International Journal of
Molecular Sciences

MDPI

Article

Bioengineering of the Marine Diatom *Phaeodactylum tricornutum* with Cannabis Genes Enables the Production of the Cannabinoid Precursor, Olivetolic Acid

Fatima Awwad [1,†], Elisa Ines Fantino [1,†], Marianne Héneault [1], Aracely Maribel Diaz-Garza [1], Natacha Merindol [1,2], Alexandre Custeau [1], Sarah-Eve Gélinas [1], Fatma Meddeb-Mouelhi [1,2], Jessica Li [3], Jean-François Lemay [4], Bogumil J. Karas [3] and Isabel Desgagne-Penix [1,2,*]

[1] Department of Chemistry, Biochemistry and Physics, Université du Québec à Trois-Rivières, 3351 Boulevard des Forges, Trois-Riviere, QC G9A 5H7, Canada
[2] Groupe de Recherche en Biologie Végétale, Université du Québec à Trois-Rivières, Trois-Riviere, QC G9A 5H7, Canada
[3] Department of Biochemistry, Schulich School of Medicine and Dentistry, Western University, London, ON N6A 5C1, Canada
[4] Centre National en Électrochimie et en Technologies Environnementales Inc., 2263 Avenue du Collège, Shawinigan, QC G9N 6V8, Canada
* Correspondence: isabel.desgagne-penix@uqtr.ca
† These authors contributed equally to this work.

Abstract: The increasing demand for novel natural compounds has prompted the exploration of innovative approaches in bioengineering. This study investigates the bioengineering potential of the marine diatom *Phaeodactylum tricornutum* through the introduction of cannabis genes, specifically, tetraketide synthase (TKS), and olivetolic acid cyclase (OAC), for the production of the cannabinoid precursor, olivetolic acid (OA). *P. tricornutum* is a promising biotechnological platform due to its fast growth rate, amenability to genetic manipulation, and ability to produce valuable compounds. Through genetic engineering techniques, we successfully integrated the cannabis genes *TKS* and *OAC* into the diatom. *P. tricornutum* transconjugants expressing these genes showed the production of the recombinant TKS and OAC enzymes, detected via Western blot analysis, and the production of cannabinoids precursor (OA) detected using the HPLC/UV spectrum when compared to the wild-type strain. Quantitative analysis revealed significant olivetolic acid accumulation (0.6–2.6 mg/L), demonstrating the successful integration and functionality of the heterologous genes. Furthermore, the introduction of TKS and OAC genes led to the synthesis of novel molecules, potentially expanding the repertoire of bioactive compounds accessible through diatom-based biotechnology. This study demonstrates the successful bioengineering of *P. tricornutum* with cannabis genes, enabling the production of OA as a precursor for cannabinoid production and the synthesis of novel molecules with potential pharmaceutical applications.

Keywords: diatom; metabolic engineering; olivetolic acid cyclase; tetraketide synthase; synthetic biology; cannabinoids; microalgae

Citation: Awwad, F.; Fantino, E.I.; Héneault, M.; Diaz-Garza, A.M.; Merindol, N.; Custeau, A.; Gélinas, S.-E.; Meddeb-Mouelhi, F.; Li, J.; Lemay, J.-F.; et al. Bioengineering of the Marine Diatom *Phaeodactylum tricornutum* with Cannabis Genes Enables the Production of the Cannabinoid Precursor, Olivetolic Acid. *Int. J. Mol. Sci.* **2023**, *24*, 16624. https://doi.org/10.3390/ijms242316624

Academic Editors: Chiara Lauritano and Assunta Saide

Received: 2 September 2023
Revised: 19 November 2023
Accepted: 20 November 2023
Published: 22 November 2023

1. Introduction

Over a thousand of the molecules produced in *Cannabis sativa* plants have been reported in the literature, including more than a hundred cannabinoids (CBs) [1]. The antinociceptive and appetite-stimulating properties of CBs have been studied thoroughly, as well as their association with relieving the symptoms associated with different diseases such as multiple sclerosis and the relief of chemotherapy side effects [2–5]. The most studied CBs are Δ9-tetrahydrocannabinol (THC), which has psychoactive potential, and cannabidiol (CBD), which acts on human endocannabinoid receptors to modulate pain and is already contained in many pharmaceutical products available on the market, such

as Sativex and Bediol [6–8]. Other therapeutic properties of specific CBs that could be administrated in a purified form are still under investigation and offer a promising alternative for different human pathology applications [5]. In planta, phytocannabinoids are produced in glandular trichomes via fatty acid and terpenoid precursors [9]. The genes and enzymes implicated in the cannabinoid pathway were characterized by several plant biologists shortly after the sequencing of the *C. sativa* genome [9–12]. The first step in the CB biosynthetic pathway is the formation of olivetolic acid (OA) from malonyl-CoA and hexanoyl-CoA, through a two-step reaction performed by a type-III polyketide synthase named tetraketide synthase (*Cs*TKS) followed by a cyclase named olivetolic acid cyclase (*Cs*OAC), both of which are localized in the cytosol or cells from the glandular trichomes (Figure 1). The enzymes belonging to the type-III polyketide synthase (PKS) superfamily are responsible for producing a wide array of central frameworks that are found in specialized plant metabolites of significant medicinal value. These metabolites encompass various categories, such as flavonoids, stilbenes, chromones, pyrones, phloroglucinols, resorcinols, xanthones, acridones, and quinolones [13,14]. Examples of PKS include chalcone synthase, which accepts *p*-coumaroyl-CoA as the starter substrate to catalyze three successive condensations with malonyl-CoA to generate naringenin chalcone [15,16] and stilbene synthase, which generates resveratrol [17]. The polyketide synthase from *C. sativa* (also named tetraketide synthase or olivetol synthase) was characterized and has been shown to be dependent on OAC for the formation of OA (Figure 1) [9]. The 3D structure of *Cs*OAC is most accurately described as a dimeric α + β barrel (DABB) protein, with similarities to other plant DABBs, especially those implicated in stress responses and to the DABB proteins found in *Streptomyces* bacteria [9]. The aromatic precursor OA is a resorcinol carrying an alkyl chain (named alkylresorcinolic acid), which is condensed with a monoterpenoid moiety (geranyl diphosphate, GPP) to yield the general structure of the most well-known CBs, as reported in feeding studies of 13C-labeled glucose in the early 1960s [18,19]. Although not all CBs are derived from the OA precursor, it remains the most widely characterized pathway in *C. sativa* [20]. The prenyl moiety of GPP is transferred to OA by an aromatic prenyltransferase (*Cs*APT) to form cannabigerolic acid (CBGA) (Figure 1). This core intermediate is the backbone of the remaining reaction, which diverges to form cannabidiolic acid (CBDA) and tetrahydrocannabinolic acid (THCA) through CBDA-synthase (CBDAS) and THCA-synthase (THCAS), respectively (Figure 1), as well as other CBs such as cannabichromenic acid (CBCA) via cannabichromenic acid cyclase (CBCAS) [9,11,21,22]. The neutral phytocannabinoids THC and CBD are generated from cannabinolic acids by non-enzymatic decarboxylation and are often produced upon exposure to light or heat (Figure 1) [21]. Aside from CBs, the cannabis plants also produce other specialized metabolites from shared precursors with cannabinoids, such as fatty acids, competing with olivetolic acid synthesis, and terpenoids such as limonene [23], competing for the GPP pool [24,25].

The yield of CB in plants is limited, both in quality and in quantity; the maximum production is due to a mixture of these metabolites accumulating in the flowers (20 mg/g fresh weight) [6,26]. In addition, the lack of scientific literature on cannabis physiology leads to the absence of standardized culturing protocols, resulting in a variation in CB profiles, quality, and quantity. In addition, cannabis pathogens such as the gray mold causal agent (*Botrytis cinerea*) affect the overall plant yield and metabolite composition [27]. As well as multiple unknown facets of CB production in planta, this suggests that the implementation of a heterologous expression system with fewer components and fewer competing enzymes might be a solution for achieving a higher production yield of single CB components, such as CBD or THC. Therefore, synthetic biology offers an important alternative method to produce CBs for the pharmaceutical field with less risk and fewer crop management challenges.

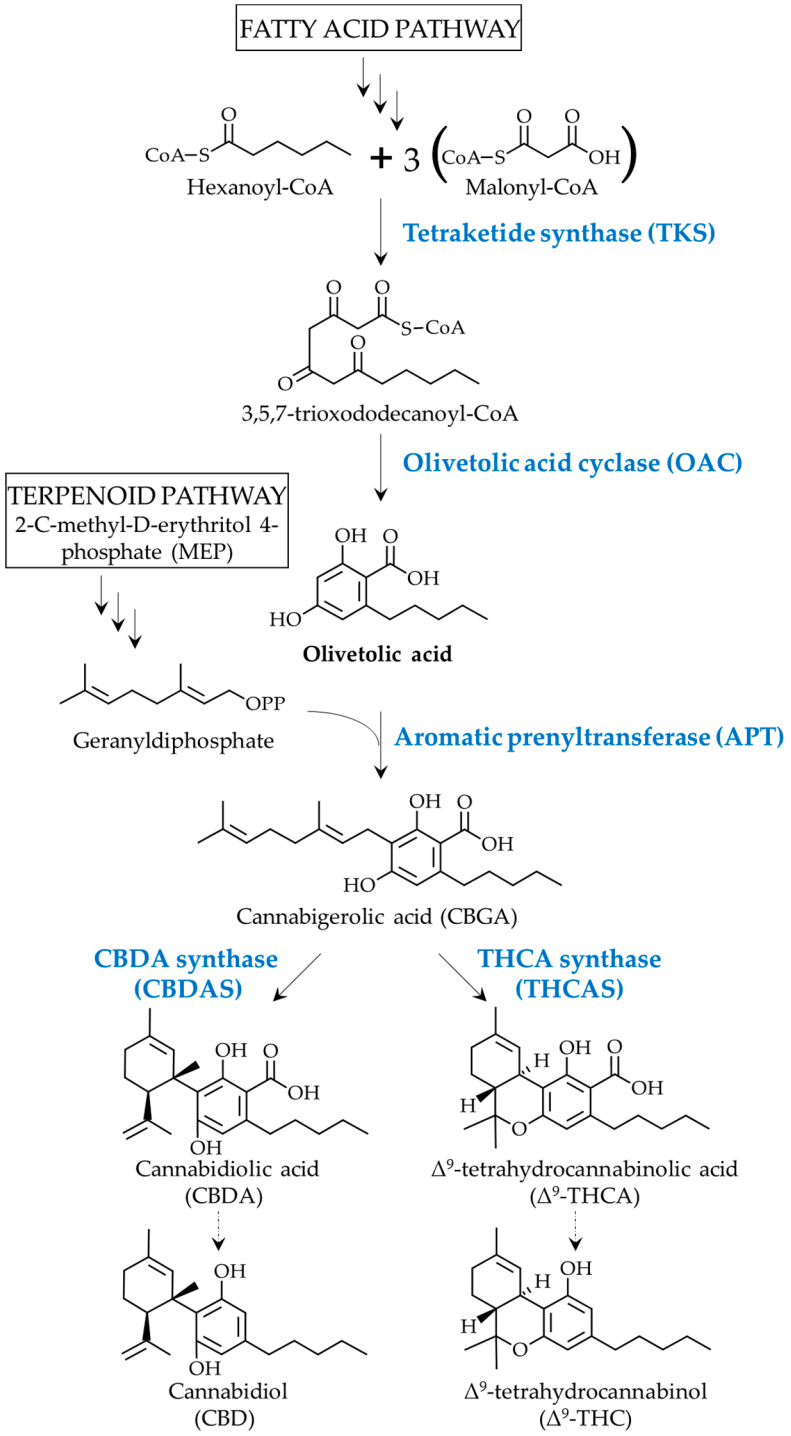


Figure 1. The proposed cannabinoid biosynthetic pathway, leading to the two major phytocannabinoids, THC and CBD. Enzymes are shown in blue. Bolded arrows represent an enzymatic reaction, whereas the dotted arrows represent non-enzymatic (heat or light) decarboxylation.

Several research groups aimed to design de novo or improve pre-existing cannabinoid production in plant tissue cultures, but apart from the transformation rates, in vitro culture viability and yield are still to be optimized [6]. For instance, few groups succeeded in transforming yeast (*Saccharomyces cerevisiae*) and bacteria (*Escherichia coli*) to produce the different metabolites of the CB pathway, such as OA or CBGA [28,29]. The production of these metabolites in yeast and bacteria often involves the addition of costly precursors such as acetyl-CoA, hexanoyl-CoA, or olivetolic acid [28,30], or involves supplementing the

media with sugars [31], such as galactose, to provide a more affordable alternative with a long series of gene stacking. For instance, the production of OA in *Dyctiostelium discoideum* on a 300-L scale led to 4.8 µg/L [32]. OA production in *E. coli* reached 80 mg/L after introducing a series of modifications to increase the CoAs pool [33]. However, the bacterial and yeast translation and post-translational mechanisms, as well as their metabolic chassis, are distinct from the mechanisms found in higher plants. This resulted in low yields, such as picomoles per unit of optical density, from these microorganisms when compared to the bioengineering and supplementation costs [31,34]. By contrast, diatoms are photosynthetic organisms that share the main metabolic paths with plants, making them promising candidates for the production of heterologous compounds [9]. All the precursor pathways, such as acetyl-CoA, hexanoyl-CoA, GPP, etc., are present in diatom cells at different times of the cell cycle, which reduces the need to supplement the culture media with costly precursors or sugars [35]. Also, the extrachromosomal transformation of diatoms via bacterial conjugation allows the fairly stable expression of heterologous genes [36,37]. Similar advances in the bioengineering of diatoms have been shown recently, allowing the production of monoterpenoids such as geraniol, an insect repellant, in the model diatom *Phaeodactylum tricornutum* [36].

Thus, the aim of this study is to produce the CB precursor OA in the marine microalga, *P. tricornutum*. Here, we inserted cannabis *Cs*TKS and *Cs*OAC genes into a stable optimized episome that was assembled in *S. cerevisiae* [38,39]. The recombinant episome was then transformed via a trans-kingdom conjugation between *E. coli* and *P. tricornutum* [38]. The transconjugant strains were screened for their ability to produce molecules of interest, using a high-performance liquid chromatograph coupled to an ultra-violet detector (HPLC-UV). We detected OA (0.6–2.6 mg/L) in *P. tricornutum* transconjugants that issued from two different expression cassettes. Although the production of the OA compounds of interest was temporary, this work sheds light on a powerful and suitable system for plant metabolite heterologous production in a cost-effective phototropic system employing *P. tricornutum*.

2. Results

2.1. Heterologous Protein Expression and Localization in the Diatom P. tricornutum

2.1.1. Transformation Validation and Transconjugant Characterization

To inform and design the metabolic engineering strategy, we first performed computational analyses to assess the availability of the lipid-derived precursors required to produce OA and cannabinoids in *P. tricornutum* with in silico tools and showed that upstream pathways existed in the host organism. According to the Biocyc database [40], hexanoyl-CoA can be formed in *P. tricornutum* in the peroxisome via the fatty acid β-oxidation pathway, mainly by long-chain acyl-CoA synthase (*Pt*ACS3 and *Pt*ACS4), while malonyl-CoA is formed via the activity of an ACYL-CoA carboxylase (ACCase). Moreover, the geranyl pyrophosphate (GPP) pool is present at detectable levels in *P. tricornutum* [36].

To produce OA in *P. tricornutum*, three different expression cassettes were designed using p*Pt*GE30 as a backbone plasmid DNA (Figure 2A), namely, *Pt*OA1, *Pt*OA2, and *Pt*OA3, with various combinations of *C. sativa* genes encoding for *Cs*TKS and *Cs*OAC (Figure 2B). Each expression cassette was codon-optimized for *P. tricornutum* and assembled (Appendix A, Figure A1, Tables A1–A3) according to the HIVE lab codon usage table (CUT) [41].

The transconjugant strains *Pt*OA1, *Pt*OA2, and *Pt*OA3 did not present significant differences in growth and shape compared to *P. tricornutum* when transformed with the p*Pt*GE30 empty vector or wild-type strains (Figure 3A–C). In addition, morphological properties such as morphotype, size, granulosity, and chlorophyll levels did not significantly differ between the studied strains, as evaluated by microscopy (Figure 3A) and flow cytometry (Figure 4).

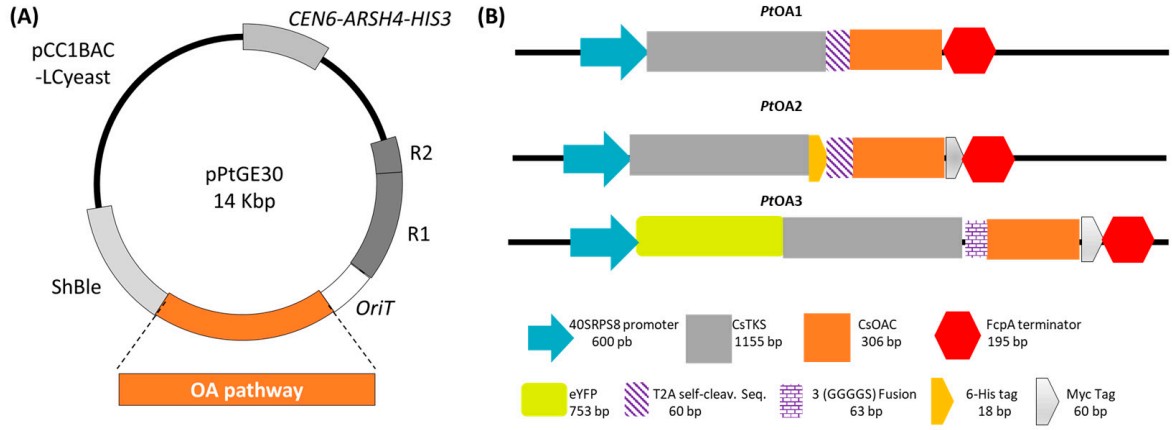


Figure 2. Design of the expression vectors *Pt*OA1, *Pt*OA2, and *Pt*OA3. (**A**) Schematic representation of the p*Pt*GE30 expression vector. (**B**) Description of the recombinant cassettes used to express *C. sativa* polyketide synthase (*Cs*TKS) and olivetolic acid cyclase (*Cs*OAC) under *P. tricornutum*'s constitutive promoter 40SRPS8, the fucoxanthin–chlorophyll binding protein A (FcpA) terminator, with or without tags (His = histidine, eYFP = enhanced yellow fluorescent protein, Seq = sequence, T2A = the *Thosea asigna* virus 2A cleavable sequence), and reporter genes of different sizes (base pair = bp).

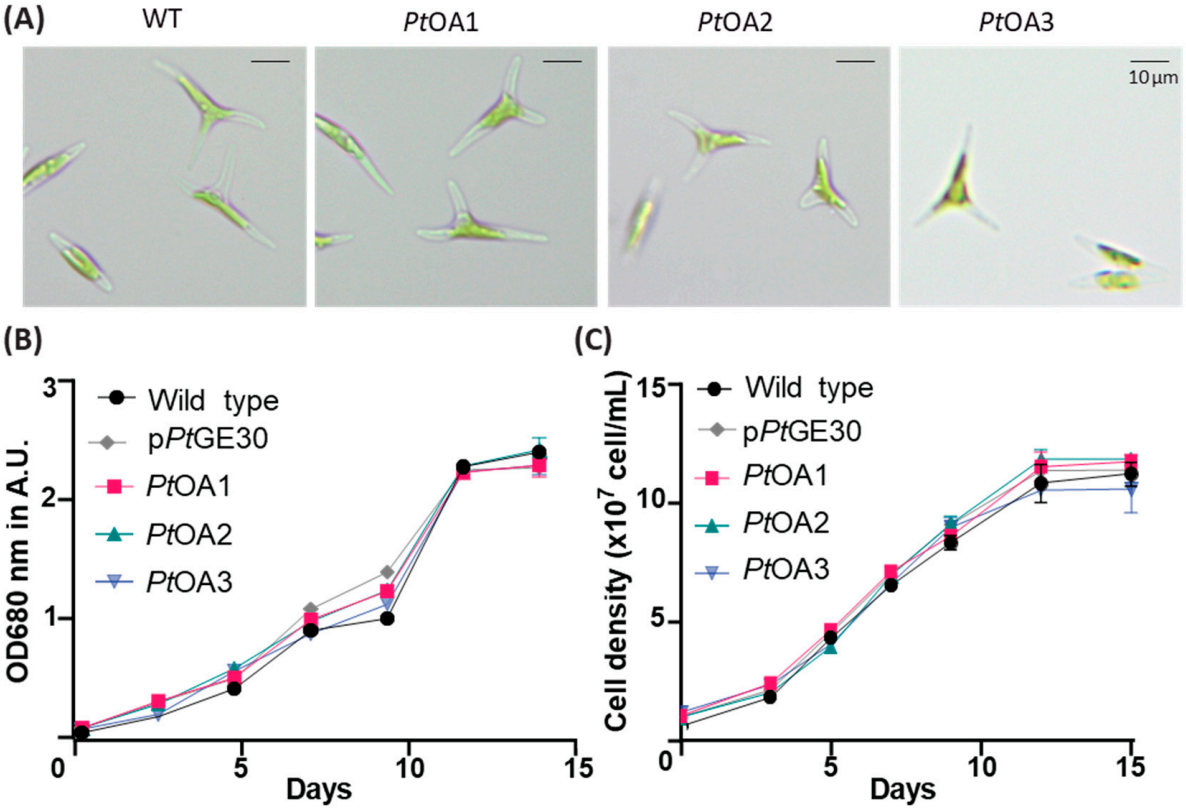


Figure 3. Cell shapes and growth curves of the transformed clones in comparison with wild-type *P. tricornutum*. (**A**) Optical microscope images of wild-type (WT) *P. tricornutum* and the transformed cells *Pt*OA1, *Pt*OA2, and *Pt*OA3. (**B**,**C**) *P. tricornutum* WT, transformed with pPtGE30, *Pt*OA1, *Pt*OA2, and *Pt*OA3 cell culture growth curves, followed by absorbance at 680 nm (**B**) or via cell count (**C**) during the 15 days of the cycle. The plotted values represent the means of three replicates and error bars represent the standard deviation.

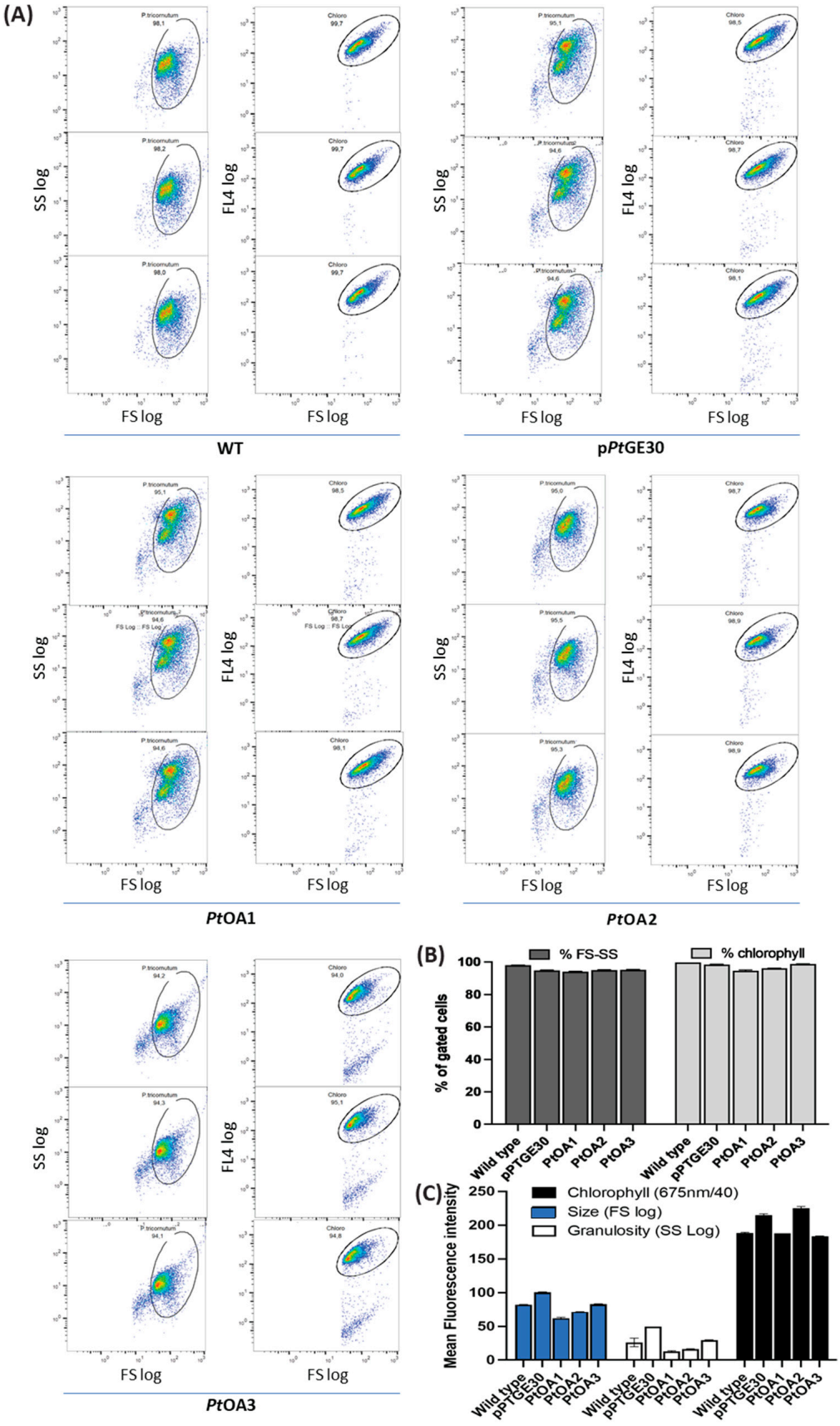


Figure 4. Flow cytometry observation for cell size and chlorophyll auto-fluorescence among *P. tricornutum* transconjugants and wild-type cultures. (**A**) Scatter plots for each strain and (**B**) a graphical

representation of chlorophyll and size percentages among gated cells of wild-type *P. tricornutum* (WT), and *P. tricornutum*, harboring the empty vector p*Pt*GE30, *Pt*OA1, *Pt*OA2, and *Pt*OA3. (**C**) Histogram representing the mean chlorophyll fluorescence intensity, mean granulosity, and mean size scatter.

2.1.2. Heterologous Protein Detection and Localization

The successful expression of cannabis genes and the accumulation of the corresponding enzymes was observed in *P. tricornutum* transconjugants. Indeed, the accumulation of the *Cs*TKS-T2A (45 kDa band) protein in *Pt*OA1C1 (Figure 5A), a 68 kDa uncleaved protein corresponding to CsTKS-His-T2A-CsOAC-cMyc in *Pt*OA2C2 (Figure 5B), and an 85 kDa band corresponding to YFP-CsTKS-3(GGGGS)-CsOAC-cMyc in *Pt*OA3C1 (Figure 5C) was detected via Western blot analysis in PCR-positive *P. tricornutum* transconjugants (Appendix A, Figure A1B). However, the construction with the self-cleavable sequence did not yield detectable successful single proteins in *Pt*OA2. This could be caused by the tag (6xHis) used in the OA2 construct but not in the OA1 construct, changing the efficiency of cleavage (Figure 5B). Then, we studied the localization of *Cs*TKS in *PtOA*3, taking advantage of the presence of the fluorescent YFP fusion protein. We observed under confocal microscopy that the heterologous recombinant protein (YFP-*Cs*TKS-3(GGGGS)-CsOAC-cMyc) produced in *PtOA*3 accumulated and was localized to the cytosol (Figure 5D).

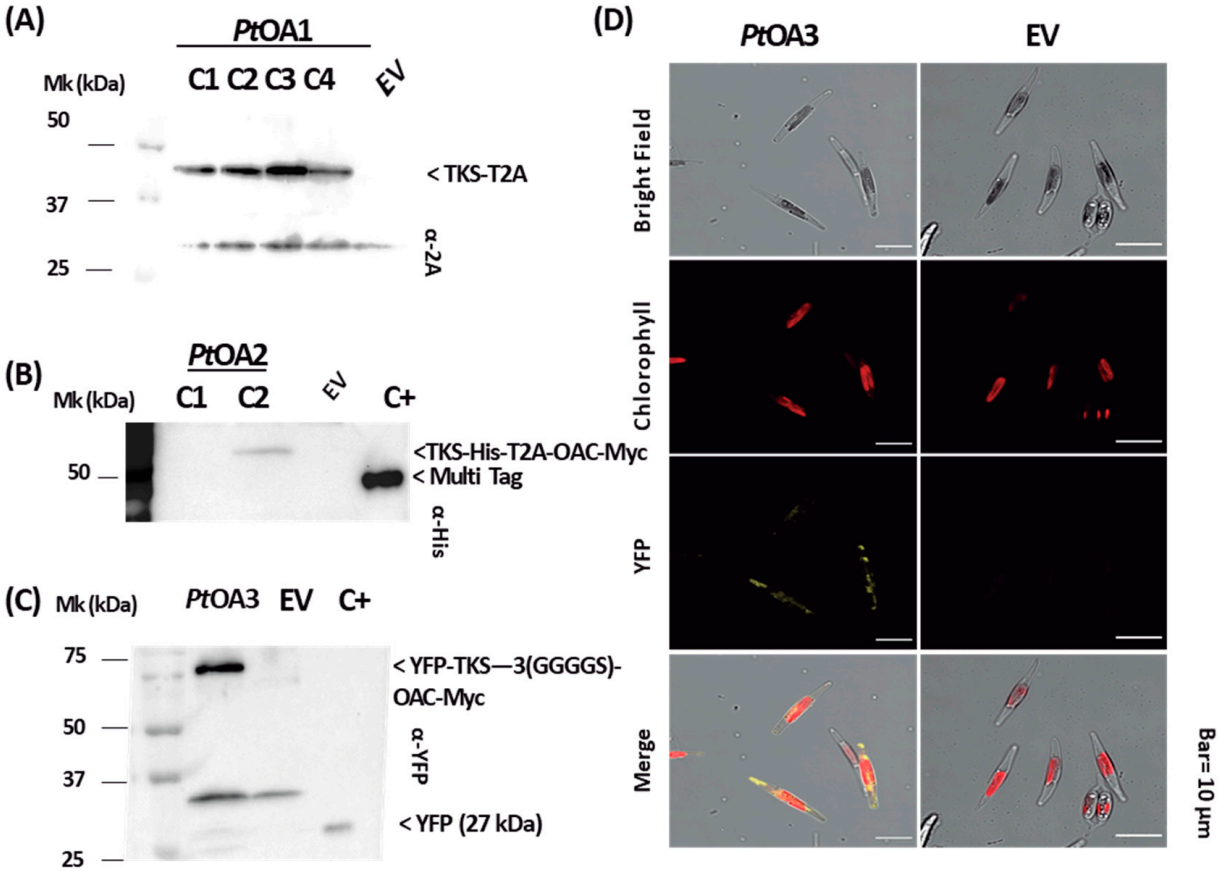


Figure 5. Expression and localization of heterologous proteins in transconjugant *P. tricornutum* strains. (**A**) Whole-cell extraction of four *Pt*OA1 (TKS-T2A-OAC) strains 1–4; TKS protein detection (45 kDa) was performed using anti-T2A antibodies. (**B**) Whole-cell extraction of two *Pt*OA2 (TKS-His-T2A-OAC-cMyc) clones 1 and 2; TKS:T2A:OAC protein detection (68 kDa) was performed using anti-His antibodies. (**C**) Whole-cell extraction of *Pt*OA3 (YFP-TKS-3(GGGGS)-OAC-cMyc) strains; YFP:TKS:OAC protein detection (85 kDa) was performed using anti-YFP antibodies. Mk: protein ladder; whole-cell extracts of the pPtGE30 culture were used as negative control (empty vector, EV);

C+: Multi Tag was used as a positive control. All Western blots were obtained from 12% SDS-PAGE gel. For each construct, four clones were characterized and the ones showing protein production are presented in this figure. (**D**) Confocal microscopic images of *Pt*OA3 and pPtGE30 cells in a bright field, with autofluorescence of the chloroplast, YFP:TKS:OAC fluorescence, and the merging of three fields are shown.

2.2. Temporal Metabolite Production

2.2.1. Metabolite Detection and Identification

Selected positive transconjugant strains, displaying protein accumulation via Western blot analysis, were cultivated, extracted, and used for targeted metabolite analysis. OA was detected at 2.6 mg/kg in *Pt*OA1 and 0.54 mg/kg in *Pt*OA2, using HPLC-UV, whereas no peaks at the OA retention time were detected in the *Pt*OA3 extracts (Figure 6 and Appendix A, Figure A3).

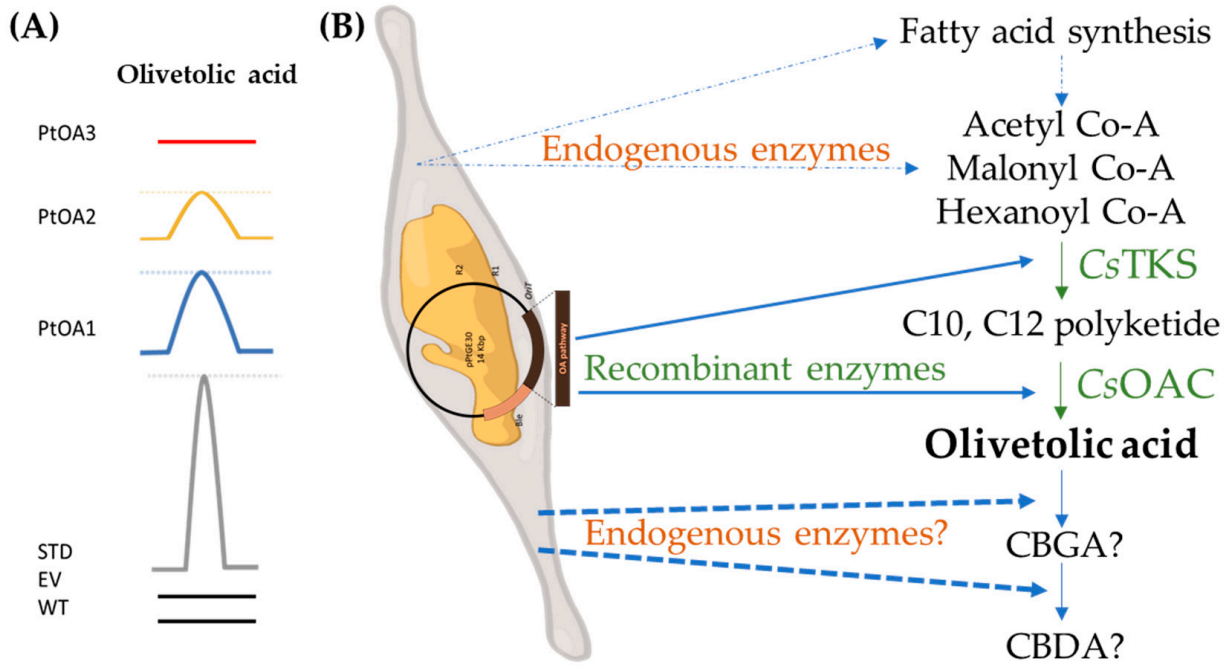


Figure 6. Olivetolic acid production, obtained via HPLC-UV analysis at 280 nm. (**A**) Curves represent the peak at the OA retention time of chromatograms, obtained following the injection of 10 µL of each extracted sample. At the bottom, a concentration of 10 µg/mL of commercial standards (STD) is shown: metabolite extraction from the *P. tricornutum* wild type (WT); *P. tricornutum,* transformed with an empty vector (EV = p*Pt*GE30), *Pt*OA1, *Pt*OA2, and *Pt*OA3. The peaks represent olivetolic acid. (**B**) Schematic representation of the bioengineered *P. tricornutum* for the production of cannabinoid precursors.

Interestingly, additional peaks with similar retention times to the cannabinoid standards of CBN, CBD, and THC, and their corresponding acid forms were detected in the extracts of OA-producing transconjugants (*Pt*OA1C1 and *Pt*OA2C2) (Appendix A, Figure A4). Liquid chromatography–mass spectrometry (LC-MS) analyses using electrospray ionization in the positive mode (ESI+) were performed to validate if the mass-to-charge ratio (m/z) of these peaks corresponded to protonated CBs $(M + H)^+$. The compounds were first scanned for specific masses corresponding to the metabolites of interest, then, in the case of a positive scan, another LC-MS analysis was performed to make sure that the fragments corresponded to the target compound. Unexpectedly, the daughter m/z obtained after employing collision-induced energy (CID) at 40 V of the selected RT and m/z corresponded to the cannabinoid standards (Appendix A, Figure A4). This suggests that *P. tricornutum* may contain endogenous enzymes that are capable of converting OA into CB-like metabo-

lites such as CBGA, CBDA, and THCA. Such enzymes could be mimicking the activity of cannabis APT (using the endogenous pool of GPP) and CBDAS/THCAS (Figure 1). Candidate enzymes from the diatom database were searched and low-similarity sequences for APT and THCAS were identified (Appendix B, Figure A6). However, further characterization is needed and might help find closer candidates among microalgae. Alternatively, candidates from a completely different type of enzyme could be involved.

To test for the presence of such endogenous enzymes, we performed a supplementation assay with *P. tricornutum* transconjugants harboring the p*Pt*GE30, which were devoid of cannabis genes (Table 1). We first added OA and GPP for 16 h and could subsequently detect small amounts (circa 3–5 mg/kg) of CBGA and other CB-like metabolites using HPLC analysis. This suggests the presence of an APT-like enzyme in *P. tricornutum* (Table 1). Next, we supplemented the sample with CBGA and detected metabolites that behaved in a similar way to CB on an HPLC chromatogram, but we were unable to confirm their identity (not all were appropriate m/z fragments) via LC-MS.

Table 1. Results of a supplementation assay of *P. tricornutum* with olivetolic acid (OA), geranyl diphosphate (GPP), and cannabigerolic acid (CBGA). The results show the mean detection levels that were found in three different strains of *P. tricornutum* transconjugants harboring p*Pt*GE30.

Supplemented mM			Detected			
OA	**GPP**	**CBGA**	**HPLC**	**MS (*m/z*) Confirmation**	**HPLC**	**MS (*m/z*) Confirmation**
0.45	1	0	CBGA	yes	CBDA, CBNA	No
0	0	0.27	CBDA	yes	THCA, CBNA	No

2.2.2. Metabolite Production Timeline

Olivetolic acid and cannabinoid-like metabolites were detected up to three months after obtaining and subculturing each transconjugant in optimal conditions with no supplementation of any kind. However, a gradual decrease in the concentration of the OA and CB-like metabolites was observed. After six months, the transconjugants completely lost the ability to produce these desired metabolites. Along with production loss, we noticed that the *P. tricornutum* transconjugant morphotype also changed from triradiate (Figure 3A) to fusiform (Figure 5D), and that the initial triradiate morphotype of the transconjugants was never restored.

3. Discussion

The bioengineering of *P. tricornutum* with cannabis genes represents a promising approach to addressing the challenges associated with traditional cannabinoid production. By harnessing the unique capabilities of marine diatoms, this study paves the way for sustainable and controlled cannabinoid biosynthesis. Furthermore, the successful expression of cannabis genes in a non-native host highlights the versatility of synthetic biology in the context of metabolic engineering. This work represents the first report of cannabinoid engineering in brown microalgae and provides a proof-of-concept evaluation of *P. tricornutum* for the heterologous production of OA and cannabinoids. We provide evidence that the natural metabolism of the marine diatom *P. tricornutum* is well-suited for use in metabolic engineering to produce the cannabinoid precursor, OA, and CB-like metabolites.

The episomal vectors, *Pt*OA1 and *Pt*OA2, contained both *CsTKS* and *CsOAC* genes, separated by a self-cleaving peptide sequence and under the control of the same promoter, to ensure similar levels of expression of both genes and, subsequently, comparable levels of accumulated enzymes for performing the two-step reactions. The transconjugants *Pt*OA1 and *Pt*OA2 only differed in terms of the addition of a 6xHis tag on *Cs*TKS for easier detection of the protein. In *Pt*OA3, both genes were linked together with a linker and tagged with YFP for visualization and localization studies.

Although a previous study demonstrated the occurrence of episomal rearrangements [42], in this study, the *P. tricornutum* transconjugant strains obtained for each episomal vector did not show significant DNA sequence rearrangement after 1 year (Appendix A, Figure A1 and Table A4). This suggests that episomal rearrangements could be sequence- or DNA-motif-specific. Also, the *Pt*OA1-3 transconjugant strains grew similarly to the WT and EV strains (Figures 3 and 4). This suggests that the presence and expression of the transgenes and the produced metabolites did not significantly affect the growth and division of *P. tricornutum*. Genome analysis of green algae, diatoms, and higher plants revealed similarities in lipid metabolism [43,44]. Therefore, the fact that there is little to no effect of *Cs*TKS and *Cs*OAC expression on *P. tricornutum* growth and fitness could be a result of the nature of the pathway, as an extension of fatty acid metabolism that is fairly close to the diatom chassis, and a preference for energy conversion into lipids.

Heterologous protein detection was not consistent among the different transconjugant strains (Figure 5 and Appendix A, Figure A2). For example, the four analyzed transconjugant strains of *Pt*OA1 show a similar pattern of accumulation, whereas, in the *Pt*OA2 strains, the T2A sequence did not appear to be cleaved. This could be explained either by the possible presence of negatively acting *cis* elements in the sequence of *Pt*OA2 transgenes [43] or by the nature of the amino acid environment that surrounds the T2A sequence cleavage site [45]. It was also reported that the gene following the T2A sequence would be affected on the expression level [46] and, thus, would yield fewer *Cs*OAC transcripts, but the effect of this phenomenon would be validated if the C10–C12 polyketide product was available as a standard for HPLC detection, which is not the case.

When compared to previous studies, the production of OA in *Yarrowia lypolytica* slightly changed when *Cs*TKS and *Cs*OAC were expressed under different promoters, compared to when the two genes were fused [42]; in *P. tricornutum*, the *Pt*OA1 showed a higher quantity of OA (Figure A3) compared to there being none in *Pt*OA3 when the two genes were fused by a glycine-serine linker. This is not applicable in *Pt*OA2 since the Western blot did not show cleaved proteins, which could indicate that the cleaved proteins were below detection levels.

To our knowledge, this is the first time that a brown microalgae host has demonstrated the capacity to accumulate OA and use it for the synthesis of CB-like metabolites. The heterologous production of OA was observed in the different transconjugants of *P. tricornutum*, transformed with episomal constructions and harboring various combinations of *C. sativa TKS* and *OAC* genes (Figure 2; Appendix A, Figure A1 and Table A2). The cassette designs in this study were fairly simple and direct, with only two main cannabis genes being introduced. When compared to the literature, the first approaches regarding yeast achieved the production of only 0.48 mg/L of OA by expressing *Cs*TKS and *Cs*OAC, and by feeding the culture with sodium hexanoate [9]. Here, we obtained higher yields (0.56–2.6 mg/kg) compared to yeast with and without hexanoate supplementation [9]. In *Y. lipolytica*, it was bioengineered with *Pseudomonas* sp. *LvaE*, encoding a short-chain acyl-CoA synthetase, acetyl-CoA carboxylase, pyruvate dehydrogenase bypass, NADPH-generating malic enzyme, as well as the activation of the peroxisomal β-oxidation pathway and the ATP export pathway to redirect the carbon flux toward OA synthesis [47]; the yield was also less than 0.5 mg/L until batch culture optimization.

Further optimization of the *P. tricornutum* transconjugant strains by boosting the pool of precursors could increase the production of OA. In *E. coli*, besides introducing *Cs*TKS and *Cs*OAC, metabolic engineering consisted of the overexpression of the acetyl-CoA carboxylase subunits A, B, C, and D from the same host to ensure the high production of malonyl-CoA, and a reversal of beta-oxidation that allowed the accumulation of hexanoyl-CoA [33]. This led to a higher quantity of OA (46.3 mg/L) after the optimization of culture conditions and induction parameters [33]. In *P. tricornutum*, hexanoyl-CoA and malonyl-CoA occur naturally, due to the long-chain acyl-CoA synthase (*Pt*ACS3 and *Pt*ACS4) for hexanoyl-CoA and to the acetyl-CoA ATP-dependent carboxylase (*Pt*ACCase) for malonyl-CoA [43]. *P. tricornutum* was also shown to accumulate intracellular amounts

of GPP, along with neryl diphosphate (NPP) [36,48]; both precursors are donors of the prenyl moiety of several cannabinoids, such as CBGA and cannabinerolic acid (CBNA) in cannabis [46]. However, overexpression of these upstream genes could eliminate the need for supplementation while maintaining the necessary pools of precursors for more stable production of cannabinoids, which is similar to the process used with yeast [31,34].

Along with precursor bioavailability, endogenous enzymes are required to enable the conversion of OA into CB or CB-like compounds, as observed in this study, either by direct production (Figure A4) or by supplementation assays (Table 1). Bioinformatic analyses led to the identification of two putative enzymes (Appendix B, Figure A6) that are encoded by Phatr3_J37858 and Phatr3_J2738.t1 and are capable of prenyltransferase activity, thereby allowing the production of CBGA. Both enzymes are UbiA prenyltransferase family members that could be used originally by the diatom to produce terpenes or flavonoids [49,50]. The in silico identification of THCAS and CBDAS-like endogenous candidate genes is more intricate, given the specificity of these enzymes. Our approach can also be used to further characterize endogenous diatom enzymes with multiple functions by testing the reactivity of diatoms to drug precursors in higher-level plants.

The episomal sequences were well-conserved in the timeframe of this study in all three constructions of interest (in DNA and proteins). However, after such a timeframe, the production of OA was not detectable, perhaps due to silencing via an epigenetic mechanism [42], mRNA or protein stability [37], and, less probably, episomal rearrangements, as seen in other cases [42]. It would be useful to characterize a larger number of transformed *P. tricornutum* using a method similar to the process used with yeast [31,34], or to try randomly integrated chromosomal expression (RICE). Interestingly, a loss of production was observed in parallel with a change in cell morphotype from a triradiate to a fusiform population, which could also be associated with changes in the intracellular levels of certain precursors related to general metabolism [51]. A loss of heterologous production of specialized metabolites was previously observed for *E. coli*, where, after the 31st generation, lycopene production decreased from 3.3% of its initial yield [52] due to segregational plasmid instability. Similarly, the use of cyanobacteria as a heterologous production platform faces many challenges, such as a loss of productivity due to the burden or loss of the metabolic pathway and genetic instability [52,53]. In yeast, low expression-cassette instability in large fermentation batches led to a decrease in recombinant protein production [54]. Taken together, these studies show that a loss of production seems to occur elsewhere in other systems, as a result of different underlying problems. In addition, these further demonstrate that like other systems, diatoms need further optimization of the metabolic pathway-encoding constructs discussed below to ensure a suitable platform for the heterologous production of specialized metabolites. Recently, cloning the *Metarhizium anisopliae ARSEF23* highly reducing PKS enzyme (*MaOvaA*), in addition to a non-reducing PKS (*MaOvaB)* and a domain PKS (*MaOvaC)* consisting of an acyl carrier protein and a thioesterase, in the fungus *Aspergillus nidulans* yielded high titers of OA (80 mg/L) [55]. Thus, it would be interesting to compare the expression and behavior of cannabis enzymes regarding fungal enzymes such as highly reducing and non-reducing PKS, along with thioesterase and acyl carrier enzymes [50]. Another approach to increase OA titers could be the construction of alternative pathways using enzymes found in other OA-producing plants than cannabis, in order to achieve OA and other CB precursors while using fungal enzymes or mutation-enhanced enzymes. It is possible, from an evolutionary point of view, that other sources of transgenes are more suitable for heterologous expression in diatoms.

Moreover, in this study, low OA yield could be explained by the detection of cannabinoid-like compounds in higher titers (Figure A4) similar to those observed in other studies, where flavonoid precursors were consumed to form flavonoids in bioengineered yeast [56]. Further steps of compartmentalization of the products would be an interesting strategy, given the possible cytotoxicity of the cannabinoids in many systems [12,57].

The application of *P. tricornutum* as a production platform for OA raises intriguing questions about its metabolic compatibility with the broader CB biosynthesis pathway. Further research is warranted to understand how the engineered diatom handles the downstream conversion of OA into various cannabinoids. Additionally, optimizing the growth conditions, such as light intensity, nutrient availability, and cultivation scale, will be crucial for maximizing production yields [51,58,59]. The successful synthesis of olivetolic acid in *P. tricornutum* opens the door to a range of future possibilities. Further pathway engineering could enable the production of specific compounds, offering a renewable and controllable source of cannabinoids for medical and industrial applications. The engineered diatom strains could be integrated into larger bioprocessing systems, harnessing their natural growth advantages for cost-effective production.

4. Materials and Methods

4.1. Microbial Strains and Growth Conditions

Saccharomyces cerevisiae VL6-48 (ATCC MYA-3666: MATα his3-Δ200 trp1-Δ1 ura3-52 lys2 ade2-1 met14 cir0) was used for the yeast assembly, as described previously [38]. Positive yeast strains containing the His selection were grown on minimal yeast media without histidine. A pool of the grown yeasts was harvested 5 days after assembly and the total DNA was extracted, as described previously [60]. The assembled plasmids were then transformed via electroporation and amplified in *Escherichia coli* (Epi300, Epicenter, Biosearch Technologies, Guelph, ON, Canada), which was grown on Luria Broth (LB) media supplemented with the appropriate antibiotic chloramphenicol (25 mg·L^{-1}) overnight at 37 °C. The plasmids were then extracted from the chloramphenicol-*E. coli* colonies, which were tested by cPCR using a miniprep kit, allowing the extraction of large vectors (EZ10 DNA miniprep kit, Bio Basic Inc., Markham, ON, Canada).

4.2. Diatom Transformation, Selection, and Subculturing

The plasmids were verified by next-generation sequencing and were then amplified in an *E. coli* Epi 300 strain containing the pTA-MOB plasmid to enable conjugation with wild-type diatoms, as described in the literature [38]. *P. tricornutum* (Culture Collection of Algae and Protozoa, CCAP 1055/1), was kindly provided by Prof. Bogumil Karas. The transfer of plasmid DNA to *P. tricornutum* via conjugation from *E. coli* was performed as described by Karas et al. [38]. For this process, 250 µL of wild-type *P. tricornutum* culture was adjusted to a density of 10^8 cells/mL, while the *P. tricornutum* cell density was obtained by plating 1 mL of wild-type *P. tricornutum* on ½ × L1 1% agar plates and growing them at 18 °C under cool fluorescent lights (75 µE $m^{-2}s^{-1}$) on a light/dark cycle of 16/8 h for 4 days. Prior to transformation, 1 mL of L1 media was added to each agar plate, and the cells were first scraped and then harvested via pipetting in a sterile tube. The cells were then diluted and mounted in an improved Neubauer hemacytometer (BLAUBRAND® counting chamber, MilliporeSigma Canada Ltd., Oakville, ON, Canada) to be counted, then the cell concentration was adjusted to 5.0 × 10^8 cells mL^{-1}. A volume of 50 mL of *E. coli* culture containing the assembled and pTA-MOB plasmids was grown at 37 °C under agitation to reach an OD_{600} of 0.9, then the volume was centrifuged at 3000× *g* for 10 min and resuspended in 500 µL of SOC media. Conjugation was initiated by adding 200 µL of *P. tricornutum* to 200 µL of *E. coli* cells. The cell mixture was then plated on ½ L1, 5% LB, 1% agar plates, incubated at 30 °C for 90 min in the dark, and then transferred to 18 °C in the light and grown for 48 h. Two days later, 1 mL of L1 media was added to the plates and the cells were collected by scraping, then a volume of 200 µL of cells was plated on ½ L1, 1% agar plates supplemented with zeocin at 50 µg mL^{-1} for selection and were incubated at 18 °C under light (75 µE $m^{-2}s^{-1}$). Two weeks later, positive colonies appeared and were streaked again on ½ L1, 1% agar plates supplemented with zeocin at 50 µg mL^{-1} for the verification of plasmid stability. The recombinant colonies of *P. tricornutum* were checked via cPCR. Four positive colonies were retained from each construct, from which the plasmids were extracted for further confirmation. The extracted vectors were then

amplified in *E. coli* Epi300; these were extracted using the EZ10 kit and sequenced using next-generation sequencing via Illumina MiSeq technology at the Massachusetts General Hospital's Center of Computational and Integrative Biology (MGH CCIB DNA Core, Boston, MA, USA). In parallel, transformed *P. tricornutum* culture was launched from a single colony until obtaining an OD_{680} nm of 0.06–0.1. The bioengineered *P. tricornutum* strains were then sub-cultured every 10 days in a 1/3 volume ratio of fresh L1 Si^{-} medium containing zeocin (50 mg·L^{-1}).

4.3. Transgene Sequences Validation

C. sativa tetraketide synthase (*Cs*TKS) and olivetolic acid cyclase (*Cs*OAC) amino acid sequences were obtained from *C. sativa* through Genbank, under the respective accession numbers of ACD76855.1 and JN679224.1 (Appendix A, Table A1). The amino acid sequence was then converted into a DNA sequence with respect to *P. tricornutum*'s codon usage table, which is accessible through NCBI Genbank, as well as the HIVE laboratory's codon usage table (CUT) [41]. Several cassettes containing the genes of interest were designed; in this article, we describe three of them. (1) *Pt*OA1 was a construction where TKS and OAC were linked by a self-cleavable sequence from the *Thosea asigna* virus (T2A) [45] without tags, to ensure a similar enzymatic site exposition as that seen in *C. sativa*. (2) *Pt*OA2 is a construction where TKS and OAC were tagged, respectively, in the C-terminal with 6xHis and c-Myc tags for possible protein detection or purification and were linked by a T2A sequence. (3) *Pt*OA3 is a construction designed to present TKS coupled with a reporter gene fluorescent protein eYFP on the N-terminal fused by $(GlyGlyGlyGlySer)_3$ to OAC, which is tagged with c-Myc on the C-terminal. Each inserted construction was designed to be expressed under the strong constitutive promoter 40SRPS8 and the FcpA terminator in a p*Pt*GE30 backbone vector [61] that contained the zeocin resistance gene *ShBle* as a selection marker. The plasmid p*Pt*GE30 contains a centromeric yeast fragment, allowing it to remain extrachromosomal [38]. All gene sequences were codon-optimized for *P. tricornutum*'s optimal codon usage, with a GC content ranging from 48 to 55%. A representation of each plasmid design is detailed in Figure 1. The primers used for the gene amplification and DNA fragment assembly are listed in Appendix A, Table A1. The fragments and transgene sequences that have been optimized for higher expression in *P. tricornutum* are detailed in Appendix A, Table A2. The sequencing results of the plasmid DNA obtained from *P. tricornutum* clones described in this study are detailed in Appendix A, Table A3.

4.4. Heterologous Protein Detection

Protein detection in the positively transformed strains of *P. tricornutum* and the negative control (wild type, transformed with p*Pt*GE30) was performed on 50 mL (OD_{680} nm 0.1 a.u.) cultures that were 6 days old with OD_{680} nm values between 0.9 and 1.7 a.u. The cells were centrifuged at 4000 rpm for 20 min at 4 °C. The resulting pellets were weighed and resuspended in loading buffer 1× (0.8% sodium dodecyl sulfate, 0.05 M Tris-HCl pH 6.8, 6% glycerol, bromophenol blue, and 3.2% β-mercaptoethanol) to a final concentration of 500 mg FW·mL^{-1}. The cell lysates were boiled at 95 °C for 5 min and were then centrifuged at maximum speed for 5 min at room temperature. A volume of 35 µL of the supernatant was loaded into 12% SDS-PAGE gel and migrated at 80 volts until the proteins passed through to the stacking gel, then the voltage was raised to 120 volts. The gel was then transferred using the BioRad Trans-Blot Turbo Transfer system at 2.5 A for 15 min. The blot was equilibrated in 1× TBS solution and blocked with 5% milk for 1 h at room temperature, before washing it three times with TBST solution and adding the primary antibody overnight at 4 °C. In this study, the various primary antibodies used for the T2A sequence were purchased from Sigma Aldrich (cat. #ABS31, Boston, MA, USA), the 6X-His Tag Monoclonal antibody from Thermofisher (cat. #MA1-21315, Waltham, MA, USA), and anti-GFP/CFP/YFP from Cedarlane (cat. #CLH106AP, Burlington, ON, Canada), all at a 1:1000 dilution in 3% BSA. After three washes with TBST solution, the *Pt*OA1 sample blot was incubated for 1 h in a 1:20,000 dilution in 5% milk of the Immun-Star Goat Anti-Rabbit

(GAR)-HRP conjugate from Bio-Rad (Mississauga, ON, Canada, cat. #1705046). Meanwhile, the *Pt*OA2 and *Pt*OA3 blots were incubated under the same conditions with Immun-Star Goat Anti-Mouse (GAM)-HRP conjugate from BioRad (cat. #1705047, Hercules, CA, USA). The multiple Tag protein (GenScript, cat. # M0101, Piscataway, NJ, USA) and the recombinant YFP purified protein (10 ng) from *E. coli* were used as the positive controls. After three washes with TBST solution, protein detection was performed using the Clarity Max Western ECL Substrate-Luminol solution from Bio-Rad (cat #1705062S, CA, USA). The blots and Coomassie-stained gels were visualized using a ChemiDoc Imaging System with Image Lab Touch software 2.4 (Bio-Rad, cat #12003153, Hercules, CA, USA) and Image Lab™ 5.2 software (Bio-Rad, cat # 1709690, Hercules, CA, USA).

4.5. Subcellular Localization of YFP

Live cell images of the four-day-old culture were captured using a Leica SP8 confocal laser microscope (Leica, Wetzlar, Germany) with an HCX PL APO 60×/1.25–0.75 Oil CS objective. The excitation of YFP and chlorophyll fluorescence occurred at 488 nm with a 65 mW argon laser. YFP fluorescence emission was detected between 552 and 560 nm, whereas chlorophyll fluorescence was detected at a bandwidth of 625–720 nm. Bright-field light microscopy images were also taken. The images were analyzed using the Fiji software (https://imagej.net/software/fiji/) for Windows (64-bit) [62].

4.6. Flow Cytometry Analysis

The BD FACS Melody device (BD Biosciences, La Jolla, CA, USA), equipped with blue (488 nm), red (640 nm), and violet (405 nm) lasers, and a Beckn Cytoflex FC500 equipped with Argon (488 nm) and HeNe (633 nm) lasers were used to measure YFP fluorescence on the FITC channel (527/32 and 525/15 nm) and chloroplast autofluorescence on the APC channel (660/10 nm and 675/15 nm, respectively). The *P. tricornutum* samples were diluted to an OD_{680} of 0.1 and were then analyzed at a fixed flow rate of 1 for at least 10,000 events per sample, conducted with 3 replicates. The diatoms were first gated using side-scattered light (SSC) versus forward scatter plots to determine the targeted population and were then gated using the chlorophyll levels (650 nm) (Figure 4). All data acquisition and analysis were carried out with the BD FlowJo version 10 software (BD Biosciences, La Jolla, CA, USA, 2020). All FACS experiments were conducted in triplicate.

4.7. Metabolite Extraction Method

The wild-type and positive transconjugant strains were pre-cultured in L1 medium, as mentioned previously. Approximately 100 mg of wet biomass from each culture was harvested by centrifuging 50 mL of culture at 1500× *g* and at 4 °C for 10 min. The pellets were resuspended in 5 mL of ethanol 95% (cat.# P016EA95), vortexed 5 times for 1 min each time, and stored at −20 °C overnight. The metabolite extracts were separated from the cells by centrifuging at 1500× g and at 4 °C for 10 min. The supernatants were transferred to clean tubes and ethanol fractions were evaporated via a speedvac vacuum concentrator (Thermo Savant SPD 2010) over 3 h at ramp level 5, without heat. The dried extracts were reconstituted in 250 µL of mobile phase solution, consisting of formic acid at 0.1% *v*/*v* in methanol and water (85:15), for the purposes of HPLC analysis. After homogenization, 200 µL of the extract was transferred to an HPLC vial with a 300 µL adaptor.

4.8. HPLC Detection Method for Cannabinoid Precursors

Olivetolic acid (CAS 491-72-5) and olivetol (CAS 500-66-3) standards were purchased from Santa Cruz Biotechnologies (Dallas, TX, USA). Extracts from *P. tricornutum* transconjugant strains and the wild type were then analyzed via HPLC (Agilent 1260 Infinity II) with a DAD detector. The column used was a ZORBAX Eclipse XDB-C18, 80Å, 4.6 × 250 mm, and 5 µm (Agilent; PN: 990967-902, Missisauga, ON, Canada). The sampler harvested and injected 10 µL of the sample into the system; the analysis conditions were the following: the temperature applied to the column was 30 °C, while the isocratic mobile phase comprised

methanol at 85% and formic acid at 0.1%, mixed with 15% pure water at a flow rate of 0.4 mL per minute. The total run time of the analysis was 70 min per sample. The analyzed wavelengths were chosen based on the maximal peak of each standard (220, 230, and 280 nm).

4.9. Mass Spectrum Validation

All the LC-MS analyses in this study were performed at the Centre National en Électrochimie et en Technologies Environnementales, Inc. (CNETE, Shawinigan, QC, Canada). To validate the presence of the compounds of interest, the samples were analyzed under the same LC conditions but used the MS detector, achieving this by diluting each sample 10 times in methanol at LC-MS grade containing 0.1% formic acid and injecting the samples into a Thermo Scientific UPLC Dionex Ultimate 3000 MS LTQ XL UPLC-MS system equipped with a ZORBAX Eclipse XDB-C18 at 80 Å, 4.6 × 250 mm, and 5 µm (Agilent; PN.: 990967-902). The MS detector was set to analyze five scan events during the same period of analysis, taking one SIM event for each compound and one for large-scale TIC analysis (the event parameters are detailed in Appendix A, Figure A5). The analysis conditions used were the following: the temperature applied to the column was 30 °C and the isocratic mobile phase used comprised methanol 85% and formic acid 0.1%, mixed with 15% pure water at a flow rate of 0.4 mL per minute. The total run time of an analysis was 70 min per sample. The sample injection volume was 2.5 µL. The events observed at 220 nm were analyzed for 5 main events via the MS spectra corresponding to each of the main cannabinoids, while fragmentation was compared to the corresponding standard and theoretical fragmentation using the software program Mass Frontier 7.0.

4.10. Precursor Supplementation Assay

P. tricornutum's empty vector (p*Pt*GE30) was grown for a full cycle from an established liquid culture until the OD_{680} was greater than 2 a.u. Afterward, six new subcultures were made with a starting inoculum of 0.1 OD_{680} in L1 media and then supplemented with 50 mg/L of zeocin for 9 days. The final volume of all subcultures was around 360 mL. Then, this volume was distributed as nine 30 mL cultures and placed into 50 mL Falcon tubes. The first set of triplicates was supplemented with ethanol as a negative control, the second set was supplemented with 1 mM of geranyl diphosphate ammonium salt (Sigma, CAS no 763-10-0) and 0.45 mM of olivetolic acid dissolved in L1 media, and the third set was supplemented with 0.27 mM of CBGA diluted in ethanol. Along with the p*Pt*GE30 subcultures, this assay was performed on at least two other recombinant strains containing at least one of the *Pt*OA cassettes. The Falcon tubes were returned to normal growth conditions for 16 h; the next day, each culture was centrifuged to separate the supernatant, then the pellet was washed and the washing liquid was kept for analysis. The metabolites were extracted from the cell pellets, as described in Section 4.7. The supernatant, washes, and cell extracts were analyzed via HPLC, and the signals of interest were confirmed later via MS analysis.

5. Conclusions

In conclusion, the bioengineering of *P. tricornutum* with cannabis genes represents a significant advancement in CB production technology. This innovative approach not only provides a sustainable alternative to traditional cannabis cultivation but also highlights the potential of marine microorganisms as versatile bioreactors for valuable natural products. As synthetic biology and metabolic engineering continue to evolve, the intersection of marine biology and cannabis biosynthesis could shape the future of CB production.

Author Contributions: F.A., E.I.F., N.M., F.M.-M. and I.D.-P. contributed to the study's conception and design. Methodology, material preparation, data collection, and analysis were performed by F.A., E.I.F., M.H., A.M.D.-G., N.M., A.C., S.-E.G. and F.M.-M. Episomal constructions, diatom conjugation, and molecular biology experiments were planned and conducted by F.A., E.I.F., A.M.D.-G., J.L., N.M., F.M.-M., B.J.K. and I.D.-P. HPLC analyses were performed by M.H., A.C. and S.-E.G., whereas all LC-MS analyses were performed by J.-F.L. The first draft of the manuscript was written by F.A., E.I.F., A.M.D.-G. and N.M. F.M.-M. and I.D.-P. edited the previous versions of the manuscript. I.D.-P. supervised the project and secured the funding. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Sciences and Engineering Research Council of Canada through the Alliance Program (award nos. ALLRP 554429-20 and ALLRP 570476-2021) to I.D.-P. Additional support in the form of scholarships was awarded to F.A. from MITACS-Elevation IT15971, to E.I.F. and A.M.D.-G. from the MITACS-Acceleration program (grant nos. IT12310, IT16463, and IT19432), and to I.D.-P. is also acknowledged. B.J.K. is supported by Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-06172).

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors would like to acknowledge Patrik Quessy from the Centre National en Électrochimie et en Technologies Environnementales Inc. for LC-MS analyses, and Mélodie B. Plourde, Pamela Lavoie, and Hugo Germain from the University du Québec à Trois-Rivières for their guidance, mentoring, and assistance.

Conflicts of Interest: The authors declare that this study received funding from Algae-C. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Abbreviations

Au	Absorbance units
CB	Cannabinoid
CBD	Cannabidiol
CBDA	Cannabidiolic acid
CBGA	Cannabigerolic acid
cPCR	Colony polymerase chain reaction
FACS	Fluorescence-activated cell sorting
Gly	Glycine
GPP	Geranyl pyrophosphate
His	Histidine
OA	Olivetolic acid
OAC	Olivetolic acid cyclase
OD	Optical density
Ser	Serine
T2A	*Thosea asigna* virus self-cleavable sequence
THC	Δ-9-tetrahydrocannabinol
THCA	Delta (9)-tetrahydrocannabinolic acid
TKS	Tetraketide synthase
YFP	Yellow fluorescent protein

Appendix A

Our study showed the construction of three episomes in the p*Pt*GE30 backbone containing *C. sativa* genes for the production of OA (Appendix A, Figure A1 and Tables A1–A3). Each episome was inserted into *P. tricornutum* via bacterial conjugation and the corresponding enzymes were successfully accumulated (Appendix A, Figure A2 and Table A4).

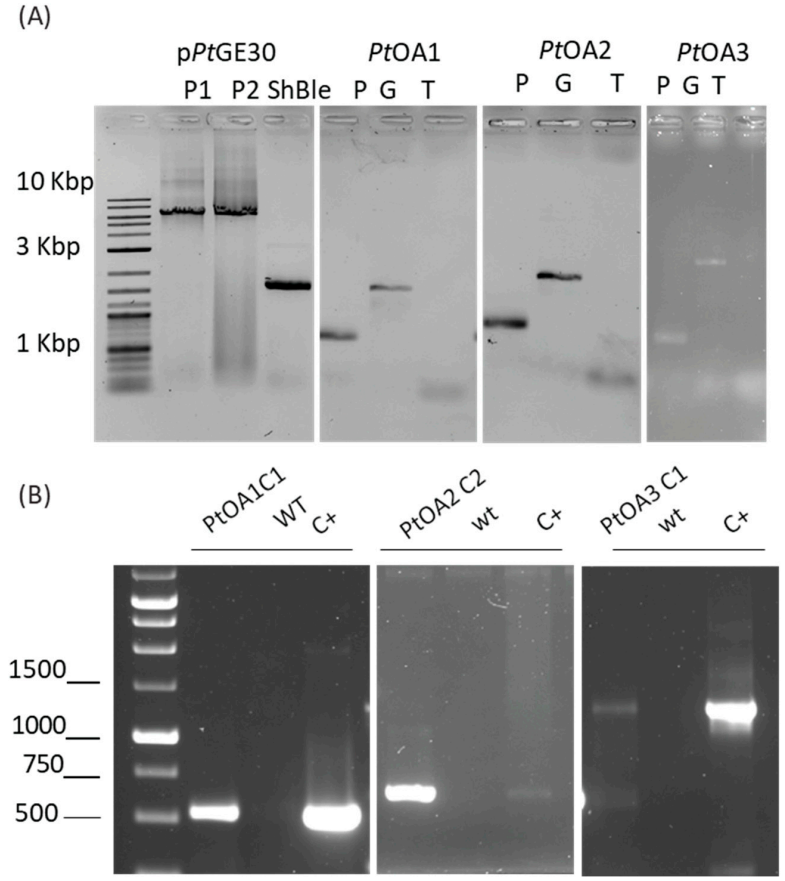


Figure A1. *P. tricornutum* plasmid fragments for yeast assembly and the colony PCR of recombinant *P. tricornutum* after bacterial conjugation. (**A**) Agarose gel electrophoresis of the amplified fragments (the backbone moieties (P1 and P2), promoter (P), gene (G), terminator (T), and selection marker *ShBle* encoding for zeocin resistance) for the plasmid constructs *Pt*OA1 and *Pt*OA2. (**B**) Agarose gel (2%) electrophoresis of the colonies. PCR was performed on the *P. tricornutum* colonies for each construct, with the cPCR primers from Appendix A, Table A4.

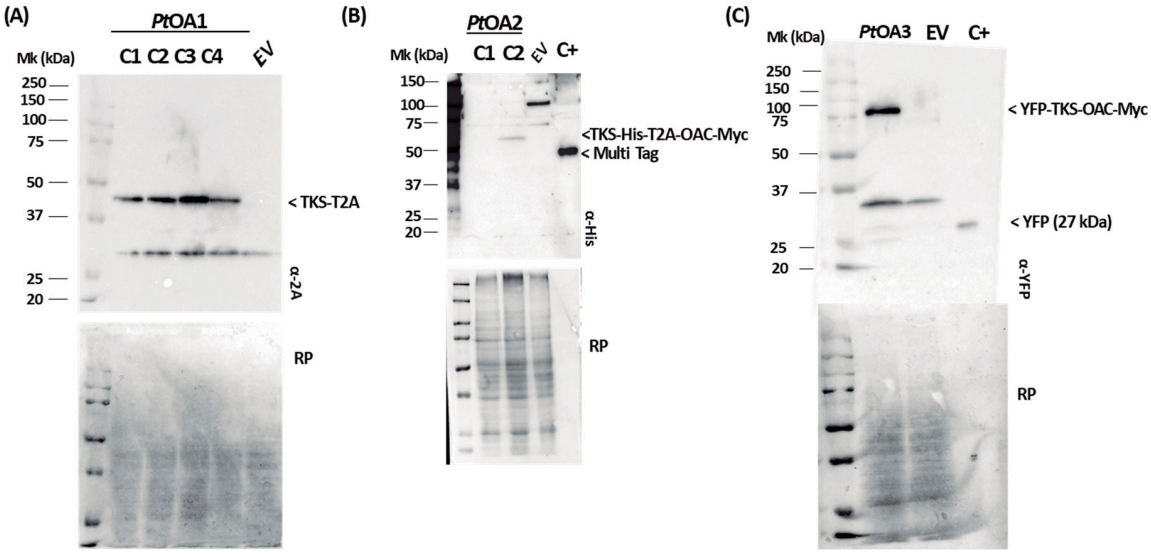


Figure A2. Western blot membranes and the Ponceau red coloration of gels of (**A**) *Pt*OA1, (**B**) *Pt*OA2, and (**C**) *Pt*OA3. The membranes were obtained from different 12% SDS gels, and the antibodies used are described in the main text.

Table A1. The amino acid sequences of *Cs*TKS and *Cs*OAC used in this study to build the *Pt*OA1, *Pt*OA2, and *Pt*OA3 cassettes. CDS: coding sequence.

Sequence Name	Amino Acid Sequence	Length (a.a.)	Accession Number (CDS)	Accession Number AA Seq
C. sativa type 3 tetraketide synthase (TKS)	MNHLRAEGPASVLAIGTANPENILLQDEFPDYYFRVTKSEHMTQLKEKFRKICDK SMIRKRNCFLNEEHLKQNPRLVEHEMQTLDARQDMLVVEVPKLGKDACAKAIKEWGQPKS KITHLIFTSASTTDMPGADYHCAKLLGLSPSVKRVMMYQLGCYGGGTVLRIAKDIAENN KGARVLAVCCDIMACLFRGPSESDLELLVGQAIFGDGAAAVIVGAEPDESVGERPIFELVSTG QTILPNSEGTIGGHIREAGLIFDLHKDVPMLISNNIEKCLIEAFTPIGISDWNSIFWITH PGGKAILDKVEEKLHLKSDKFVDSRHVLSEHGNMSSSTVLFVMDELRKRSLEEGKSTTGDGFE WGVLFGFGPGLTVERVVVRSVPIKY	385	EU551165.1	ACD76855.1
C. sativa olivetolic acid cyclase (OAC)	MAVKHLIVLKFKDEITEAQKEEFFKTYVNLVNIIPAMKDVYWGKDVTQKNKEEGYTHIVEV TFESVETIQDYIIHPAHVGFGDVYRSFWEKLLIFDYTPRK	101	NM_001397939.1	NP_001384868.1

Table A2. List of the DNA sequence parts for the plasmid constructs, with the corresponding references and accession numbers. The OA1, OA2, and OA3 parts are color-coded as follows: *Cannabis sativa tetraketide synthase* CsTKS (AB164375), *Cannabis sativa olivetolic acid cyclase* CsOAC (JN679224), self-cleavable sequence from the *Thosea asigna* virus T2A, c-Myc and 6His tags, 3(GGGGS) fusion, and the yellow fluorescent protein, YFP. All the mentioned sequences are codon-optimized for the best *Phaeodactylum tricornutum* usage.

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
p*Pt*GE30-part 1	CAGGGTAATATAGATCTTCCGCTGCATAACCCTGCTTCGGGGTCATTATAGCGATTTTTTCGGTATATCCATCCTTTTTCGCA CGATATACAGGATTTTGCCAAAGGGTTCGTGTAGACTTTCCTTGGTGTATCCAACGGCGTCAGCCGGGCAGGATAGGTGAAGTAG GCCCACCCGCGAGCGGGTGTTCCTTCTTCACTGTCCCTTATTCGCACCTGGCGGTGCTCAACGGGAATCCTGCTCTGCGAGGCTG GCCGGCTACCGCCGGCGTAACAGATGAGGGCAAGCGGATGGCTGATGAAACCAAGCCAACCAGGAAGGGCAGCCCACCTATCAAG GTGTACTGCCTTCCAGACGAACGAAGAGCGATTGAGGAAAAGGCGGCGGCGGCCGGCATGAGCCTGTCGGCCTACCTGCTGGCCG TCGGCCAGGGCTACAAAATCACGGGCGTCGTGGACTATGAGCACGTCCGCGAGCTGGCCCGCATCAATGGCGACCTGGGCCGCCT GGGCGGCCTGCTGAAACTCTGGCTCACCGACGACCCGCGCACGGCGCGGTTCGGTGATGCCACGATCCTCGCCCTGCTGGCGAAG ATCGAAGAGAAGCAGGACGAGCTTGGCAAGGTCATGATGGGCGTGGTCCGCCCGAGGGCAGAGCCATGACTTTTTTAGCCGCTAA AACGGCCGGGGGGTGCGCGTGATTGCCAAGCACGTCCCCATGCGCTCCATCAAGAAGAGGCACTTCGAGCTGTAAGTACATCACC GACGAGCAAGGCAAGACGATCCGCGCCTGTTTTATTGAGAACGTTGTTCGTGTTGGCCTCAATGGTAGCGATGCGTCATTCAGCG AAGTTCTGGTGCTGATGATGTGGTTCGCTTTGCCACTGGTCAATGTGGTAAGCCCGTGTAATGTCAGTAACCTTTTTACTGATCT CAGCTTGAGCACGGTCGCTGATGAGCTTATCCATGGCCCCACGGTAACGGATATGATCCTCTAGGGCGTTGACAAATTCTTTGTC GGTTTTCATGGGGTAGACGTCAGTGACCAGGGAACGTCTCCCTACAAAAAGTTGCGCATACTTTGCTCCGTTGTCGACGGCAGGG GTGTCGGACCAGATTGTATCTGTGGCAACGGCCTCATTGCGTCAATGGACCTTTAAAGCCGGGAAACGAGACTTGAAATGTTTAC GTAGAGGTGCATTATAGACCTCACGCGCATATTGAGTGGTTGCAAAAATGGTTTTGCGAACGGTATCACTGGAGACCCATGCCAG GCAAGGACGGAGGGCATCATAATCATGTTCATTGCGTTGGACAGCATGCTTGTTACACGTAAGTATACGGTCGAGAGCATCGTGA CGTTGAGAATTACAGATGGCAACGTGTCGGTGCATGTATTGGCCGAGATTGTCAAATCGTGGCTCAATATATGGCATGCCAGGTA AGTCATGGATGTCGGTATGCCATTCAGCCTTCATGTCGATTTCGTTGTCGAGAACTGATGGGTCCCAGTCAACGTCAGAAGTAAG AACAACATGGGGAAGGGAATGAAGTTCATGACTGGTGGGTGCGCGCATATCCATGTACGGAAGACCCTGTCGTATGTTTAGTGGG ATGATGTAATCATCCAGGGTAACAATACGTTGGTTGCCGCCCACGGTACGAGACCGATCTTGAACACAATTGTGGCAATGTTCAA	Slattery et al., 2018 [61]

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
p*Pt*GE30-part 1	GTTGGGCGCTGGAATGGATGGTTTTACCTTTACCAAGATGCGCGTATTGATGCATGATAACAACAATGGGCCCTCGTTGCGATTC GACAAGACCGGCAGCCGTGACAATGTCTAGATTGGACAAGGTATGTTCATTTATACCGGTTATGTTGGCTGAACGTCCAGTTTTG TGCAGTACGGTAACATCACTCCCCGCAAGTCCACCGTTGGCACCGCGATCGACAAGTGCAGATGTGGTATTTTGTACTGTGTGTCG GGAGACTTGGTACTGTAACACATTTGACTGAAGTGACTTGGTGGTACCATGAGGACGGAGGGAACCCGTTGATACGGTGTTGCGT GACGCTGCAAGTACCTTACGTATGTCGCCTGGATCCATACGTCCAACCCGATCAGAAAGATGTGCTAGGAGTTCCGTTTCTGGTG CACAATCATGAAAGGTATCCACGGGAGTTCCATTGTTGTCGTCCGGGGCTGCGTCGCGACTGCCAGTTGCATGTACATGCGCCAT CAAGGGACGTGCCAATCCATTACTAACCGGGAGAACCTTAGCTGGGGCAGCCAATCCCTGAAGTATAGCCTTAGCATCATCTGAG AGCTGGTTCCACATATCTCTAGGGATATAAGGTCGTTCACGGTTAGTGGGCGTACTATGATGTCGGTTATTAGTACAATTCTCCC TGCGGGCATGCGCATTGGCTTCATAGAGTACGGAAGGTGACAAGTCAATATTGTAGTTAACATCCGGATCAGTGTCAAAGTCTGT AGGATGGTAAGAAAGATCAGTAGAATGAATACTACGCTTTTCCTTGGGACTACGGGAATTAGAGAAGTTGTTTCCTTTATTGTAG AGTGACGCTGAAGCAAGTAGAAGACTAAGGTAACTCTCGTAGCTAATAGGATTACCTCCTTTGGCTAGGTCAAGAGTGGCTGTGA TCTTCACTTGACAAAGTTCCGGTACATTGTGGACAGCATTCTCCAAAAGACTAAGACACAGTTGCTTTGGGAGTTGCTCAGCCAT TGGTACAGTATTGTGGTAGATACAAAGGTGGTTCTTCCAATGAAGGATAAATCCTTCTGCTGTGCCTGTCCATGAGGATCCATAT TTCGCCGTAGTTAGGTAACCAAGCGTGGTGGCTGAACTGATCTTCGCACTTGCTGATTCCGTATAGTGTTTGACAACTTTACAAA ACACTTCTTGCGCAGTTCGCTCTAGCCTGGACAGAGAAGGGACAATCTTCGTCGTTAAGACTCGTGCGAATAGCAAAAGATCACA AAATAGCACATCGGCACCGACCAACGATTATTTCCAAGGAAAAAAAGAATGCTTCACTACAAGAAATTGTGTCATCCCTATACAG AGTCTTGTTACTGTGACAGAAAATTGATGGAAGATGTGGCGGATTGCCTTTACACTAGCCAACTTGTTCGACTAATTGCAGCTTCT TCTGAGAGGCTTCACCGAGTAACGCGAAGAACACCGGTGTCTCGTACATGCTCGTCGGTGAACGCTCGTCCAATGACACCCCCCA CTTTGTATCAATATCCCAACTTGGTAGTGAACTGGAATGATACATGCAATTTCGCGCCGCATCAACAGCCACGGGCACCATCGACG AATAGACTCGGTCGAGCTGGTTGCCCTCGCCGCTGGGCTGGCGGCCGTCTATGGCCCTGCAAACGCGCCAGAAACGCCGTCGAAGC CGTGTGCGAGACACCGCGGCCGGCCGCCGGCGTTGTGGATACCTCGCGGAAAACTTGGCCCTCACTGACAGATGAGGGGCGGACGT TGACACTTGAGGGGCCGACTCACCCGGCGCGGCGTTGACAGATGAGGGGCAGGCTCGATTTCGGCCGGCGACGTGGAGCTGGCCAG CCTCGCAAATCGGCGAAAACGCCTGATTTTACGCGAGTTTCCCACAGATGATGTGGACAAGCCTGGGGATAAGTGCCCTGCGGTAT TGACACTTGAGGGGCGCGACTACTGACAGATGAGGGGCGCGATCCTTGACACTTGAGGGGCAGAGTGCTGACAGATGAGGGGCGCA CCTATTGACATTTGAGGGGCTGTCCACAGGCAGAAAATCCAGCATTTGCAAGGGTTTCCGCCCGTTTTTCGGCCACCGCTAACCTG TCTTTTAACCTGCTTTTAAACCAATATTTATAAACCTTGTTTTTTAACCAGGGCTGCGCCCTGTGCGCGTGACCGCGCACGCCGAAG GGGGGTGCCCCCCCTTCTCGAACCCTCCCGGTCGAGTGAGCGAGGAAGCACCAGGGAACAGCACTTATATATTCTGCTTACACACG ATGCCTGAAAAAACTTCCCTTGGGGTTATCCACTTATCCACGGGGATATTTTTATAATTATTTTTTTTATAGTTTTTAGATCTTCT TTTTTAGAGCGCCTTGTAGGCCTTTATCCATGCTGGTTCTAGAGAAGGTGTTGTGACAAATTGCCCTTTCAGTGTGACAAATCACC CTCAAATGACAGTCCTGTCTGTGACAAATTGCCCTTAACCCTGTGACAAATTGCCCTCAGAAGAAGCTGTTTTTTCACAAAGTTAT CCCTGCTTATTGACTCTTTTTTATTTAGTGTGACAATCTAAAAACTTGTCACACTTCACATGGATCTGTCATGGCGGAAACAGCGG TTATCAATCACAAGAAACGTAAAAATAGCCCGCGAATCGTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGAT CAAAAACGTATGCTGTATCTGTTCGTTGACCAGATCAGAAAATCTGATGGCACCCTACAGGAACATGACGGTATCTGCGAGATCCA TGTTGCTAAATATGCTGAAATATTCGGATTGACCTCTGCGGAAGCCAGTAAGGATATACGGCAGGCATTGAAGAGTTTCGCGGGGA AGGAAGTGGTTTTTTATCGCCCTGAAGAGGATGCCGGCGATGAAAAAGGCTATGAATCTTTTCCTTGGTTTATCAAACGTGCGCAC AGTCCATCCAGAGGGCTTTACAGTGTACATATCAACCCATATCTCATTCCCTTCTTTATCGGGTTACAGAACCGGTTTACGCAGTTT CGGCTTAGTGAAACAAAAGAAATCACCAATCCGTATGCCATGCGTTTATACGAATCCCTGTGTCAGTATCGTAAGCCGGATGGCTCA GGCATCGTCTCTCTGAAAATCGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGCCTGACTTCCGCCGCCG CTTCCTGCAGGTCTGTGTTAATGAGATCAACAGCAGAACTCCAATGCGCCTCTCATACATTGAGAAAAAGAAAGGCCGCCAGACGA CTCATATCGTATTTTCCTTCCGCGATATCACTTCCATGACGACAGGATAGTCTGAGGGTTATCTGTCACAGATTTGAGGGTGGTTCG	Slattery et al., 2018 [61]

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
p*Pt*GE30-part 1	TCACATTTGTTCTGACCTACTGAGGGTAATTTGTCACAGTTTTGCTGTTTCCTTCAGCCTGCATGGATTTTCTCATACTTTTTGAA CTGTAATTTTTAAGGAAGCCAAATTTGAGGGCAGTTTGTCACAGTTGATTTCCTTCTCTTTCCCTTCGTCATGTGACCTGATATC GGGGGTTAGTTCGTCATCATTGATGAGGGTTGATTATCACAGTTTATTACTCTGAATTGGCTATCCGCGTGTGTACCTCTACCTGG AGTTTTTCCCACGGTGGATATTTCTTCTTGCGCTGAGCGTAAGAGCTATCTGACAGAACAGTTCTTCTTTGCTTCCTCGCCAGTT CGCTCGCTATGCTCGGTTACACGGCTGCGGCGAGCATCACGTGCTATAAAAATAATTATAATTTAAATTTTTTAATATAAATATA TAAATTAAAAATAGAAAGTAAAAAAAAGAAATTAAAGAAAAAATAGTTTTTGTTTTCCGAAGATGTAAAAGACTCTAGGGGGATCG CCAACAAATACTACCTTTTACCTTGCTCTTCCTGCTCTCAGGTATTAATGCCGAATTGTTTCATCTTGTCTGTGTAGAAGACCAC ACACGAAAATCCTGTGATTTTACATTTTACTTATCGTTAATCGAATGTATATCTATTTAATCTGCTTTTCTTGTCTAATAAATATA TATGTAAAGTACGCTTTTTGTTGAAATTTTTTAAACCTTTGTTTATTTTTTTTTCTTCATTCCGTAACTCTTCTACCTTCTTTATT TACTTTCTAAAATCCAAATACAAAACATAAAAATAAATAAACACAGAGTAAATTCCCAAATTATTCCATCATTAAAAGATACGAGG CGCGTGTAAGTTACAGGCAAGCGATCCTAGTACACTCTATATTTTTTTATGCCTCGGTAATGATTTTCATTTTTTTTTTCCACCTA GCGGATGACTCTTTTTTTTTCTTAGCGATTGGCATTATCACATAATGAATTATACATTATATAAAGTAATGTGATTTCTTCGAA GAATATACTAAAAAATGAGCAGGCAAGATAAACGAAGGCAAAGATGACAGAGCAGAAAGCCCTAGTAAAGCGTATTACAAATGA AACCAAGATTCAGATTGCGATCTCTTTAAAGGGTGGTCCCCTAGCGATAGAGCACTCGATCTTCCCAGAAAAAGAGGCAGAAGCA GTAGCAGAACAGGCCACACAATCGCAAGTGATTAACGTCCACACAGGTATAGGGTTTCTGGACCATATGATACATGCTCTGGCCA AGCATTCCGGCTGGTCGCTAATCGTTGAGTGCATTGGTGACTTACACATAGACGACCATCACACCACTGAAGACTGCGGGATTGC TCTCGGTCAAGCTTTTAAAGAGGCCCTAG	Slattery et al., 2018 [61]
p*Pt*GE30-part 2	TCGAGCTGGTTGCCCTCGCCGCTGGGCTGGCGGCCGTCTATGGCCCTGCAAACGCGCCAGAAACGCCGTCGAAGCCGTGTGCGA GACACCGCGGCCGGCCGCCGGCGTTGTGGATACCTCGCGGAAAACTTGGCCCTCACTGACAGATGAGGGGCGGACGTTGACACTT GAGGGGCCGACTCACCCGGCGCGGCGTTGACAGATGAGGGGCAGGCTCGATTTCGGCCGGCGACGTGGAGCTGGCCAGCCTCGCA AATCGGCGAAAACGCCTGATTTTACGCGAGTTTCCCACAGATGATGTGGACAAGCCTGGGGATAAGTGCCCTGCGGTATTGACAC TTGAGGGGCGCGACTACTGACAGATGAGGGGCGCGATCCTTGACACTTGAGGGGCAGAGTGCTGACAGATGAGGGGCGCACCTAT TGACATTTGAGGGGCTGTCCACAGGCAGAAAATCCAGCATTTGCAAGGGTTTCCGCCCGTTTTTCGGCCACCGCTAACCTGTCTT TTAACCTGCTTTTAAACCAATATTTATAAACCTTGTTTTTAACCAGGGCTGCGCCCTGTGCGCGTGACCGCGCACGCCGAAGGGG GGTGCCCCCCCTTCTCGAACCCTCCCGGTCGAGTGAGCGAGGAAGCACCAGGGAACAGCACTTATATATTCTGCTTACACACGAT GCCTGAAAAAACTTCCCTTGGGGTTATCCACTTATCCACGGGGATATTTTTATAATTATTTTTTTTATAGTTTTTAGATCTTCTT TTTTAGAGCGCCTTGTAGGCCTTTATCCATGCTGGTTCTAGAGAAGGTGTTGTGACAAATTGCCCTTTCAGTGTGACAAATCACC CTCAAATGACAGTCCTGTCTGTGACAAATTGCCCTTAACCCTGTGACAAATTGCCCTCAGAAGAAGCTGTTTTTTCACAAAGTTA TCCCTGCTTATTGACTCTTTTTTATTTAGTGTGACAATCTAAAAACTTGTCACACTTCACATGGATCTGTCATGGCGGAAACAGC GGTTATCAATCACAAGAAACGTAAAAATAGCCCGCGAATCGTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGG GATCAAAAACGTATGCTGTATCTGTTCGTTGACCAGATCAGAAAATCTGATGGCACCCTACAGGAACATGACGGTATCTGCGAGA TCCATGTTGCTAAATATGCTGAAATATTCGGATTGACCTCTGCGGAAGCCAGTAAGGATATACGGCAGGCATTGAAGAGTTTCGC GGGGAAGGAAGTGGTTTTTTATCGCCCTGAAGAGGATGCCGGCGATGAAAAAGGCTATGAATCTTTTCCTTGGTTTATCAAACGT GCGCACAGTCCATCCAGAGGGCTTTACAGTGTACATATCAACCCATATCTCATTCCCTTCTTTATCGGGTTACAGAACCGGTTTA CGCAGTTTCGGCTTAGTGAAACAAAAGAAATCACCAATCCGTATGCCATGCGTTTATACGAATCCCTGTGTCAGTATCGTAAGCC GGATGGCTCAGGCATCGTCTCTCTGAAAATCGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGCCTGAC TTCCGCCGCCGCTTCCTGCAGGTCTGTGTTAATGAGATCAACAGCAGAACTCCAATGCGCCTCTCATACATTGAGAAAAAGAAAG GCCGCCAGACGACTCATATCGTATTTTCCTTCCGCGATATCACTTCCATGACGACAGGATAGTCTGAGGGTTATCTGTCACAGAT TTGAGGGTGGTTCGTCACATTTGTTCTGACCTACTGAGGGTAATTTGTCACAGTTTTGCTGTTTCCTTCAGCCTGCATGGATTTT CTCATACTTTTTGAACTGTAATTTTTAAGGAAGCCAAATTTGAGGGCAGTTTGTCACAGTTGATTTCCTTCTCTTTCCCTTCGTC	Slattery et al., 2018 [61]

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
p*Pt*GE30-part 2	ATGTGACCTGATATCGGGGGTTAGTTCGTCATCATTGATGAGGGTTGATTATCACAGTTTATTACTCTGAATTGGCTATCCGCGT GTGTACCTCTACCTGGAGTTTTTCCCACGGTGGATATTTCTTCTTGCGCTGAGCGTAAGAGCTATCTGACAGAACAGTTCTTCTT TGCTTCCTCGCCAGTTCGCTCGCTATGCTCGGTTACACGGCTGCGGCGAGCATCACGTGCTATAAAAATAATTATAATTTAAATT TTTTAATATAAATATATAAATTAAAAATAGAAAGTAAAAAAAAGAAATTAAAGAAAAAATAGTTTTTGTTTTCCGAAGATGTAAAA GACTCTAGGGGGATCGCCAACAAATACTACCTTTTACCTTGCTCTTCCTGCTCTCAGGTATTAATGCCGAATTGTTTCATCTTGT CTGTGTAGAAGACCACACACGAAAATCCTGTGATTTTACATTTTACTTATCGTTAATCGAATGTATATCTATTTAATCTGCTTTT CTTGTCTAATAAATATATATGTAAAGTACGCTTTTTGTTGAAATTTTTTAAACCTTTGTTTATTTTTTTTTCTTCATTCCGTAACT CTTCTACCTTCTTTATTTACTTTCTAAAATCCAAATACAAAACATAAAAATAAATAAACACAGAGTAAATTCCCAAATTATTCCA TCATTAAAAGATACGAGGCGCGTGTAAGTTACAGGCAAGCGATCCTAGTACACTCTATATTTTTTTATGCCTCGGTAATGATTTT CATTTTTTTTTTCCACCTAGCGGATGACTCTTTTTTTTTCTTAGCGATTGGCATTATCACATAATGAATTATACATTATATAAAG TAATGTGATTTCTTCGAAGAATATACTAAAAAATGAGCAGGCAAGATAAACGAAGGCAAAGATGACAGAGCAGAAAGCCCTAGTA AAGCGTATTACAAATGAAACCAAGATTCAGATTGCGATCTCTTTAAAGGGTGGTCCCCTAGCGATAGAGCACTCGATCTTCCCAG AAAAAGAGGCAGAAGCAGTAGCAGAACAGGCCACACAATCGCAAGTGATTAACGTCCACACAGGTATAGGGTTTCTGGACCATAT GATACATGCTCTGGCCAAGCATTCCGGCTGGTCGCTAATCGTTGAGTGCATTGGTGACTTACACATAGACGACCATCACACCACTG AAGACTGCGGGATTGCTCTCGGTCAAGCTTTTAAAGAGGCCCTAGGGGCCGTGCGTGGAGTAAAAAGGTTTGGATCAGGATTTGC GCCTTTGGATGAGGCACTTTCCAGAGCGGTGGTAGATCTTTCGAACAGGCCGTACGCAGTTGTCGAACTTGGTTTGCAAAGGGAG AAAGTAGGAGATCTCTCTTGCGAGATGATCCCGCATTTTCTTGAAAGCTTTGCAGAGGCTAGCAGAATTACCCTCCACGTTGATT GTCTGCGAGGCAAGAATGATCATCACCGTAGTGAGAGTGCGTTCAAGGCTCTTGCGGTTGCCATAAGAGAAGCCACCTCGCCCAA TGGTACCAACGATGTTCCCTCCACCAAAGGTGTTCTTATGTAGTTTTACACAGGAGTCTGGACTTGACGCTAGTGATAATAAGTG ACTGAGGTATGTGCTCTTCTTATCTCCTTTTGTAGTGTTGCTCTTATTTTAAACAACTTTGCGGTTTTTTGATGACTTTGCGATTT TGTTGTTGCTTTGCAGTAAATTGCAAGATTTAATAAAAAAACGCAAAGCAATGATTAAAGGATGTTCAGAATGAAACTCATGGAA ACACTTAACCAGTGCATAAACGCTGGTCATGAAATGACGAAGGCTATCGCCATTGCACAGTTTAATGATGACAGCCCGGAAGCGA GGAAAATAACCCGGCGCTGGAGAATAGGTGAAGCAGCGGATTTAGTTGGGGTTTCTTCTCAGGCTATCAGAGATGCCGAGAAAGCA GGGCGACTACCGCACCCGGATATGGAAATTCGAGGACGGGTTGAGCAACGTGTTGGTTATACAATTGAACAAATTAATCATATGCG TGATGTGTTTGGTACGCGATTGCGACGTGCTGAAGACGTATTTCCACCGGTGATCGGGGTTGCTGCCCATAAAGGTGGCGTTTACA AAACCTCAGTTTCTGTTCATCTTGCTCAGGATCTGGCTCTGAAGGGGCTACGTGTTTTGCTCGTGGAAGGTAACGACCCCCAGGGA ACAGCCTCAATGTATCACGGATGGGTACCAGATCTTCATATTCATGCAGAAGACACTCTCCTGCCTTTCTATCTTGGGGAAAAGG ACGATGTCACTTATGCAATAAAGCCCACTTGCTGGCCGGGGCTTGACATTATTCCTTCCTGTCTGGCTCTGCACCGTATTGAAACT GAGTTAATGGGCAAATTTGATGAAGGTAAACTGCCCACCGATCCACACCTGATGCTCCGACTGGCCATTGAAACTGTTGCTCATG ACTATGATGTCATAGTTATTGACAGCGCGCCTAACCTGGGTATCGGCACGATTAATGTCGTATGTGCTGCTGATGTGCTGATTGT TCCCACGCCTGCTGAGTTGTTTGACTACACCTCCGCACTGCAGTTTTTCGATATGCTTCGTGATCTGCTCAAGAACGTTGATCTTA AAGGGTTCGAGCCTGATGTACGTATTTTGCTTACCAAATACAGCAATAGCAATGGCTCTCAGTCCCCGTGGATGGAGGAGCAAAT TCGGGATGCCTGGGGAAGCATGGTTCTAAAAAATGTTGTACGTGAAACGGATGAAGTTGGTAAAGGTCAGATCCGGATGAGAACTG TTTTTGAACAGGCCATTGATCAACGCTCTTCAACTGGTGCCTGGAGAAATGCTCTTTCTATTTGGGAACCTGTCTGCAATGAAATT TTCGATCGTCTGATTAAACCACGCTGGGAGATTAGATAATGAAGCGTGCGCCTGTTATTCCAAAACATACGCTCAATACTCAACCG GTTGAAGATACTTCGTTATCGACACCAGCTGCCCCGATGGTGGATTCGTTAATTGCGCGCGTAGGAGTAATGGCTCGCGGTAATGC CATTACTTTGCCTGTATGTGGTCGGGATGTGAAGTTTACTCTTGAAGTGCTCCGGGGTGATAGTGTTGAGAAGACCTCTCGGGTAT GGTCAGGTAATGAACGTGACCAGGAGCTGCTTACTGAGGACGCACTGGATGATCTCATCCCTTCTTTTCTACTGACTGGTCAACAG ACACCGGCGTTCGGTCGAAGAGTATCTGGTGTCATAGAAATTGCCGATGGGAGTCGCCGTCGTAAAGCTGCTGCACTTACCGAAAG TGATTATCGTGTTCTGGTTGGCGAGCTGGATGATGAGCAGATGGCTGCATTATCCAGATTGGGTAACGATTATCGCCCAACAAGTG	Slattery et al., 2018 [61]

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
p*Pt*GE30-part 2	CTTATGAACGTGGTCAGCGTTATGCAAGCCGATTGCAGAATGAATTTGCTGGAAATATTTCTGCGCTGGCTGATGCGGAAAATATT TCACGTAAGATTATTACCCGCTGTATCAACACCGCCAAATTGCCTAAATCAGTTGTTGCTCTTTTTTCTCACCCCGGTGAACTATC TGCCCGGTCAGGTGATGCACTTCAAAAAGCCTTTACAGATAAAGAGGAATTACTTAAGCAGCAGGCATCTAACCTTCATGAGCAGA AAAAAGCTGGGGTGATATTTGAAGCTGAAGAAGTTATCACTCTTTTAACTTCTGTGCTTAAAACGTCATCTGCATCAAGAACTAGT TTAAGCTCACGACATCAGTTTGCTCCTGGAGCGACAGTATTGTATAAGGGCGATAAAATGGTGCTTAACCTGGACAGGTCTCGTGTT CCAACTGAGTGTATAGAGAAAATTGAGGCCATTCTTAAGGAACTTGAAAAGCCAGCACCCTGATGCGACCTCGTTTTAGTCTACGT TTATCTGTCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTGGCCTGAATATTCTCTCTGGGCCCACTGTTCCACTTG TATCGTCGGTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCG TATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATAATCAGACTGGGACCACGGTCCCACTCG TATCGTCGGTCTGATTATTAGTCTGGGACCATGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCG TATCGTCGGTCTGATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCG TATCGTCGGTCTGATTATTAGTCTGGGACCACGATCCCACTCGTGTTGTCGGTCTGATTATCGGTCTGGGACCACGGTCCCACTTG TATTGTCGATCAGACTATCAGCGTGAGACTACGATTCCATCAATGCCTGTCAAGGGCAAGTATTGACATGTCGTCGTAACCTGTAG AACGGAGTAACCTCGGTGTGCGGTTGTATGCCTGCTGTGGATTGCTGCTGTGTCCTGCTTATCCACAACATTTTGCGCACGGTTAT GTGGACAAAATACCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATCCGATGCAAGTGTGTCGCTGTCGACGAGCTCG CGAGCTCGGACATGAGGTTGCCCCGTATTCAGTGTCGCTGATTTGTATTGTCTGAAGTTGTTTTTACGTTAAGTTGATGCAGATCA ATTAATACGATACCTGCGTCATAATTGATTATTTGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGATAATCATT ATCACTTTACGGGTCCTTTCCGGTGATCCGACAGGTTACGGGGCGGCGACCTCGCGGGTTTTCGCTATTTATGAAAATTTTCCGGTT TAAGGCGTTTCCGTTCTTCTTCGTCATAACTTAATGTTTTTATTTAAAATACCCTCTGAAAAGAAAGGAAACGACAGGTGCTGAAA GCGAGCTTTTTGGCCTCTGTCGTTTCCTTTCTCTGTTTTTGTCCGTGGAATGAACAATGGAAGTCCGAGCTCATCGCTAATAACTT CGTATAGCATACATTATACGAAGTTATATTCGATGCGGCCGCAAGGGGTTCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCT TAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCACACAGATGCGTAAGGAGAAAATACC GCATCAGGCGCCATTCGCCATTCAGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAA AGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGT AATACGACTCACTATAGGGCGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGAGTATTCT ATAGTCTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCT TTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGAACCCCTTGCGGCCGCCCGGGCCGTCGACCAATT CTCATGTTTGACAGCTTATCATCGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCAACCAGGCGTTTAAGGG CACCAATAACTGCCTTAAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCCGACATGG AAGCCATCACAAACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGCCTTGCGTATAATATTTGCCCATGGTGAA AACGGGGGCGAAGAAGTTGTCCATATTGGCCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCTGAGACGAAAAACA TATTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAATATATGTGTAGAAACTGCCGGA AATCGTCGTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAACAAGGGTGAACACTATCCCAT ATCACCAGCTCACCGTCTTTCATTGCCATACGAAATTCCGGATGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATA AAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACT GACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATTTTAGC TTCCTTAGCTCCTGAAAATCTCGATAACTCAAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTT ACGTGCCGATCAACGTCTCATTTTCGCCAAAAGTTGGCCCAGGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTATTCTGC	Slattery et al., 2018 [61]

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
p*Pt*GE30-part 2	GAAGTGATCTTCCGTCACAGGTATTTATTCGCGATAAGCTCATGGAGCGGCGTAACCGTCGCACAGGAAGGACAGAGAAAGCGCGGA TCTGGGAAGTGACGGACAGAACGGTCAGGACCTGGATTGGGGAGGCGGTTGCCGCCGCTGCTGCTGACGGTGTGACGTTCTCTGTT CCGGTCACACCACATACGTTCCGCCATTCCTATGCGATGCACATGCTGTATGCCGGTATACCGCTGAAAGTTCTGCAAAGCCTGAT GGGACATAAGTCCATCAGTTCAACGGAAGTCTACACGAAGGTTTTTGCGCTGGATGTGGCTGCCCGGCACCGGGTGCAGTTTGCGA TGCCGGAGTCTGATGCGGTTGCGATGCTGAAACAATTATCCTGAGAATAAATGCCTTGGCCTTTATATGGAAATGTGGAACTGAGTG GATATGCTGTTTTTGTCTGTTAAACAGAGAAGCTGGCTGTTATCCACTGAGAAGCGAACGAAACAGTCGGGAAAATCTCCCATTAT CGTAGAGATCCGCATTATTAATCTCAGGAGCCTGTGTAGCGTTTATAGGAAGTAGTGTTCTGTCATGATGCCTGCAAGCGGTAACGA AAACGATTTGAATATGCCTTCAGGAACAATAGAAATCTTCGTGCGGTGTTACGTTGAAGTGGAGCGGATTATGTCAGCAATGGACA GAACAACCTAATGAACACAGAACCATGATGTGGTCTGTCCTTTTACAGCCAGTAGTGCTCGCCGCAGTCGAGCGACAGGGCGAAGCCC	Slattery et al., 2018 [61]
ShBle cassette	TCGAGCTGGTTGCCCTCGCCGCTGGGCTGGCGGCCGTCTATGGCCCTGCAAACGCGCCAGAAACGCCGTCGAAGCCGTGTGCGAGA CACCGCGGCCGGCCGCCGGCGTTGTGGATACCTCGCGGAAAACTTGGCCCTCACTGACAGATGAGGGGCGGACGTTGACACTTGAG GGGCCGACTCACCCGGCGCGGCGTTGACAGATGAGGGGCAGGCTCGATTTCGGCCGGCGACGTGGAGCTGGCCAGCCTCGCAAATC GGCGAAAACGCCTGATTTTACGCGAGTTTCCCACAGATGATGTGGACAAGCCTGGGGATAAGTGCCCTGCGGTATTGACACTTGAG GGGCGCGACTACTGACAGATGAGGGGCGCGATCCTTGACACTTGAGGGGCAGAGTGCTGACAGATGAGGGGCGCACCTATTGACAT TTGAGGGGCTGTCCACAGGCAGAAAATCCAGCATTTGCAAGGGTTTCCGCCCGTTTTTCGGCCACCGCTAACCTGTCTTTTAACCT GCTTTTAAACCAATATTTATAAACCTTGTTTTTAACCAGGGCTGCGCCCTGTGCGCGTGACCGCGCACGCCGAAGGGGGGTGCCCC CCCTTCTCGAACCCTCCCGGTCGAGTGAGCGAGGAAGCACCAGGGAACAGCACTTATATATTCTGCTTACACACGATGCCTGAAAA AACTTCCCTTGGGGTTATCCACTTATCCACGGGGATATTTTTATAATTATTTTTTTTATAGTTTTTAGATCTTCTTTTTTAGAGCG CCTTGTAGGCCTTTATCCATGCTGGTTCTAGAGAAGGTGTTGTGACAAATTGCCCTTTCAGTGTGACAAATCACCCTCAAATGACA GTCCTGTCTGTGACAAATTGCCCTTAACCCTGTGACAAATTGCCCTCAGAAGAAGCTGTTTTTTCACAAAGTTATCCCTGCTTATT GACTCTTTTTTATTTAGTGTGACAATCTAAAAACTTGTCACACTTCACATGGATCTGTCATGGCGGAAACAGCGGTTATCAATCAC AAGAAACGTAAAAATAGCCCGCGAATCGTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGATCAAAAACGTAT GCTGTATCTGTTCGTTGACCAGATCAGAAAATCTGATGGCACCCTACAGGAACATGACGGTATCTGCGAGATCCATGTTGCTAAAT ATGCTGAAATATTCGGATTGACCTCTGCGGAAGCCAGTAAGGATATACGGCAGGCATTGAAGAGTTTCGCGGGGAAGGAAGTGGTT TTTTATCGCCCTGAAGAGGATGCCGGCGATGAAAAAGGCTATGAATCTTTTCCTTGGTTTATCAAACGTGCGCACAGTCCATCCAG AGGGCTTTACAGTGTACATATCAACCCATATCTCATTCCCTTCTTTATCGGGTTACAGAACCGGTTTACGCAGTTTCGGCTTAGTG AAACAAAAGAAATCACCAATCCGTATGCCATGCGTTTATACGAATCCCTGTGTCAGTATCGTAAGCCGGATGGCTCAGGCATCGTC TCTCTGAAAATCGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGCCTGACTTCCGCCGCCGCTTCCTGCA GGTCTGTGTTAATGAGATCAACAGCAGAACTCCAATGCGCCTCTCATACATTGAGAAAAAGAAAGGCCGCCAGACGACTCATATCG TATTTTCCTTCCGCGATATCACTTCCATGACGACAGGATAGTCTGAGGGTTATCTGTCACAGATTTGAGGGTGGTTCGTCACATTT GTTCTGACCTACTGAGGGTAATTTGTCACAGTTTTGCTGTTTCCTTCAGCCTGCATGGATTTTCTCATACTTTTTGAACTGTAATT TTTAAGGAAGCCAAATTTGAGGGCAGTTTGTCACAGTTGATTTCCTTCTCTTTCCCTTCGTCATGTGACCTGATATCGGGGGTTAG TTCGTCATCATTGATGAGGGTTGATTATCACAGTTTATTACTCTGAATTGGCTATCCGCGTGTGTACCTCTACCTGGAGTTTTTCC CACGGTGGATATTTCTTCTTGCGCTGAGCGTAAGAGCTATCTGACAGAACAGTTCTTCTTTGCTTCCTCGCCAGTTCGCTCGCTAT GCTCGGTTACACGGCTGCGGCGAGCATCACGTGCTATAAAAATAATTATAATTTAAATTTTTTAATATAAATATATAAATTAAAAA TAGAAAGTAAAAAAAGAAATTAAAGAAAAAATAGTTTTTGTTTTCCGAAGATGTAAAAGACTCTAGGGGGATCGCCAACAAATACT ACCTTTTACCTTGCTCTTCCTGCTCTCAGGTATTAATGCCGAATTGTTTCATCTTGTCTGTGTAGAAGACCACACACGAAAATCCT GTGATTTTACATTTTACTTATCGTTAATCGAATGTATATCTATTTAATCTGCTTTTCTTGTCTAATAAATATATATGTAAAGTACGC TTTTTGTTGAAATTTTTTAAACCTTTGTTTATTTTTTTTTCTTCATTCCGTAACTCTTCTACCTTCTTTATTTACTTTCTAAAATC	Karas et al., 2015 [38]

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
ShBle cassette	CAAATACAAAACATAAAAATAAATAAACACAGAGTAAATTCCCAAATTATTCCATCATTAAAAGATACGAGGCGCGTGTAAGTTAC AGGCAAGCGATCCTAGTACACTCTATATTTTTTTATGCCTCGGTAATGATTTTCATTTTTTTTTTCCACCTAGCGGATGACTCTTT TTTTTTCTTAGCGATTGGCATTATCACATAATGAATTATACATTATATAAAGTAATGTGATTTCTTCGAAGAATATACTAAAAAAT GAGCAGGCAAGATAAACGAAGGCAAAGATGACAGAGCAGAAAGCCCTAGTAAAGCGTATTACAAATGAAACCAAGATTCAGATTGC GATCTCTTTAAAGGGTGGTCCCCTAGCGATAGAGCACTCGATCTTCCCAGAAAAAGAGGCAGAAGCAGTAGCAGAACAGGCCACAC AATCGCAAGTGATTAACGTCCACACAGGTATAGGGTTTCTGGACCATATGATACATGCTCTGGCCAAGCATTCCGGCTGGTCGCTA ATCGTTGAGTGCATTGGTGACTTACACATAGACGACCATCACACCACTGAAGACTGCGGGATTGCTCTCGGTCAAGCTTTTAAAGA GGCCCTAGGGGCCGTGCGTGGAGTAAAAAGGTTTGGATCAGGATTTGCGCCTTTGGATGAGGCACTTTCCAGAGCGGTGGTAGATC TTTCGAACAGGCCGTACGCAGTTGTCGAACTTGGTTTGCAAAGGGAGAAAGTAGGAGATCTCTCTTGCGAGATGATCCCGCATTTT CTTGAAAGCTTTGCAGAGGCTAGCAGAATTACCCTCCACGTTGATTGTCTGCGAGGCAAGAATGATCATCACCGTAGTGAGAGTGC GTTCAAGGCTCTTGCGGTTGCCATAAGAGAAGCCACCTCGCCCAATGGTACCAACGATGTTCCCTCCACCAAAGGTGTTCTTATGT AGTTTTACACAGGAGTCTGGACTTGACGCTAGTGATAATAAGTGACTGAGGTATGTGCTCTTCTTATCTCCTTTTGTAGTGTTGCT CTTATTTTAAACAACTTTGCGGTTTTTTGATGACTTTGCGATTTTGTTGTTGCTTTGCAGTAAATTGCAAGATTTAATAAAAAAAC GCAAAGCAATGATTAAAGGATGTTCAGAATGAAACTCATGGAAACACTTAACCAGTGCATAAACGCTGGTCATGAAATGACGAAGG CTATCGCCATTGCACAGTTTAATGATGACAGCCCGGAAGCGAGGAAAATAACCCGGCGCTGGAGAATAGGTGAAGCAGCGGATTTA GTTGGGGTTTCTTCTCAGGCTATCAGAGATGCCGAGAAAGCAGGGCGACTACCGCACCCGGATATGGAAATTCGAGGACGGGTTGA GCAACGTGTTGGTTATACAATTGAACAAATTAATCATATGCGTGATGTGTTTGGTACGCGATTGCGACGTGCTGAAGACGTATTTC CACCGGTGATCGGGGTTGCTGCCCATAAAGGTGGCGTTTACAAAACCTCAGTTTCTGTTCATCTTGCTCAGGATCTGGCTCTGAAG GGGCTACGTGTTTTGCTCGTGGAAGGTAACGACCCCCAGGGAACAGCCTCAATGTATCACGGATGGGTACCAGATCTTCATATTCAT GCAGAAGACACTCTCCTGCCTTTCTATCTTGGGGAAAAGGACGATGTCACTTATGCAATAAAGCCCACTTGCTGGCCGGGGCTTGAC ATTATTCCTTCCTGTCTGGCTCTGCACCGTATTGAAACTGAGTTAATGGGCAAATTTGATGAAGGTAAACTGCCCACCGATCCACA CCTGATGCTCCGACTGGCCATTGAAACTGTTGCTCATGACTATGATGTCATAGTTATTGACAGCGCGCCTAACCTGGGTATCGGCAC GATTAATGTCGTATGTGCTGCTGATGTGCTGATTGTTCCCACGCCTGCTGAGTTGTTTGACTACACCTCCGCACTGCAGTTTTTCG ATATGCTTCGTGATCTGCTCAAGAACGTTGATCTTAAAGGGTTCGAGCCTGATGTACGTATTTTGCTTACCAAATACAGCAATAGC AATGGCTCTCAGTCCCCGTGGATGGAGGAGCAAATTCGGGATGCCTGGGGAAGCATGGTTCTAAAAAATGTTGTACGTGAAACGGA TGAAGTTGGTAAAGGTCAGATCCGGATGAGAACTGTTTTTGAACAGGCCATTGATCAACGCTCTTCAACTGGTGCCTGGAGAAATGC TCTTTCTATTTGGGAACCTGTCTGCAATGAAATTTTCGATCGTCTGATTAAACCACGCTGGGAGATTAGATAATGAAGCGTGCGCCT GTTATTCCAAAACATACGCTCAATACTCAACCGGTTGAAGATACTTCGTTATCGACACCAGCTGCCCCGATGGTGGATTCGTTAATT GCGCGCGTAGGAGTAATGGCTCGCGGTAATGCCATTACTTTGCCTGTATGTGGTCGGGATGTGAAGTTTACTCTTGAAGTGCTCCGG GGTGATAGTGTTGAGAAGACCTCTCGGGTATGGTCAGGTAATGAACGTGACCAGGAGCTGCTTACTGAGGACGCACTGGATGATCTC ATCCCTTCTTTTCTACTGACTGGTCAACAGACACCGGCGTTCGGTCGAAGAGTATCTGGTGTCATAGAAATTGCCGATGGGAGTCGC CGTCGTAAAGCTGCTGCACTTACCGAAAGTGATTATCGTGTTCTGGTTGGCGAGCTGGATGATGAGCAGATGGCTGCATTATCCAGA TTGGGTAACGATTATCGCCCAACAAGTGCTTATGAACGTGGTCAGCGTTATGCAAGCCGATTGCAGAATGAATTTGCTGGAAATATT TCTGCGCTGGCTGATGCGGAAAATATTTCACGTAAGATTATTACCCGCTGTATCAACACCGCCAAATTGCCTAAATCAGTTGTTGCTC TTTTTTCTCACCCCGGTGAACTATCTGCCCGGTCAGGTGATGCACTTCAAAAAGCCTTTACAGATAAAGAGGAATTACTTAAGCAGCA GGCATCTAACCTTCATGAGCAGAAAAAAGCTGGGGTGATATTTGAAGCTGAAGAAGTTATCACTCTTTTAACTTCTGTGCTTAAAACG TCATCTGCATCAAGAACTAGTTTAAGCTCACGACATCAGTTTGCTCCTGGAGCGACAGTATTGTATAAGGGCGATAAAATGGTGCTTA ACCTGGACAGGTCTCGTGTTCCAACTGAGTGTATAGAGAAAATTGAGGCCATTCTTAAGGAACTTGAAAAGCCAGCACCCTGATGCGAC CTCGTTTTAGTCTACGTTTATCTGTCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTGGCCTGAATATTCTCTCTGGGC CCACTGTTCCACTTGTATCGTCGGTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGAC	Karas et al., 2015 [38]

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
ShBle cassette	CACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATAATCAGACTGGGACC ACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCATGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCA CGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCAC GGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGATCCCACTCGTGTTGTCGGTCTGATTATCGGTCTGGGACCACG GTCCCACTTGTATTGTCGATCAGACTATCAGCGTGAGACTACGATTCCATCAATGCCTGTCAAGGGCAAGTATTGACATGTCGTCGT AACCTGTAGAACGGAGTAACCTCGGTGTGCGGTTGTATGCCTGCTGTGGATTGCTGCTGTGTCCTGCTTATCCACAACATTTTGCGC ACGGTTATGTGGACAAAATACCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATCCGATGCAAGTGTGTCGCTGTCGAC GAGCTCGCGAGCTCGGACATGAGGTTGCCCCGTATTCAGTGTCGCTGATTTGTATTGTCTGAAGTTGTTTTTACGTTAAGTTGATGC AGATCAATTAATACGATACCTGCGTCATAATTGATTATTTGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGATAA TCATTATCACTTTACGGGTCCTTTCCGGTGATCCGACAGGTTACGGGGCGGCGACCTCGCGGGTTTTCGCTATTTATGAAAATTTTCC GGTTTAAGGCGTTTCCGTTCTTCTTCGTCATAACTTAATGTTTTTATTTAAAATACCCTCTGAAAAGAAAGGAAACGACAGGTGCTGA AAGCGAGCTTTTTGGCCTCTGTCGTTTCCTTTCTCTGTTTTTGTCCGTGGAATGAACAATGGAAGTCCGAGCTCATCGCTAATAAC TTCGTATAGCATACATTATACGAAGTTATATTCGATGCGGCCGCAAGGGGTTCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGC TTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCACACAGATGCGTAAGGAGAAAATAC CGCATCAGGCGCCATTCGCCATTCAGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCG AAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATT GTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGAGTATT CTATAGTCTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAA CATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCG CTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGAACCCCTTGCGGCCGCCCGGGCCGTCGACCAA TTCTCATGTTTGACAGCTTATCATCGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCAACCAGGCGTTTAAG GGCACCAATAACTGCCTTAAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCCGACA TGGAAGCCATCACAAACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGCCTTGCGTATAATATTTGCCCATGGT GAAAACGGGGGCGAAGAAGTTGTCCATATTGGCCACGTTTAAATCAAAACTGGTGAAACTCACCCAGGGATTGGCTGAGACGAAAA ACATATTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAATATATGTGTAGAAACTGCC GGAAATCGTCGTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAACAAGGGTGAACACTATCCC ATATCACCAGCTCACCGTCTTTCATTGCCATACGAAATTCCGGATGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGA TAAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACT GACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATTTTAGCT TCCTTAGCTCCTGAAAATCTCGATAACTCAAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACG TGCCGATCAACGTCTCATTTTCGCCAAAAGTTGGCCCAGGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTATTCTGCGAAG TGATCTTCCGTCACAGGTATTTATTCGCGATAAGCTCATGGAGCGGCGTAACCGTCGCACAGGAAGGACAGAGAAAGCGCGGATCTG GGAAGTGACGGACAGAACGGTCAGGACCTGGATTGGGGAGGCGGTTGCCGCCGCTGCTGCTGACGGTGTGACGTTCTCTGTTCCGGT CACACCACATACGTTCCGCCATTCCTATGCGATGCACATGCTGTATGCCGGTATACCGCTGAAAGTTCTGCAAAGCCTGATGGGACA TAAGTCCATCAGTTCAACGGAAGTCTACACGAAGGTTTTTGCGCTGGATGTGGCTGCCCGGCACCGGGTGCAGTTTGCGATGCCGGA GTCTGATGCGGTTGCGATGCTGAAACAATTATCCTGAGAATAAATGCCTTGGCCTTTATATGGAAATGTGGAACTGAGTGGATATGC TGTTTTTGTCTGTTAAACAGAGAAGCTGGCTGTTATCCACTGAGAAGCGAACGAAACAGTCGGGAAAATCTCCCATTATCGTAGAGA TCCGCATTATTAATCTCAGGAGCCTGTGTAGCGTTTATAGGAAGTAGTGTTCTGTCATGATGCCTGCAAGCGGTAACGAAAACGATT TGAATATGCCTTCAGGAACAATAGAAATCTTCGTGCGGTGTTACGTTGAAGTGGAGCGGATTATGTCAGCAATGGACAGAACAACCT AATGAACACAGAACCATGATGTGGTCTGTCCTTTTACAGCCAGTAGTGCTCGCCGCAGTCGAGCGACAGGGCGAAGCCC	Karas et al., 2015 [38]

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
40SRPS8 promoter	CCCTGCGATAGACCTTTTCCAAACTCACGCAGTCCAAGAAAACAAAGGGGTGAGAAGTATACGCACCTTTCGGTTTCGGCATAATTCT TAAACTCTTGTGGTCACTTTCTTGTGAAGAAGCTAGGGGCACTCGTTTTCCCTCAGAGCCTGCAAACACAAAATTCCTGCAGTCAAT TGTCCCAACACTCGGCAAACCGTATGCGCAAGCAACGATGCGCAGAAGGCCGTGGATGGATGGCGACTCGCGATATGGCTTCTTGGG TCGCCAGTGTGGTACGTCCGGCGTATGTCAATACGCGAATTCGGACGACTGGCATCTCTAGGAGGAGGATTCCTTCTTTTATGACAT GTTTATTTTATATACATTGATGCTTTCCGACAGTCGGAAGTAATAAATGAATTTATTTCAAGACTACCTATACTCCTTTGACTTGTT CGACTAATCTTACCGCTTACTAAAATCTCGAAATCACGCTTGACCTCTCGCACGCAAATTTTTGCTGCTGGACGCTACGCACTCGGC CCAATTCTTCTCGGTCCTCGTCGTCGCAATTGTCGTTGCGTTGATCTTGCACCGAAGGAATCAGAGAATAGAATACC	Slattery et al., 2018 [61]
FcpA terminator	CCGCAACAACTACCTCGACTTTGGCTGGGACACTTTCAGTGAGGACAAGAAGCTTCAGAAGCGTGCTATCGAACTCAACCAGGGACGTG CGGCACAAATGGGCATCCTTGCTCTCATGGTGCACGAACAGTTGGGAGTCTCTATCCTTCCTTAAAAATTTAATTTTCATTAGTTGCAGTCA- CTCCGCTTTGGTTT	Niu et al., 2013 [63]
OA1	ATGAACCACCTCCGTGCCGAAGGACCCGCCTCCGTCCTCGCCATTGGAACCGCCAACCCCGAAAACATTCTCCTCCAGGACGAATTCCC GGACTACTACTTCCGCGTCACCAAGTCGGAACACATGACCCAGCTCAAGGAAAAGTTCCGTAAGATCTGTGACAAGTCGATGATTCG TAAGCGCAACTGCTTCCTCAACGAAGAACACTTGAAGCAGAACCCCCGTCTCGTCGAACACGAAATGCAGACCCTCGACGCCCGTCA GGACATGCTCGTCGTCGAAGTCCCCAAGCTCGGAAAGGATGCCTGCGCCAAGGCCATCAAGGAATGGGGTCAGCCCAAGTCCAAGAT CACCCACCTCATTTTCACCTCCGCCTCCACCACCGACATGCCCGGAGCTGACTACCACTGCGCCAAGTTGCTCGGCCTCTCCCCCTC CGTCAAGCGCGTCATGATGTACCAGCTTGGCTGCTACGGCGGAGGTACCGTCTTGCGTATTGCCAAGGACATCGCCGAAAACAACAA GGGAGCCCGTGTCTTGGCCGTCTGCTGTGACATTATGGCCTGCCTCTTCCGTGGCCCCTCCGAATCCGACCTCGAACTCTTGGTCGGT CAGGCCATCTTTGGCGATGGAGCCGCCGCCGTCATTGTCGGTGCCGAACCCGACGAATCCGTCGGAGAACGTCCCATTTTTGAACTT GTCTCCACCGGACAGACCATTCTCCCCAACTCCGAAGGAACCATTGGCGGCCACATTCGTGAAGCCGGACTCATCTTTGACCTCCAC AAGGACGTCCCGATGCTCATCTCCAACAACATCGAAAAGTGCCTCATTGAAGCCTTCACTCCGATCGGAATTTCCGACTGGAACTCC ATCTTCTGGATTACCCACCCCGGCGGAAAGGCCATTTTGGACAAGGTCGAAGAAAAGCTCCACCTCAAGTCCGATAAGTTTGTCGAC TCCCGTCACGTCCTCTCCGAACACGGTAACATGTCCTCCTCCACCGTCCTCTTTGTCATGGACGAACTCCGTAAGCGTTCCTTGGAA GAAGGCAAGTCGACCACCGGCGACGGCTTCGAATGGGGAGTCTTGTTTGGCTTCGGTCCCGGATTGACCGTCGAACGTGTCGTCGTC CGTTCCGTTCCCATCAAGTACGGATCCGGAGAAGGACGCGGTTCGTTGCTCACCTGCGGAGACGTCGAAGAAAACCCCGGACCCATGG CCGTCAAGCACCTCATTGTCTTGAAGTTCAAGGATGAAATCACCGAAGCTCAGAAGGAAGAATTCTTCAAGACCTACGTCAACCTCG TCAACATCATTCCCGCCATGAAGGACGTCTACTGGGGAAAGGACGTCACCCAGAAGAACAAGGAAGAAGGATACACCCACATTGTCG AAGTCACCTTTGAATCCGTCGAAACCATCCAGGACTACATCATCCACCCCGCCCACGTCGGATTCGGTGACGTCTACCGCTCCTTCT GGGAAAAGCTCTTGATTTTCGACTACACCCCCCGTAAATAA	This study
OA2	ATGAACCACCTCCGTGCCGAAGGTCCCGCCTCCGTCCTCGCCATTGGAACCGCCAACCCCGAAAACATTCTCCTCCAAGACGAATTCCC TGACTACTACTTCCGCGTGACTAAGAGTGAGCACATGACGCAATTGAAAGAAAAGTTCCGTAAGATCTGTGATAAGTCCATGATCCG TAAGCGTAACTGTTTCCTGAATGAGGAACACCTTAAGCAGAACCCCCGCTTGGTCGAACATGAAATGCAGACCTTGGATGCCCGCCA GGACATGCTCGTCGTTGAAGTCCCTAAGTTGGGTAAAGACGCTTGCGCCAAGGCCATTAAGGAATGGGGACAGCCCAAGAGTAAGAT CACCCATTTGATTTTCACATCCGCCTCTACGACCGATATGCCTGGAGCCGATTACCACTGCGCCAAGCTTCTGGGACTGTCCCCCTC CGTTAAGCGCGTCATGATGTACCAGCTTGGATGTTACGGAGGTGGCACCGTCCTCCGTATTGCCAAGGACATTGCCGAAAACAACAA GGGTGCCCGTGTCCTCGCCGTCTGCTGCGACATCATGGCCTGCCTCTTCCGCGGACCCTCCGAATCCGACTTGGAACTCCTCGTGGG ACAGGCCATTTTTGGTGACGGCGCCGCCGCCGTCATTGTCGGAGCTGAACCCGACGAATCCGTCGGAGAAAGGCCCATTTTTGAATT GGTCTCCACCGGACAGACCATTCTCCCGAACTCCGAAGGAACCATCGGCGGACACATTCGTGAAGCCGGTCTCATCTTTGACTTGCA CAAGGACGTTCCCATGCTCATTTCCAACAACATTGAAAAGTGCCTCATTGAAGCCTTCACCCCCATTGGAATCTCCGACTGGAATAG TATCTTTTGGATTACCCACCCCGGAGGAAAGGCCATCTTGGATAAGGTGGAAGAAAAGCTCCACCTCAAGTCCGACAAGTTCGTCGA	This study

Table A2. *Cont.*

Name	Nucleotide Sequence (Codon Optimized)	Description and/or References
OA2	CTCCCGTCACGTCTTGTCCGAACACGGAAACATGTCGTCCTCCACCGTCCTCTTTGTCATGGATGAATTGCGTAAGCGCTCGCTCGA AGAAGGAAAGTCCACCACCGGCGATGGTTTCGAATGGGGCGTTCTCTTCGGATTTGGACCCGGTCTCACCGTCGAACGTGTCGTCGTC CGTTCTGTCCCGATTAAGTACCACCATCACCACCACCACGGATCGGGCGAAGGACGTGGTTCCCTCCTTACGTGCGGAGACGTCGAA GAAAACCCCGGTCCCATGGCCGTCAAGCACTTGATCGTTCTCAAGTTCAAGGACGAAATCACCGAAGCCCAGAAGGAAGAATTCTTC AAGACCTACGTCAACCTCGTCAACATCATTCCCGCCATGAAGGATGTTTACTGGGGAAAGGACGTCACCCAGAAGAACAAGGAAGAA GGTTACACCCACATCGTCGAAGTCACCTTTGAATCCGTTGAAACCATCCAGGACTACATTATCCACCCGGCCCACGTCGGCTTCGGA GACGTCTACCGTTCCTTTTGGGAAAAGTTGTTGATTTTCGACTACACCCCCCGTAAGGAACAGAAGTTGATTTCCGAAGAAGACCTCTAA	This study
OA3	ATGCGTAAGGGTGAAGAACTCTTCACCGGAGTCGTCCCCATCTTGGTGGAATTGGACGGAGACGTCAACGGCCATAAGTTCTCGGTT TCCGGTGAAGGTGAAGGTGACGCCACTAACGGAAAGCTCACCCTCAAGTTCATTTGCACCACCGGCAAGCTCCCGGTCCCCTGGCCC ACCCTCGTCACCACCTTGACTTACGGTGTCCAGTGCTTCGCCCGTTACCCCGACCACATGAAGCAGCACGACTTCTTTAAGTCCGCC ATGCCCGAAGGATACGTCCAGGAACGCACCATTTCCTTCAAGGACGACGGAACCTACAAGACCCGTGCCGAAGTCAAGTTTGAAGGAG ATACCCTCGTCAACCGCATTGAACTCAAGGGCATTGATTTCAAAGAAGATGGCAACATTCTCGGACACAAGCTCGAATACAACTTCAA CTCCCACAACGTCTACATCACCGCCGATAAGCAAAAGAACGGTATTAAGGCCAACTTCAAGATTCGCCACAACGTCGAAGACGGTTC CGTCCAGTTGGCCGACCACTACCAGCAGAACACCCCGATTGGTGACGGTCCGGTCCTCCTCCCCGACAACCACTACCTCTCGTACCAG TCCGCCCTCTCCAAGGATCCGAACGAGAAGCGTGACCACATGGTCCTCTTGGAATTCGTCACCGCCGCCGGAATCACCCACGGCATGG ACGAACTCTACAAGCGTCCCGCCGCCAACGACGAGAACTACGCCGCCTCCGTCGGTATGAACCACTTGCGCGCCGAAGGACCGGCTTC CGTCTTGGCCATTGGTACCGCCAACCCGGAAAACATCCTCCTCCAGGACGAATTCCCCGACTACTACTTCCGTGTCACCAAGTCCGAA CACATGACCCAGTTGAAGGAAAAATTCCGTAAGATTTGTGATAAGTCCATGATTCGTAAGCGTAACTGCTTCTTGAACGAAGAACACC TCAAGCAAAACCCCCGTCTCGTCGAACACGAAATGCAAACCCTCGACGCCCGTCAGGACATGCTCGTTGTCGAAGTCCCGAAGCTTGG CAAGGACGCCTGCGCCAAGGCCATTAAGGAATGGGGCCAGCCCAAGAGCAAGATTACCCACTTGATTTTCACCTCCGCCTCCACCACC GATATGCCCGGCGCTGACTACCACTGCGCTAAGCTCCTCGGTCTCTCCCCCTCCGTCAAGCGTGTCATGATGTACCAGCTCGGATGCT ACGGAGGTGGAACCGTCCTCCGCATTGCCAAGGACATTGCTGAAAACAACAAGGGCGCCCGTGTCCTCGCCGTCTGCTGTGACATTAT GGCCTGCTTGTTTCGTGGTCCCTCCGAATCCGATTTGGAACTCCTCGTCGGCCAGGCCATTTTTGGTGACGGTGCCGCCGCCGTCATC GTCGGCGCCGAACCGGACGAATCCGTCGGTGAACGTCCGATTTTTGAACTCGTCTCCACCGGTCAAACCATTCTCCCCAACTCCGAAG GAACCATCGGAGGACACATCCGTGAAGCCGGACTCATTTTTGACCTCCACAAGGATGTCCCCATGCTCATCTCCAACAACATCGAAAA GTGCCTCATCGAAGCCTTCACCCCCATTGGCATCTCCGACTGGAACTCCATTTTCTGGATTACCCACCCCGGTGGAAAGGCCATTCTC GACAAGGTCGAAGAAAAGCTCCACTTGAAGTCCGATAAGTTCGTCGACTCCCGCCACGTCCTCTCCGAACACGGAAACATGTCCTCCT CCACCGTCCTCTTCGTCATGGACGAACTCCGTAAGCGTTCCCTCGAAGAAGGAAAGTCCACGACGGGTGACGGATTCGAGTGGGGAGT CCTCTTCGGCTTCGGTCCCGGACTCACCGTCGAACGTGTCGTCGTCCGTTCCGTGCCCATCAAGTACGGAGGAGGAGGATCCGGTGGT GGCGGATCCGGAGGTGGTGGATCGATGGCCGTCAAGCACCTCATTGTCTTGAAGTTTAAGGACGAAATCACCGAAGCCCAGAAGGAAG AATTCTTCAAGACTTACGTCAACTTGGTCAACATCATCCCCGCCATGAAGGACGTCTACTGGGGAAAGGACGTCACCCAAAAGAACAA GGAAGAAGGATACACCCACATTGTCGAAGTCACCTTTGAATCCGTCGAAACCATTCAGGACTACATTATTCACCCCGCCCACGTCGGTT TCGGTGACGTCTACCGTTCCTTCTGGGAAAAGCTCTTGATTTTTGACTACACCCCCCGTAAGGAACAGAAGTTGATTTCCGAAGAAGACTTGTAA	This study

Table A3. List of the primers used in this study for the *P. tricornutum* episome, with the design of and description of the obtained product.

Primer Name	Sequence (5′ to 3′)	Amplicon Size (bp)	Description of Product
bk1-FP	CAGGGTAATATAGATCTTCCGCTGCATAACCCTGCTTCGGGGTCATTATAGCGATTTTTT	6158	pPTGE30 backbone part 1
bk1-RP	TTTGCAAACCAAGTTCGACAACTGCGTACGGCCTGTTCGAAAGATCTACCACCGCTCTGG		
bk1-FP	GGGCCGTGCGTGGAGTAAAAAGGTTTGGATCAGGATTTGCGCCTTTGGATGAGGCACTTT	6318	pPTGE30 backbone part 2
bk2-RP	ATGGGCTTCGCCCTGTCGCTCGACTGCGGCGAGCACTACTGGCTGTAAAAGGACAGACCA		
ShBle-FP	TGGTCTGTCCTTTTACAGCCAGTAGTGCTCGCCGCAGTCGAGCGACAGGGCGAAGCCCATGAGCACAAGAGGTGACAAAA	1466	*P. t* selection marker region
ShBle-RP	ATACTTCTCACCCCTTTGTTTTCTTGGACTGCGTGAGTTTGGAAAAGGTCTATCGCAGGGATTGCAGCTTGTTGCAGAAG		
40SRPS8-FP	TTGCATGGTTAGACCTCCTTGACGACTGTGAGCCTACATCCTTCTGCAACAAGCTGCAATCCCTGCGATAGACCTTTTCC	700	40SRPS8 Promoter region
40SRPS8-RP	CGAGGACGGAGGCGGGTCCTTCGGCACGGAGGTGGTTCATGGTATTCTATTCTCTGATTC		
FcpA terminator-FP	GGAAAAGCTCTTGATTTTCGACTACACCCCCCGTAAATAACCGCAACAACTACCTCGACT	275	FcpA Terminator region
FcpA terminator-RP	CCGAAGCAGGGTTATGCAGCGGAAGATCTATATTACCCTGAAACCAAAGCGGAGTGACTG		
OA1-FP	GCGTTGATCTTGCACCGAAGGAATCAGAGAATAGAATACCATGAACCACCTCCGTGCCGA	1524	OA1 pathway genes
OA1-RP	ACTGAAAGTGTCCCAGCCAAAGTCGAGGTAGTTGTTGCGGTTATTTACGGGGGGTGTAGT		
OA2-FP	GCGTTGATCTTGCACCGAAGGAATCAGAGAATAGAATACCATGAACCACCTCCGTGCCGA	1572	OA2 pathway genes
OA2-RP	AGTCGAGGTAGTTGTTGCGGTTAGAGGTCTTCTTCGGAAATCAACTTCTGTTCCTTACGG		
OA3-FP	GCGTTGATCTTGCACCGAAGGAATCAGAGAATAGAATACCATGCGTAAGGGTGAAGAACT	2352	OA3 pathway genes
OA3-RP	AGTCGAGGTAGTTGTTGCGGTTACAAGTCTTCTTCGGAAATCAACTTCTGTTCCTTACGG		
Q-TKS1-Gn2-F	CGA CTG GAA CTC CAT CTT CTG	527	Pt OA1 colony PCR
Q-OAC1-Gn2-R	GGGTGTATCCTTCTTCCTTGTT		
Q-TKS3-Gn2-F	ATTGGAATCTCCGACTGGAATAG	594	Pt OA2 colony PCR
Q-OAC3-Gn2-R	GATGGTTTCAACGGATTCAAAGG		
Q-TKS7-Gn2-F	GTCAGGACATGCTCGTTGT	1211	Pt OA3 colony PCR
Q-OAC7-Gn2-R	TTCCCAGAAGGAACGGTAGA		

P. t: Phaeodactylum tricornutum; FcpA: fucoxanthin-chlorophyll binding protein.

Table A4. Olivetolic acid biosynthesis construct analysis. The plasmid constructs were rescued from positive *P. tricornutum*-transformed clones and amplified in *E. coli* before being extracted and fully sequenced at the Massachusetts General Hospital (MGH), Center for Computational and Integrative Biology (CCIB) DNA Core, using Illumina MiSeq technology. ON: original nucleotide; SN substitute nucleotide.

Sequence ID	Identity	Mismatch Position (ON-SN)	Effect on Protein	Number of Insertion
*Pt*OA1	>99%	9305 (T-C)	No	60 bp insertion (after terminator, no effect)
*Pt*OA2	>99%	8526 (C-T) 9304 (T-C)	No	0
*Pt*OA3	>99%	8192 (G-A) 9304 (T-C)	No	0

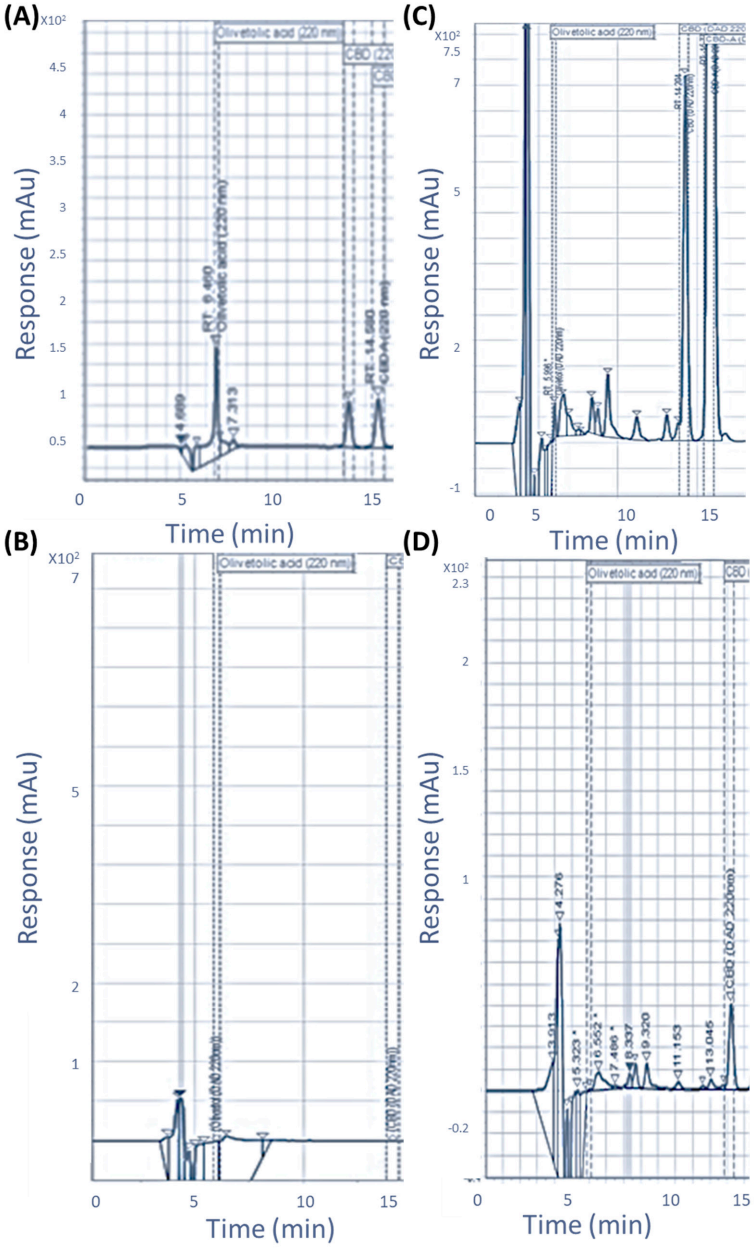
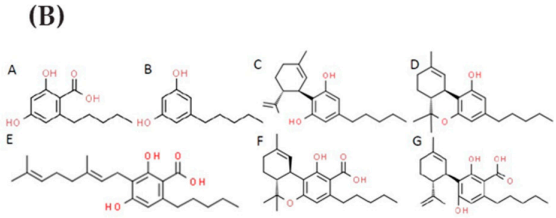
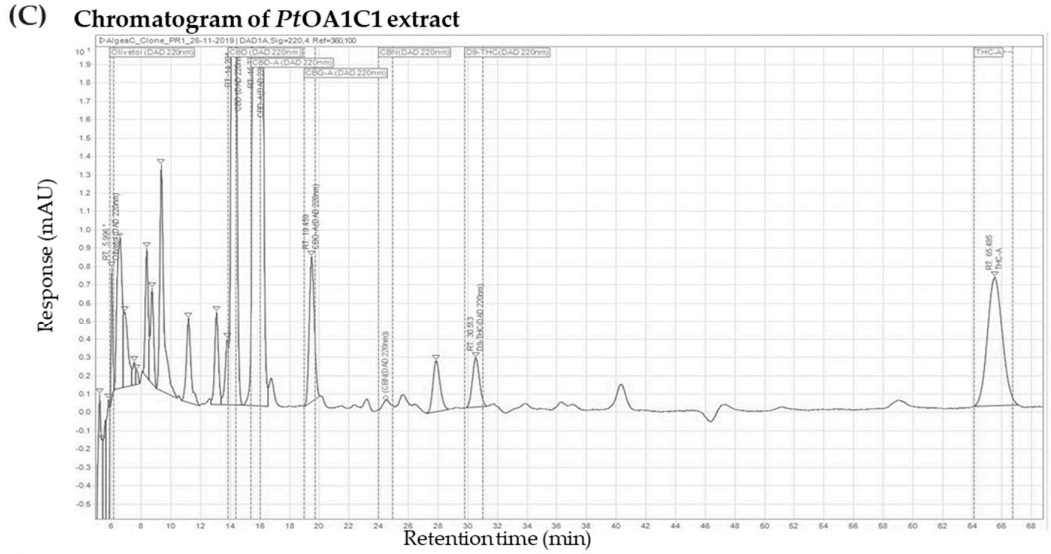


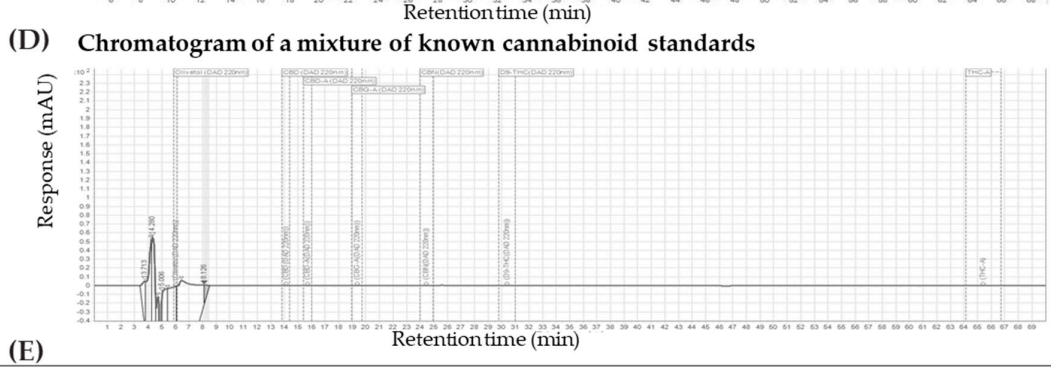
Figure A3. Olivetolic acid detection by HPLC-UV. Curves obtained by the injection of 10 µL of each sample. A concentration of 10 mg/kg of commercial standards (**A**); metabolite extraction from the *P. tricornutum* wild type (**B**); *Pt*OA1 (**C**) and *Pt*OA2 (**D**).

(A)

Compound	Molecular formula	Structure assignment	Parent m/z
Olivetolic acid	$C_{12}H_{16}O_4$	A	225.104
Olivetol	$C_{11}H_{16}O_2$	B	181.115
Cannabidiol	$C_{21}H_{30}O_2$	C	315.224
Δ-9-Tetrahydrocannabinol	$C_{21}H_{30}O_2$	D	315.224
Cannabigerolic acid	$C_{22}H_{32}O_4$	E	361.230
Δ-9-Tetrahydrocannabinolic acid	$C_{22}H_{30}O_4$	F	319.214
Cannabidiolic acid	$C_{22}H_{30}O_4$	G	319.214







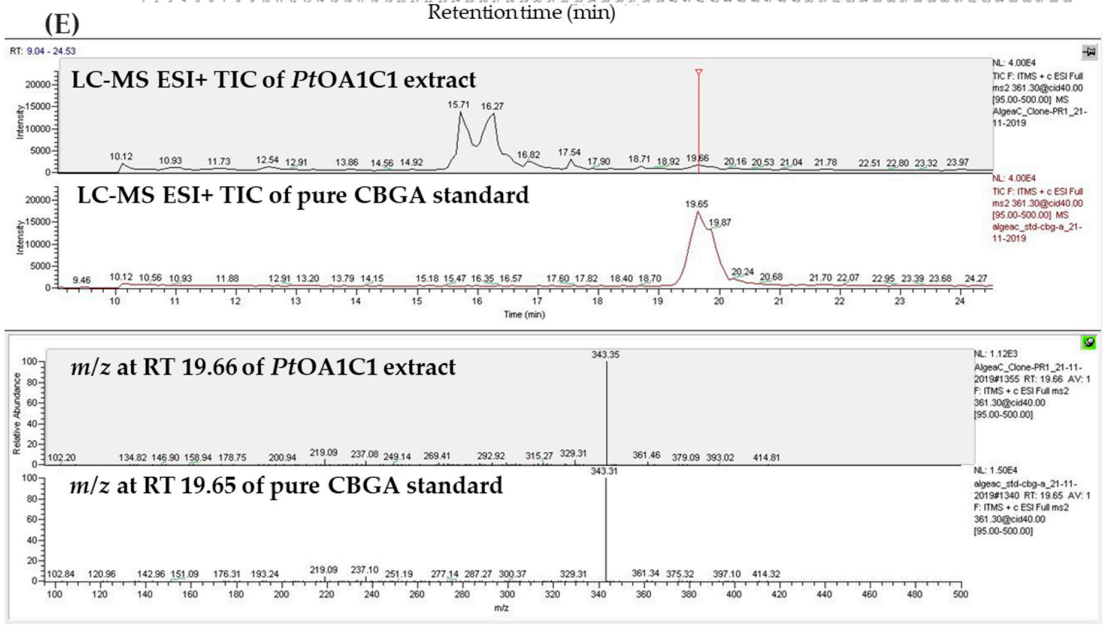


Figure A4. *Cont.*

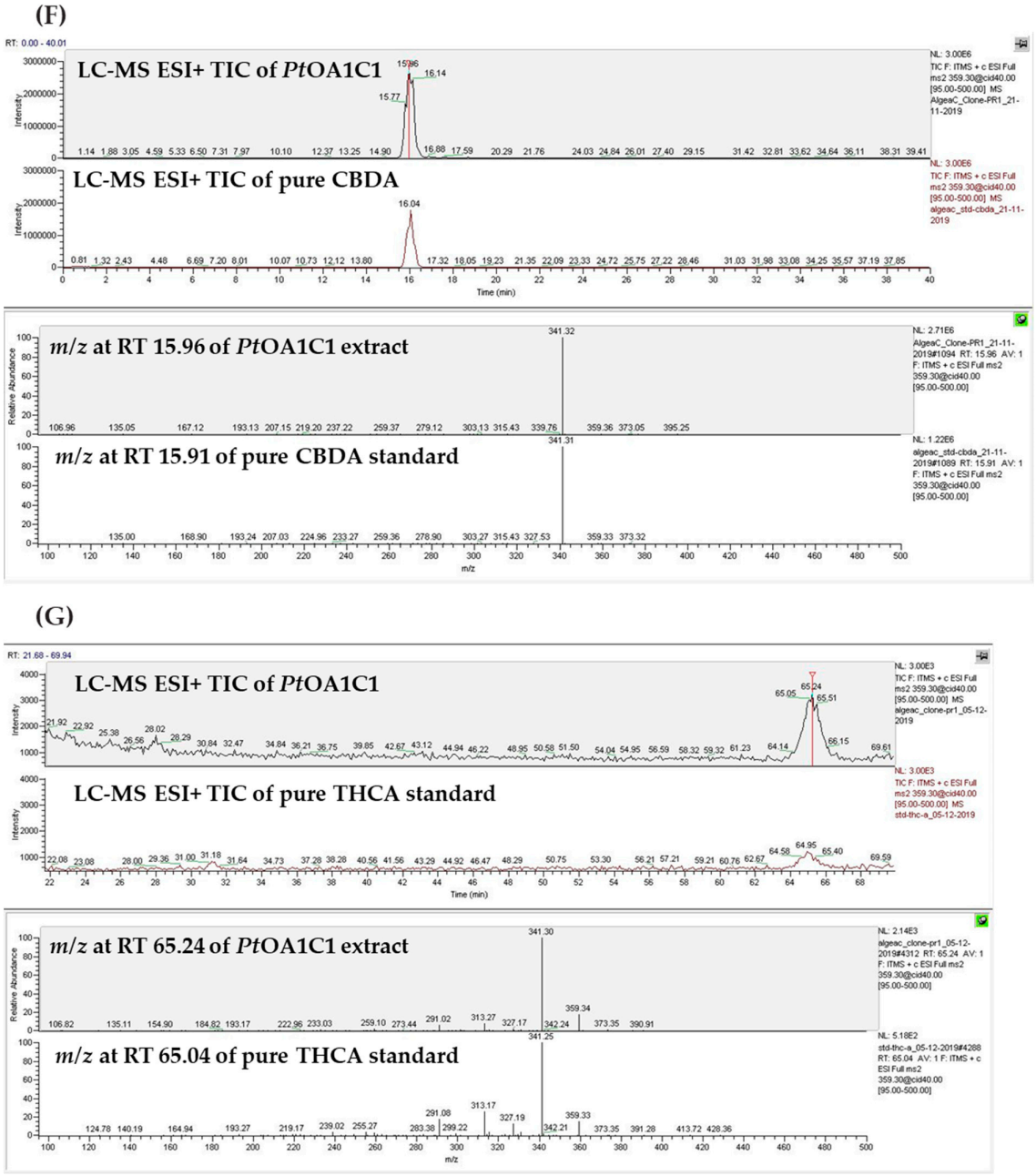


Figure A4. The cannabinoid-like metabolites detected via HPLC-UV and LC-MS analyses. (**A**) The list and (**B**) chemical structure of the targeted metabolites. A representative LC chromatogram of (**C**) 10 µL of metabolite extraction from *P. tricornutum Pt*OA1C1 and of (**D**) 10 mg/kg of a mixture of known cannabinoid commercial standards, both showing peaks corresponding to the retention times (RT) of cannabinoid standards. LC-MS in ESI+ analyses. Total ion chromatograph (TIC) of selected peaks from the corresponding cannabinoids, such as *P. tricornutum Pt*OA1C1 (**E**) peak at RT 19.66 min, with an m/z similar to the CBGA standard (RT 19.65 min). (**F**) The peak at RT 15.96 min, with an m/z similar to the CBDA standard (RT 15.91 min). (**G**) The peak at RT 19.66 min, with an m/z similar to the THCA standard (RT 65.04 min).

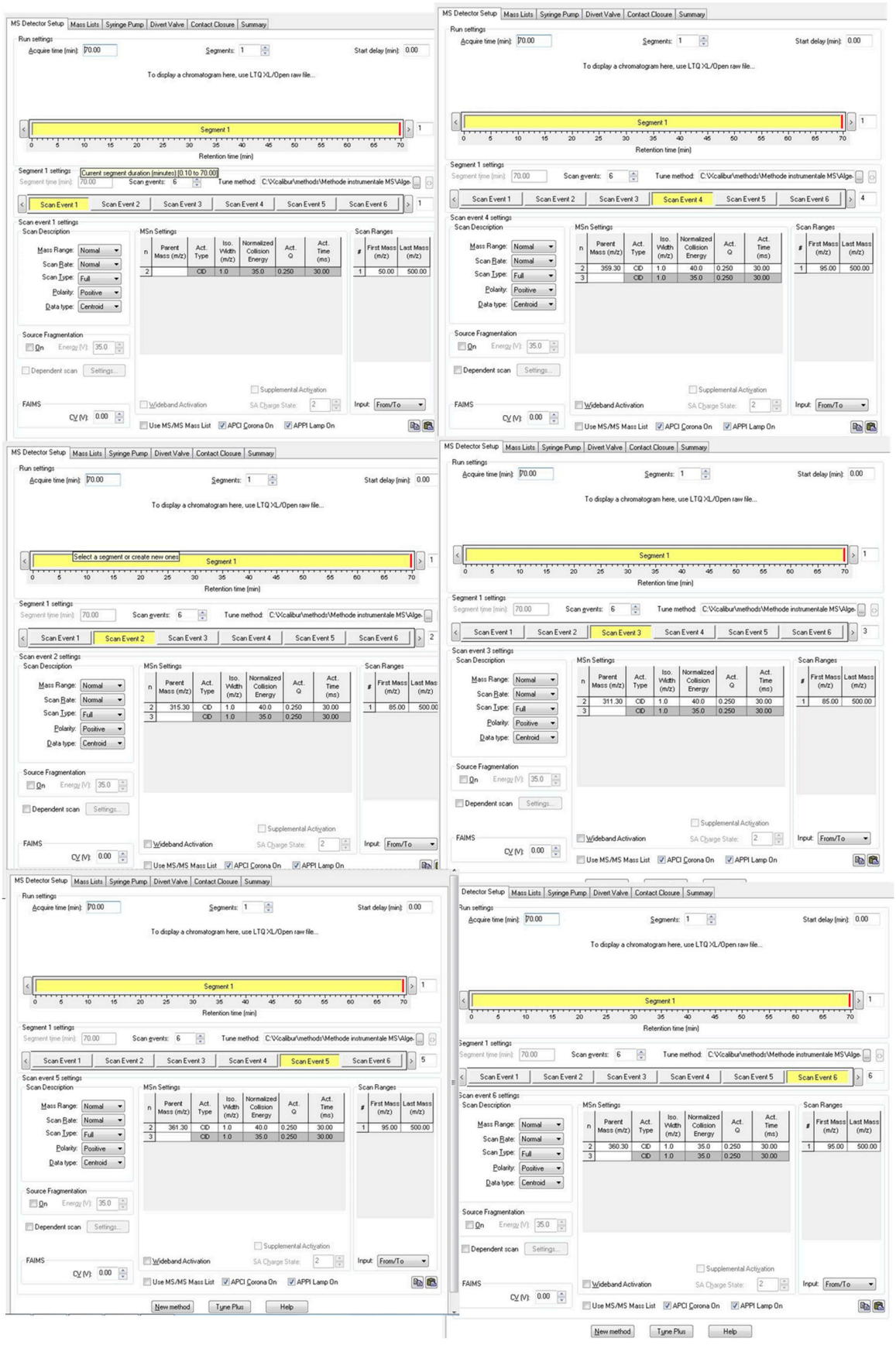


Figure A5. Details of the events used in LC-MS to confirm the identity of the cannabinoids.

Appendix B. The In Silico Analyses Identified Putative Endogenous *P. tricornutum* Candidates for CB-like Biosynthesis

To investigate the hypothesis that the endogenous enzyme could yield cannabinoids, characterization of the enzyme candidates was performed using bioinformatic tools. Putative candidate sequences for endogenous *P. tricornutum* enzymes that were able to convert OA and GPP into CBGA were obtained from the HMMER web server [64], using the CsAPT4 amino acid sequence as a query against the Protist database. We obtained five hits where the protein B7G488 (gene ID Phatr3_J37858) (57% similarity) and the protein B7FT51 (gene ID: Phatr3_J2738.t1) (51% similarity) had been annotated as the predicted protein members of the UbiA prenyltransferase family, to which CsAPT4 belongs [49]. The phylogenetic analysis indicated that these two proteins are clustered closer, according to sequence similarity, to the *C. sativa* APTs group (Appendix A, Figure A5). The soluble aromatic prenyltransferase from *Streptomyces* sp., NphB, had already been characterized by Kuzuyama et al. in 2005 [65], in a cluster with that from *Aspergillus terreus*, a fungal APT. No hits were found by using the same approach to identify endogenous diatom-soluble APT candidates, using *Streptomyces* sp. NphB as a query. Endogenous *P. tricornutum* candidate enzymes that were similar to THCA/CBDA synthase were not found in the database. One putative hypothetical candidate (with a 25.50% sequence identity compared to THCAS) from the diatom database was the violaxanthin deepoxidase-like protein (VDL), identified for its putative function in pigment deepoxidation and its requirement of a specific pH, which is similar to THCAS.

Figure A6. The phylogenetic tree of a putative candidate of endogenous *P. tricornutum* enzymes (in orange), which is similar to that of APT from cannabis and other species.

References

1. Aizpurua-Olaizola, O.; Soydaner, U.; Öztürk, E.; Schibano, D.; Simsir, Y.; Navarro, P.; Etxebarria, N.; Usobiaga, A. Evolution of the Cannabinoid and Terpene Content during the Growth of *Cannabis sativa* Plants from Different Chemotypes. *J. Nat. Prod.* **2016**, *79*, 324–331. [CrossRef] [PubMed]
2. Borrelli, F.; Pagano, E.; Romano, B.; Panzera, S.; Maiello, F.; Coppola, D.; De Petrocellis, L.; Buono, L.; Orlando, P.; Izzo, A.A. Colon Carcinogenesis Is Inhibited by the TRPM8 Antagonist Cannabigerol, a Cannabis-Derived Non-Psychotropic Cannabinoid. *Carcinogenesis* **2014**, *35*, 2787–2797. [CrossRef] [PubMed]
3. Chakravarti, B.; Ravi, J.; Ganju, R.K. Cannabinoids as Therapeutic Agents in Cancer: Current Status and Future Implications. *Oncotarget* **2014**, *5*, 5852–5872. [CrossRef] [PubMed]
4. Eibach, L.; Scheffel, S.; Cardebring, M.; Lettau, M.; Özgür Celik, M.; Morguet, A.; Roehle, R.; Stein, C. Cannabidivarin for HIV-Associated Neuropathic Pain: A Randomized, Blinded, Controlled Clinical Trial. *Clin. Pharmacol. Ther.* **2021**, *109*, 1055–1062. [CrossRef] [PubMed]
5. Brunt, T.M.; Bossong, M.G. The Neuropharmacology of Cannabinoid Receptor Ligands in Central Signaling Pathways. *Eur. J. Neurosci.* **2022**, *55*, 909–921. [CrossRef] [PubMed]

6. Schachtsiek, J.; Warzecha, H.; Kayser, O.; Stehle, F. Current Perspectives on Biotechnological Cannabinoid Production in Plants. *Planta Medica* **2018**, *84*, 214–220. [CrossRef] [PubMed]
7. Cohen, K.; Weinstein, A. The Effects of Cannabinoids on Executive Functions: Evidence from Cannabis and Synthetic Cannabinoids—A Systematic Review. *Brain Sci.* **2018**, *8*, 40. [CrossRef] [PubMed]
8. Vadivelu, N.; Kai, A.M.; Kodumudi, G.; Sramcik, J.; Kaye, A.D. Medical Marijuana: Current Concepts, Pharmacological Actions of Cannabinoid Receptor Mediated Activation, and Societal Implications. *Curr. Pain Headache Rep.* **2018**, *22*, 3. [CrossRef]
9. Gagne, S.J.; Stout, J.M.; Liu, E.; Boubakir, Z.; Clark, S.M.; Page, J.E. Identification of Olivetolic Acid Cyclase from *Cannabis sativa* Reveals a Unique Catalytic Route to Plant Polyketides. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 12811–12816. [CrossRef]
10. Valliere, M.A.; Korman, T.P.; Woodall, N.B.; Khitrov, G.A.; Taylor, R.E.; Baker, D.; Bowie, J.U. A Cell-Free Platform for the Prenylation of Natural Products and Application to Cannabinoid Production. *Nat. Commun.* **2019**, *10*, 565. [CrossRef]
11. Garfinkel, A.R.; Otten, M.; Crawford, S. SNP in Potentially Defunct Tetrahydrocannabinolic Acid Synthase Is a Marker for Cannabigerolic Acid Dominance in *Cannabis sativa* L. *Genes* **2021**, *12*, 228. [CrossRef] [PubMed]
12. Sirikantaramas, S.; Taura, F.; Tanaka, Y.; Ishikawa, Y.; Morimoto, S.; Shoyama, Y. Tetrahydrocannabinolic Acid Synthase, the Enzyme Controlling Marijuana Psychoactivity, Is Secreted into the Storage Cavity of the Glandular Trichomes. *Plant Cell Physiol.* **2005**, *46*, 1578–1582. [CrossRef]
13. Morita: How Structural Subtleties Lead to Molecular...—Google Scholar. Available online: https://scholar.google.com/scholar_lookup?title=How%20structural%20subtleties%20lead%20to%20molecular%20diversity%20for%20the%20type%20III%20polyketide%20synthases&journal=J%20Biol%20Chem&doi=10.1074%2Fjbc.REV119.006129&volume=294&pages=15121-15136&publication_year=2019&author=Morita%2CH&author=Wong%2CCP&author=Abe%2CI (accessed on 8 August 2023).
14. Abe, I. Biosynthesis of Medicinally Important Plant Metabolites by Unusual Type III Polyketide Synthases. *J. Nat. Med.* **2020**, *74*, 639–646. [CrossRef] [PubMed]
15. Pandith, S.A.; Ramazan, S.; Khan, M.I.; Reshi, Z.A.; Shah, M.A. Chalcone Synthases (CHSs): The Symbolic Type III Polyketide Synthases. *Planta* **2019**, *251*, 15. [CrossRef] [PubMed]
16. Nivina, A.; Yuet, K.P.; Hsu, J.; Khosla, C. Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chem. Rev.* **2019**, *119*, 12524–12547. [CrossRef] [PubMed]
17. Tian, B.; Liu, J. Resveratrol: A Review of Plant Sources, Synthesis, Stability, Modification and Food Application. *J. Sci. Food Agric.* **2020**, *100*, 1392–1404. [CrossRef] [PubMed]
18. Degenhardt, F.; Stehle, F.; Kayser, O. The Biosynthesis of Cannabinoids. In *Handbook of Cannabis and Related Pathologies*; Elsevier: London, UK, 2017; pp. 13–23.
19. Gaoni, Y.; Mechoulam, R. Isolation, Structure, and Partial Synthesis of an Active Constituent of Hashish. *J. Am. Chem. Soc.* **1964**, *86*, 1646–1647. [CrossRef]
20. Welling, M.T.; Deseo, M.A.; Bacic, A.; Doblin, M.S. Biosynthetic Origins of Unusual Cannabimimetic Phytocannabinoids in *Cannabis sativa* L: A Review. *Phytochemistry* **2022**, *201*, 113282. [CrossRef]
21. Taura, F.; Tanaya, R.; Sirikantaramas, S. Recent Advances in Cannabinoid Biochemistry and Biotechnology. *ScienceAsia* **2019**, *45*, 399. [CrossRef]
22. Livingston, S.J.; Quilichini, T.D.; Booth, J.K.; Wong, D.C.; Rensing, K.H.; Laflamme-Yonkman, J.; Castellarin, S.D.; Bohlmann, J.; Page, J.E.; Samuels, A.L. Cannabis Glandular Trichomes Alter Morphology and Metabolite Content during Flower Maturation. *Plant J.* **2020**, *101*, 37–56. [CrossRef]
23. Günnewich, N.; Page, J.E.; Köllner, T.G.; Degenhardt, J.; Kutchan, T.M. Functional Expression and Characterization of Trichome-Specific (-)-Limonene Synthase and (+)-α-Pinene Synthase from *Cannabis Sativa*. *Nat. Prod. Commun.* **2007**, *2*, 223–232.
24. Grof, C.P. Cannabis, from Plant to Pill. *Br. J. Clin. Pharmacol.* **2018**, *84*, 2463–2467. [CrossRef]
25. Fellermeier, M.; Zenk, M.H. Prenylation of Olivetolate by a Hemp Transferase Yields Cannabigerolic Acid, the Precursor of Tetrahydrocannabinol. *FEBS Lett.* **1998**, *427*, 283–285. [CrossRef]
26. Frontiers | Cannabis sativa: The Plant of the Thousand and One Molecules | Plant Science. Available online: https://www.frontiersin.org/articles/10.3389/fpls.2016.00019/full (accessed on 28 January 2022).
27. Balthazar, C.; Novinscak, A.; Cantin, G.; Joly, D.L.; Filion, M. Biocontrol Activity of *Bacillus* spp. and *Pseudomonas* spp. against Botrytis Cinerea and Other Cannabis Fungal Pathogens. *Phytopathology* **2022**, *112*, 549–560. [CrossRef]
28. Poulos, J.L.; Farnia, A.N. Production of Cannabidiolic Acid in Yeast 2018. U.S. Patent US20180073043A1, 15 March 2018.
29. Poulos, J.L.; Farnia, A.N. Production of Tetrahydrocannabinolic Acid in Yeast 2019. U.S. Patent US10392635B2, 27 August 2019.
30. Thomas, F.; Schmidt, C.; Kayser, O. Bioengineering Studies and Pathway Modeling of the Heterologous Biosynthesis of Tetrahydrocannabinolic Acid in Yeast. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 9551–9563. [CrossRef]
31. Luo, X.; Reiter, M.A.; d'Espaux, L.; Wong, J.; Denby, C.M.; Lechner, A.; Zhang, Y.; Grzybowski, A.T.; Harth, S.; Lin, W. Complete Biosynthesis of Cannabinoids and Their Unnatural Analogues in Yeast. *Nature* **2019**, *567*, 123–126. [CrossRef]
32. Kufs, J.E.; Reimer, C.; Steyer, E.; Valiante, V.; Hillmann, F.; Regestein, L. Scale-up of an Amoeba-Based Process for the Production of the Cannabinoid Precursor Olivetolic Acid. *Microb. Cell Fact.* **2022**, *21*, 217. [CrossRef]
33. Tan, Z.; Clomburg, J.M.; Gonzalez, R. Synthetic Pathway for the Production of Olivetolic Acid in *Escherichia coli*. *ACS Synth. Biol.* **2018**, *7*, 1886–1896. [CrossRef]
34. Zirpel, B.; Degenhardt, F.; Martin, C.; Kayser, O.; Stehle, F. Engineering Yeasts as Platform Organisms for Cannabinoid Biosynthesis. *J. Biotechnol.* **2017**, *259*, 204–212. [CrossRef]

35. Dhaouadi, F.; Awwad, F.; Diamond, A.; Desgagné-Penix, I. Diatoms' Breakthroughs in Biotechnology: *Phaeodactylum tricornutum* as a Model for Producing High-Added Value Molecules. *Am. J. Plant Sci.* **2020**, *11*, 1632–1670. [CrossRef]
36. Fabris, M.; George, J.; Kuzhiumparambil, U.; Lawson, C.A.; Jaramillo-Madrid, A.C.; Abbriano, R.M.; Vickers, C.E.; Ralph, P. Extrachromosomal Genetic Engineering of the Marine Diatom *Phaeodactylum tricornutum* Enables the Heterologous Production of Monoterpenoids. *ACS Synth. Biol.* **2020**, *9*, 598–612. [CrossRef]
37. Kassaw, T.K.; Paton, A.J.; Peers, G. Episome-Based Gene Expression Modulation Platform in the Model Diatom *Phaeodactylum tricornutum*. *ACS Synth. Biol.* **2022**, *11*, 191–204. [CrossRef]
38. Karas, B.J.; Diner, R.E.; Lefebvre, S.C.; McQuaid, J.; Phillips, A.P.; Noddings, C.M.; Brunson, J.K.; Valas, R.E.; Deerinck, T.J.; Jablanovic, J. Designer Diatom Episomes Delivered by Bacterial Conjugation. *Nat. Commun.* **2015**, *6*, 6925. [CrossRef]
39. Karas, B.J.; Molparia, B.; Jablanovic, J.; Hermann, W.J.; Lin, Y.-C.; Dupont, C.L.; Tagwerker, C.; Yonemoto, I.T.; Noskov, V.N.; Chuang, R.-Y. Assembly of Eukaryotic Algal Chromosomes in Yeast. *J. Biol. Eng.* **2013**, *7*, 30. [CrossRef]
40. Karp, P.D.; Billington, R.; Caspi, R.; Fulcher, C.A.; Latendresse, M.; Kothari, A.; Keseler, I.M.; Krummenacker, M.; Midford, P.E.; Ong, Q. The BioCyc Collection of Microbial Genomes and Metabolic Pathways. *Brief. Bioinform.* **2019**, *20*, 1085–1093. [CrossRef]
41. Athey, J.; Alexaki, A.; Osipova, E.; Rostovtsev, A.; Santana-Quintero, L.V.; Katneni, U.; Simonyan, V.; Kimchi-Sarfaty, C. A New and Updated Resource for Codon Usage Tables. *BMC Bioinform.* **2017**, *18*, 391. [CrossRef]
42. Diamond, A.; Diaz-Garza, A.M.; Li, J.; Slattery, S.S.; Merindol, N.; Fantino, E.; Meddeb-Mouelhi, F.; Karas, B.J.; Barnabé, S.; Desgagné-Penix, I. Instability of Extrachromosomal DNA Transformed into the Diatom *Phaeodactylum tricornutum*. *Algal Res.* **2023**, *70*, 102998. [CrossRef]
43. Li-Beisson, Y.; Thelen, J.J.; Fedosejevs, E.; Harwood, J.L. The Lipid Biochemistry of Eukaryotic Algae. *Prog. Lipid Res.* **2019**, *74*, 31–68. [CrossRef]
44. Liu, B.; Benning, C. Lipid Metabolism in Microalgae Distinguishes Itself. *Curr. Opin. Biotechnol.* **2013**, *24*, 300–309. [CrossRef]
45. Chng, J.; Wang, T.; Nian, R.; Lau, A.; Hoi, K.M.; Ho, S.C.; Gagnon, P.; Bi, X.; Yang, Y. Cleavage Efficient 2A Peptides for High Level Monoclonal Antibody Expression in CHO Cells. *MAbs* **2015**, *7*, 403–412. [CrossRef]
46. Liu, Z.; Chen, O.; Wall, J.B.J.; Zheng, M.; Zhou, Y.; Wang, L.; Ruth Vaseghi, H.; Qian, L.; Liu, J. Systematic Comparison of 2A Peptides for Cloning Multi-Genes in a Polycistronic Vector. *Sci. Rep.* **2017**, *7*, 2193. [CrossRef]
47. Ma, J.; Gu, Y.; Xu, P. Biosynthesis of Cannabinoid Precursor Olivetolic Acid in Genetically Engineered Yarrowia Lipolytica. *Commun. Biol.* **2022**, *5*, 1239. [CrossRef]
48. George, J.; Kahlke, T.; Abbriano, R.M.; Kuzhiumparambil, U.; Ralph, P.J.; Fabris, M. Metabolic Engineering Strategies in Diatoms Reveal Unique Phenotypes and Genetic Configurations with Implications for Algal Genetics and Synthetic Biology. *Front. Bioeng. Biotechnol.* **2020**, *8*, 513. [CrossRef]
49. Munakata, R.; Kitajima, S.; Nuttens, A.; Tatsumi, K.; Takemura, T.; Ichino, T.; Galati, G.; Vautrin, S.; Bergès, H.; Grosjean, J.; et al. Convergent Evolution of the UbiA Prenyltransferase Family Underlies the Independent Acquisition of Furanocoumarins in Plants. *New Phytol.* **2020**, *225*, 2166–2182. [CrossRef]
50. Wiles, D.; Shanbhag, B.K.; O'Brien, M.; Doblin, M.S.; Bacic, A.; Beddoe, T. Heterologous Production of *Cannabis sativa*-Derived Specialised Metabolites of Medicinal Significance—Insights into Engineering Strategies. *Phytochemistry* **2022**, *203*, 113380. [CrossRef]
51. Chuberre, C.; Chan, P.; Walet-Balieu, M.-L.; Thiébert, F.; Burel, C.; Hardouin, J.; Gügi, B.; Bardor, M. Comparative Proteomic Analysis of the Diatom *Phaeodactylum tricornutum* Reveals New Insights Into Intra- and Extra-Cellular Protein Contents of Its Oval, Fusiform, and Triradiate Morphotypes. *Front. Plant Sci.* **2022**, *13*, 673113. [CrossRef]
52. Hussain, M.H.; Hong, Q.; Zaman, W.Q.; Mohsin, A.; Wei, Y.; Zhang, N.; Fang, H.; Wang, Z.; Hang, H.; Zhuang, Y.; et al. Rationally Optimized Generation of Integrated *Escherichia coli* with Stable and High Yield Lycopene Biosynthesis from Heterologous Mevalonate (MVA) and Lycopene Expression Pathways. *Synth. Syst. Biotechnol.* **2021**, *6*, 85–94. [CrossRef]
53. Taylor, G.M.; Hitchcock, A.; Heap, J.T. Combinatorial Assembly Platform Enabling Engineering of Genetically Stable Metabolic Pathways in Cyanobacteria. *Nucleic Acids Res.* **2021**, *49*, e123. [CrossRef]
54. Hensing, M.C.M.; Rouwenhorst, R.J.; Heijnen, J.J.; van Dijken, J.P.; Pronk, J.T. Physiological and Technological Aspects of Large-Scale Heterologous-Protein Production with Yeasts. *Antonie Leeuwenhoek* **1995**, *67*, 261–279. [CrossRef]
55. Okorafor, I.C.; Chen, M.; Tang, Y. High-Titer Production of Olivetolic Acid and Analogs in Engineered Fungal Host Using a Nonplant Biosynthetic Pathway. *ACS Synth. Biol.* **2021**, *10*, 2159–2166. [CrossRef]
56. Trantas, E.; Panopoulos, N.; Ververidis, F. Metabolic Engineering of the Complete Pathway Leading to Heterologous Biosynthesis of Various Flavonoids and Stilbenoids in Saccharomyces Cerevisiae. *Metab. Eng.* **2009**, *11*, 355–366. [CrossRef]
57. Favero, G.R.; de Melo Pereira, G.V.; de Carvalho, J.C.; de Carvalho Neto, D.P.; Soccol, C.R. Converting Sugars into Cannabinoids—The State-of-the-Art of Heterologous Production in Microorganisms. *Fermentation* **2022**, *8*, 84. [CrossRef]
58. Barbosa, M.J.; Janssen, M.; Südfeld, C.; D'Adamo, S.; Wijffels, R.H. Hypes, Hopes, and the Way Forward for Microalgal Biotechnology. *Trends Biotechnol.* **2023**, *41*, 452–471. [CrossRef]
59. Sharma, N.; Fantino, E.I.; Awwad, F.; Mérindol, N.; Augustine, A.; Meddeb, F.; Desgagné-Penix, I. Impact of Different Light Characteristics on the Growth and Lipid Content of Diatom *Phaeodactylum tricornutum* Transconjugant Strains. *Am. J. Plant Sci.* **2023**, *14*, 41–63. [CrossRef]
60. Kouprina, N.; Larionov, V. Recent Advances in Chromosome Engineering. *Chromosome Res.* **2015**, *23*, 1–5. [CrossRef]

61. Slattery, S.S.; Diamond, A.; Wang, H.; Therrien, J.A.; Lant, J.T.; Jazey, T.; Lee, K.; Klassen, Z.; Desgagné-Penix, I.; Karas, B.J. An Expanded Plasmid-Based Genetic Toolbox Enables Cas9 Genome Editing and Stable Maintenance of Synthetic Pathways in *Phaeodactylum tricornutum*. *ACS Synth. Biol.* **2018**, *7*, 328–338. [CrossRef]
62. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat. Methods* **2012**, *9*, 676–682. [CrossRef]
63. Niu, Y.F.; Zhang, M.H.; Li, D.W.; Yang, W.D.; Liu, J.S.; Bai, W.B.; Li, H.Y. Improvement of neutral lipid and polyunsaturated fatty acid biosynthesis by overexpressing a type 2 diacylglycerol acyltransferase in marine diatom Phaeodactylum tricornutum. *Mar Drugs* **2013**, *11*, 4558–4569. [CrossRef]
64. Shoichet, B.; Mysinger, M.M.; Carchia, M.; Irwin, J.J.; Shoichet, B.K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–6594.
65. Kuzuyama, T.; Noel, J.P.; Richard, S.B. Structural Basis for the Promiscuous Biosynthetic Prenylation of Aromatic Natural Products. *Nature* **2005**, *435*, 983–987. [CrossRef]