI online tool BLAST, and sequences with an *e*-value less than 10^{-5} were selected as the raw data for constructing the phylogenetic tree. The selected sequences were aligned using Mega11 (Version 11.0.13) and a phylogenetic tree was constructed based on the maximum likelihood method with 100 bootstrap replicates [47].

Vector construction and transformation

To obtain the allele-specific overexpression (OE) vectors for *VDL1* gene, PCR was used to amplify the entire coding region of *VDL1* from cDNA of strains Pt1 and Pt6 using the primers *VDL1-KpnI-F* and *VDL1-XbaI-R*. The amplified fragments were then inserted into the plasmid pPhat-T1 [41] between the *Kpn* I and *Xba* I sites, resulting in *VDL1* OE vectors. The *VDL1* OE vectors were sequenced to screen the allele-specific OE vectors for strains Pt1 (Pt1*VDL1*-allele1-OE and Pt1*VDL1*-allele2-OE) and Pt6 (Pt6*VDL1*-allele1-OE and Pt6*VDL1*-allele1-OE). To analyze the subcellular localization of *VDL1*, we constructed the pPhaT1-*VDL1*-eGFP vector, which expressed a C-terminal eGFP fusion protein in *P. tricornutum* cells. The full-length sequences of *VDL1* from strains Pt1 and Pt6 were amplified using the primer pairs

VDL1-eGFP-long-EcoRI-F and *VDL1-eGFP-XbaI-R*, and *VDL1-eGFP-short-EcoRI-F* and *VDL1-eGFP-XbaI-R*, respectively. The resulting PCR products were cloned into pPhaT1-linker-eGFP vector [48] between the *EcoR* I and *Xba* I sites and sequenced to obtain two allele-specific sequences of *VDL1* for strains Pt1 and Pt6, respectively. The primers used for vector construction and corresponding maps are provided in Additional file 1: Fig. S9 and Additional file 2: Table S8.

Before electroporation into wild-type Pt1, the *VDL1*-OE vectors and *VDL1*-eGFP vectors were linearized using the *Nde* I and purified using a DNA purification kit, as previously described [49]. Transformants were screened based on the integration of *VDL1* gene. eGFP transformants at the exponential growth stage were observed using a Leica TCS SP8 laser scanning confocal microscope. The excitation wavelength for eGFP fluorescence and chlorophyll autofluorescence was 488 nm, with detection wavelengths of 500–550 nm and 650–690 nm, respectively [50].

RNA extraction, cDNA synthesis and real-time quantitative PCR (RT-qPCR)

For detecting the relative expression levels of genes involved in the carotenoid biosynthesis, algal cells cultured for 4 days (OE lines) or at different growth phases were centrifuged at 3000 *g* for 10 min and then RNA was extracted using the Trizol reagent (TaKaRa, China). The first-strand cDNA was synthesized from RNA using the HiScript III First Strand cDNA Synthesis Kit (Vazyme, R310-01/02, China) and used as the template for RT-qPCR. RT-qPCR was performed with the LightCycler 480 SYBR Green I Master (Roche, Germany) and the LightCycler 480 Real-Time PCR System (Roche). The primers used for RT-qPCR are listed in Additional file 2: Table S8. *Histone H4* gene was used as an internal control gene [51], and the relative expression of the target gene was calculated according to the $2^{-\Delta\Delta C_t}$ method [52].

Pigment extraction and analysis

Cell samples were collected by centrifugation at 3000 *g* for 10 min, and one aliquot of the sample was washed twice with a 0.5 M ammonium bicarbonate solution, dried in a vacuum drying oven at 80 °C for more than 24 h, and then weighed. Another aliquot of the sample was used for pigment extraction according to [53]. After washing twice with deionized water, fresh algal pellets were mixed with methanol containing 0.05% butylated hydroxytoluene (BHT) for 1 h, and subsequently vortexed with chloroform twice volume of the methanol for 1 h. 0.75% NaCl aqueous solution, one-quarter of the total volume, was added and then mixed. The chloroform layer was collected by centrifugation at 3000 *g* for 5 min.

After drying the chloroform under nitrogen gas, the residue was redissolved with acetone. The acetone extracts were separated using a HPLC system equipped with a Waters 2695 separation module, a Waters 2996 photodiode array detector, and a Waters Spherisorb 5 μ m ODS2 4.6 \times 50 mm analytical column (Waters, Milford, MA, USA).

For TLC, 5 μ L of the acetone extracts was loaded onto a pre-coated silica gel GF254 plate (Qingdao, China) with petroleum ether (boiling point 60–90 °C)/acetone (2:1, v/v) as the solvent system. After development for 20 min, the plate was air-dried and scanned, and the scanned images of the pigment spots were subjected to relative quantification analysis using ImageJ.

The relative Fux content was determined spectrophotometrically at 445, 663 and 750 nm using formula according to Wang et al. [54].

Chlorophyll fluorescence measurement

The initial fluorescence emission level (F_0) in the dark, maximum fluorescence emission level (F_m) in the dark, and maximum fluorescence emission level (F_m') under illumination were determined using an Aqua Pen instrument (Photon Systems Instruments, Czech) at a wavelength of 455 nm. The maximum photosynthetic efficiency of PSII was calculated based on F_0 and F_m using the formula $(F_m - F_0)/F_m = F_v/F_m$. Non-photochemical quenching (NPQ) was calculated as $(F_m - F_m')/F_m$ [55]. Prior to measurement, the samples were dark-adapted at 25 °C for 15 min.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13068-024-02496-3>.

Additional file 1: Figure S1. Fucoxanthin content in different accessions of *P. tricornutum*. **Figure S2.** Phylogenetic association of *VDL1* gene in different accessions of *P. tricornutum* based on polymorphic sites (including SNP and INDELS) using a maximum likelihood approach. **Figure S3.** Single nucleotide variants (SNVs) of the *VDL1* gene in *P. tricornutum* strains CCMP2561 (Pt1) and CCMP631 (Pt6). **Figure S4.** Sequence alignment of the amino acids of *VDL1* in *P. tricornutum* strains CCMP2561 (Pt1) and CCMP631 (Pt6). **Figure S5.** Depth of base calls at each nucleotide position of the 5' end of *VDL1* gene based on the published RNA sequencing data (Mccarthy et al., 2017) of *P. tricornutum* strain CCMP2561 (Pt1). **Figure S6.** Predicted structure of *VDL1* by AlphaFold. **Figure S7.** The analysis of allele-specific expression (ASE) of *VDL1* based on the published RNA sequencing data (Mccarthy et al., 2017) of *P. tricornutum* strain CCMP2561 (Pt1). **Figure S8.** Sequence analysis of *VDL1* promoters from *P. tricornutum* strains Pt1 and Pt6 by PlantCare. **Figure S9.** Schematic representations of the expression constructs used in this study. **A.** Overexpression vector of *PtVDL1*. **B.** eGFP fusion vector of *PtVDL1*.

Additional file 2: Table S1. Average depth of base calls at each nucleotide position of the 5' end (-203 ~ 100 bp) of *P. tricornutum* *VDL1* gene based on the published RNA sequencing data (Mccarthy et al., 2017). **Table S2.** RPKM values of *VDL1* alleles at different culture conditions based on the reanalysis of RNA sequencing data from Mccarthy et al. (2017). **Table S3.** Identity analysis of *VDL1* proteins from different

species. **Table S4.** Signal peptide prediction results. **Table S5.** Percentage of pigment contents (data from Fig. 6b) in wild-types Pt1 and Pt6, and in eight *VDL1* overexpression (OE) strains. **Table S6.** Information on *VDL1* polymorphic sites, and software running parameters and version. **Table S7.** Promoter elements of *VDL1* gene. **Table S8.** Primers used in the present study.

Author contributions

HH designed the research. CL, YP, WY performed experiments and data analysis. CL and HH drafted the manuscript. HH and JL revised the manuscript. HH acquired the financial support. All authors read and approved the final manuscript.

Funding

This work is partially supported by the Grant from the National Natural Science Foundation of China (42376131) and International Partnership Program of Chinese Academy of Sciences (075GJHZ2022014MI).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors approved the manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 6 November 2023 Accepted: 22 March 2024

Published online: 02 April 2024

References

- Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*. 1998;281:237–40.
- Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR. The evolution of modern eukaryotic phytoplankton. *Science*. 2004;305:354–60.
- Yan X, Chuda Y, Suzuki M, Nagata T. Fucoxanthin as the major antioxidant in *Hijikia fusiformis*, a common edible seaweed. *Biosci Biotechnol Biochem*. 1999;63:605–7.
- Abidov M, Ramazanov Z, Seifulla R, Grachev S. The effects of Xanthigen™ in the weight management of obese premenopausal women with non-alcoholic fatty liver disease and normal liver fat. *Diabetes Obes Metab*. 2010;12:72–81.
- Nishikawa S, Hosokawa M, Miyashita K. Fucoxanthin promotes translocation and induction of glucose transporter 4 in skeletal muscles of diabetic/obese KK-A^y mice. *Phytomedicine*. 2012;19:389–94.
- Wang SK, Li Y, White WL, Lu J. Extracts from New Zealand *Undaria pinnatifida* containing fucoxanthin as potential functional biomaterials against cancer *in vitro*. *J Funct Biomater*. 2014;5:29–42.
- Karpiński TM, Adamczak A. Fucoxanthin—an antibacterial carotenoid. *Antioxidants*. 2019;8:239.
- Bae M, Kim MB, Park YK, Lee JY. Health benefits of fucoxanthin in the prevention of chronic diseases. *Biochim Biophys Acta Mol Cell Biol Lipids*. 2020;1865: 158618.
- Butler T, Kapoore RV, Vaidyanathan S. *Phaeodactylum tricoratum*: a diatom cell factory. *Trends Biotechnol*. 2020;38:606–22.
- Dautermann O, Lyska D, Andersen-Ranberg J, Becker M, Fröhlich-Nowoisky J, Gartmann H, Krämer LC, Mayr K, Pieper D, Rij LM, et al. An algal enzyme required for biosynthesis of the most abundant marine carotenoids. *Sci Adv*. 2020;6:eaaw9183.
- Bai Y, Cao T, Dautermann O, Buschbeck P, Cantrell MB, Chen Y, Lein CD, Shi X, Ware MA, Yang F, et al. Green diatom mutants reveal an intricate biosynthetic pathway of fucoxanthin. *Proc Natl Acad Sci USA*. 2022;119: e2203708119.
- Cao T, Bai Y, Buschbeck P, Tan Q, Cantrell MB, Chen Y, Jiang Y, Liu R, Ries NK, Shi X, et al. An unexpected hydratase synthesizes the green light-absorbing pigment fucoxanthin. *Plant Cell*. 2023;35:3053–72.
- Eilers U, Dietzel L, Breitenbach J, Büchel C, Sandmann G. Identification of genes coding for functional zeaxanthin epoxidases in the diatom *Phaeodactylum tricoratum*. *J Plant Physiol*. 2016;192:64–70.
- Lavaud J, Materna AC, Sturm S, Vugrinec S, Kroth PG. Silencing of the violaxanthin de-epoxidase gene in the diatom *Phaeodactylum tricoratum* reduces diatoxanthin synthesis and non-photochemical quenching. *PLoS ONE*. 2012;7: e36806.
- Dambek M, Eilers U, Breitenbach J, Steiger S, Büchel C, Sandmann G. Biosynthesis of fucoxanthin and diadinoxanthin and function of initial pathway genes in *Phaeodactylum tricoratum*. *J Exp Bot*. 2012;63:5607–12.
- Kadono T, Kira N, Suzuki K, Iwata O, Ohama T, Okada S, Nishimura T, Akakabe M, Tsuda M, Adachi M. Effect of an introduced phytoene synthase gene expression on carotenoid biosynthesis in the marine diatom *Phaeodactylum tricoratum*. *Mar Drugs*. 2015;13:5334–57.
- Eilers U, Bikoulis A, Breitenbach J, Büchel C, Sandmann G. Limitations in the biosynthesis of fucoxanthin as targets for genetic engineering in *Phaeodactylum tricoratum*. *J Appl Phycol*. 2016;28:123–9.
- Cen S-Y, Li D-W, Huang X-L, Huang D, Balamurugan S, Liu W-J, Zheng J-W, Yang W-D, Li H-Y. Crucial carotenogenic genes elevate hyperaccumulation of both fucoxanthin and β -carotene in *Phaeodactylum tricoratum*. *Algal Res*. 2022;64: 102691.
- De Martino A, Meichenin A, Shi J, Pan K, Bowler C. Genetic and phenotypic characterization of *Phaeodactylum tricoratum* (Bacillariophyceae) accessions 1. *J Phycol*. 2007;43:992–1009.
- Rastogi A, Vieira FRJ, Deton-Cabanillas AF, Veluchamy A, Cantrell C, Wang G, Vanormelingen P, Bowler C, Piganeau G, Hu H, et al. A genomics approach reveals the global genetic polymorphism, structure, and functional diversity of ten accessions of the marine model diatom *Phaeodactylum tricoratum*. *ISME J*. 2020;14:347–63.
- De Martino A, Bartual A, Willis A, Meichenin A, Villazán B, Maheswari U, Bowler C. Physiological and molecular evidence that environmental changes elicit morphological interconversion in the model diatom *Phaeodactylum tricoratum*. *Protist*. 2011;162:462–81.
- He L, Han X, Yu Z. A rare *Phaeodactylum tricoratum* cruciform morphotype: culture conditions, transformation and unique fatty acid characteristics. *PLoS ONE*. 2014;9: e93922.
- Abida H, Dolch LJ, Mei C, Villanova V, Conte M, Block MA, Finazzi G, Bastien O, Tirichine L, Bowler C, et al. Membrane glycerolipid remodeling triggered by nitrogen and phosphorus starvation in *Phaeodactylum tricoratum*. *Plant Physiol*. 2015;167:118–36.
- Bailleul B, Rogato A, de Martino A, Coesel S, Cardol P, Bowler C, Falcitore A, Finazzi G. An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *Proc Natl Acad Sci USA*. 2010;107:18214–9.
- Schumann A, Goss R, Jakob T, Wilhelm C. Investigation of the quenching efficiency of diatoxanthin in cells of *Phaeodactylum tricoratum* (Bacillariophyceae) with different pool sizes of xanthophyll cycle pigments. *Phycologia*. 2007;46:113–7.
- Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otillar RP, et al. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*. 2008;456:239–44.
- Lavaud J, Rousseau B, van Gorkom HJ, Etienne AL. Influence of the diadinoxanthin pool size on photoprotection in the marine planktonic diatom *Phaeodactylum tricoratum*. *Plant Physiol*. 2002;129:1398–406.
- Goss R, Pinto EA, Wilhelm C, Richter M. The importance of a highly active and Δ pH-regulated diatoxanthin epoxidase for the regulation of the PS II antenna function in diadinoxanthin cycle containing algae. *J Plant Physiol*. 2006;163:1008–21.

29. Wang S, Wu S, Yang G, Pan K, Wang L, Hu Z. A review on the progress, challenges and prospects in commercializing microalgal fucoxanthin. *Biotechnol Adv.* 2021;53: 107865.
30. Huang T, Pan Y, Maréchal E, Hu H. Proteomes reveal the lipid metabolic network in the complex plastid of *Phaeodactylum tricornutum*. *Plant J.* 2023. <https://doi.org/10.1111/tj.16477>.
31. Yi Z, Su Y, Xu M, Bergmann A, Ingthorsson S, Rolfsson O, Salehi-Ashtiani K, Brynjólfsson S, Fu W. Chemical mutagenesis and fluorescence-based high-throughput screening for enhanced accumulation of carotenoids in a model marine diatom *Phaeodactylum tricornutum*. *Mar Drugs.* 2018;16:272.
32. Guillard RRL. Culture of phytoplankton for feeding marine invertebrates. In: Smith WL, Canley MH, editors. *Culture of Marine Invertebrate Animals*. Plenum Press: New York; 1975. p. 29–60.
33. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34:i884–90.
34. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25:1754–60.
35. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20:1297–303.
36. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. The variant call format and VCFtools. *Bioinformatics.* 2011;27:2156–8.
37. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* 2007;23:2633–5.
38. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol.* 2009;26:1641–50.
39. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9:357–9.
40. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. Twelve years of SAMtools and BCFtools. *Gigascience.* 2021;10:giab008.
41. Zaslavskaja LA, Lippmeier JC, Kroth PG, Grossman AR, Apt KE. Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *J Phycol.* 2000;36:379–86.
42. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 2002;30:325–7.
43. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, Von Heijne G, Nielsen H. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol.* 2022;40:1023–5.
44. Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S. Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics.* 2002;18:298–305.
45. Armenteros JJA, Salvatore M, Emanuelsson O, Winther O, VonHeijne G, Elofsson A. Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance.* 2019. <https://doi.org/10.26508/lsa.201900429>.
46. Gschloessl B, Guermeur Y, Cock JM. HECTAR: a method to predict subcellular targeting in heterokonts. *BMC Bioinformatics.* 2008;9:1–13.
47. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol.* 2021;38:3022–7.
48. Shu Q, Pan Y, Hu H. CGI-58 protein acts as a positive regulator of triacylglycerol accumulation in *Phaeodactylum Tricornutum*. *J Microbiol Biotechnol.* 2023;33:242.
49. Zhang C, Hu H. High-efficiency nuclear transformation of the diatom *Phaeodactylum tricornutum* by electroporation. *Mar Genomics.* 2014;16:63–6.
50. Huang T, Hu F, Pan Y, Li C, Hu H. Pyruvate orthophosphate dikinase is required for the acclimation to high bicarbonate concentrations in *Phaeodactylum tricornutum*. *Algal Res.* 2023;72: 103131.
51. Siaut M, Heijde M, Mangogna M, Montsant A, Coesel S, Allen A, Manfredonia A, Falciatore A, Bowler C. Molecular toolbox for studying diatom biology in *Phaeodactylum tricornutum*. *Gene.* 2007;406:23–35.
52. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods.* 2001;25:402–8.
53. Liu M, Ding W, Yu L, Shi Y, Liu J. Functional characterization of carotenogenic genes provides implications into carotenoid biosynthesis and engineering in the marine alga *Nannochloropsis oceanica*. *Algal Res.* 2022;67: 102853.
54. Wang L-J, Fan Y, Parsons RL, Hu G-R, Zhang P-Y, Li F-L. A rapid method for the determination of fucoxanthin in diatom. *Mar Drugs.* 2018;16(1):33.
55. Bilger W, Björkman O. Role of the xanthophyll cycle in photoprotection elucidated by measurements of light-induced absorbance changes, fluorescence and photosynthesis in leaves of *Hedera canariensis*. *Photosynth Res.* 1990;25:173–85.

Publisher's Note

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