

RESEARCH

Open Access



# Impact of heterologous expression of *Cannabis sativa* tetraketide synthase on *Phaeodactylum tricornutum* metabolic profile

Nicolas Sene<sup>1</sup>, Karen Cristine Gonçalves dos Santos<sup>1</sup>, Natacha Merindol<sup>1</sup>, Sarah-Eve Gélina, S.-E., Alexandre Custeau<sup>1</sup>, Fatima Awwad<sup>1</sup>, Elisa Fantino<sup>1</sup>, Fatma Meddeb-Mouelhi<sup>1,2</sup>, Hugo Germain<sup>1,2</sup> and Isabel Desgagné-Penix<sup>1,2\*</sup>

## Abstract

**Background** Pharmaceutical safety is an increasing global priority, particularly as the demand for therapeutic compounds rises alongside population growth. Phytocannabinoids, a class of bioactive polyketide molecules derived from plants, have garnered significant attention due to their interaction with the human endocannabinoid system, offering potential benefits for managing a range of symptoms and conditions. Traditional extraction from cannabis plants poses regulatory, environmental, and yield-related challenges. Consequently, microbial biosynthesis has emerged as a promising biotechnological alternative to produce cannabinoids in a controlled, scalable, and sustainable manner. Developing diatom-based biofactories represent a crucial step in advancing this biotechnology, enabling the efficient production of high-valued compounds such as cannabinoids.

**Results** We engineered the diatom *Phaeodactylum tricornutum*, a unicellular photosynthetic model organism prized for its naturally high lipid content, to produce olivetolic acid (OA), a key metabolic precursor to most cannabinoids. The genes encoding tetraketide synthase and olivetolic acid cyclase from cannabis were cloned onto episomal vectors and introduced using bacterial conjugation in two separate *P. tricornutum* transconjugant lines to evaluate enzyme activity and OA production in vivo. Both genes were successfully expressed, and the corresponding enzymes accumulated within the transconjugant lines. However, despite testing the cell extracts individually and in combination, OA accumulation was not detected suggesting potential conversion or utilization of OA by endogenous metabolic pathways within the diatoms. To investigate this further, we analyzed the impact of C5TKS expression on the diatom's metabolome, revealing significant alterations that may indicate metabolic flux redirection or novel pathway interactions.

**Conclusions** Our study demonstrates the successful expression of cannabinoid biosynthetic genes in *P. tricornutum* but highlights challenges in OA accumulation, likely due to endogenous metabolic interactions. These findings underscore the complexity of metabolic engineering in diatoms and suggest the need for further pathway optimization and metabolic flux analysis to achieve efficient cannabinoid biosynthesis. This research contributes to advancing sustainable biotechnological approaches for cannabinoid production.

**Keywords** Diatoms, Cannabinoids, Metabolic engineering, Polyketides, Photosynthetic biofactory, Untargeted metabolomic, Shikimate pathway

\*Correspondence:

Isabel Desgagné-Penix

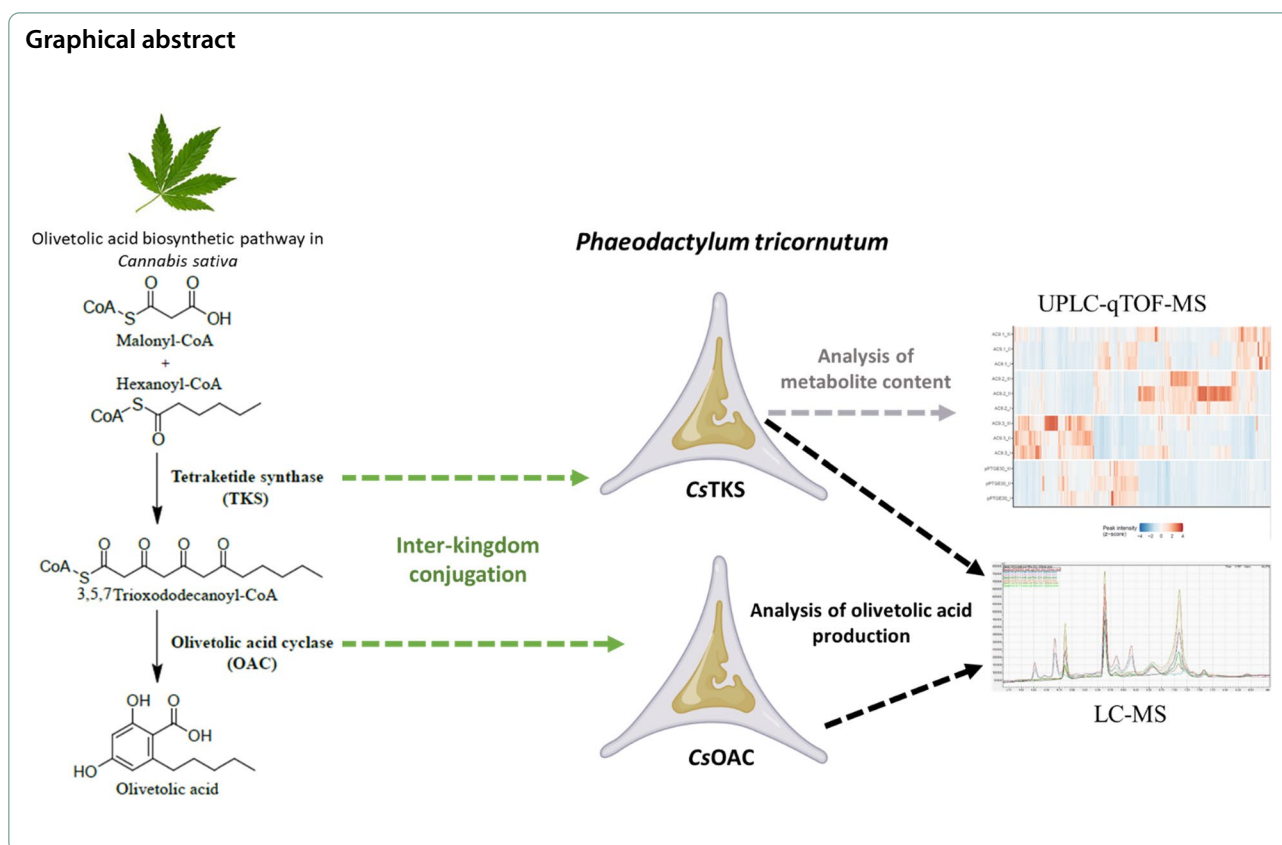
[Isabel.Desgagne-Penix@uqtr.ca](mailto:Isabel.Desgagne-Penix@uqtr.ca)

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Graphical abstract

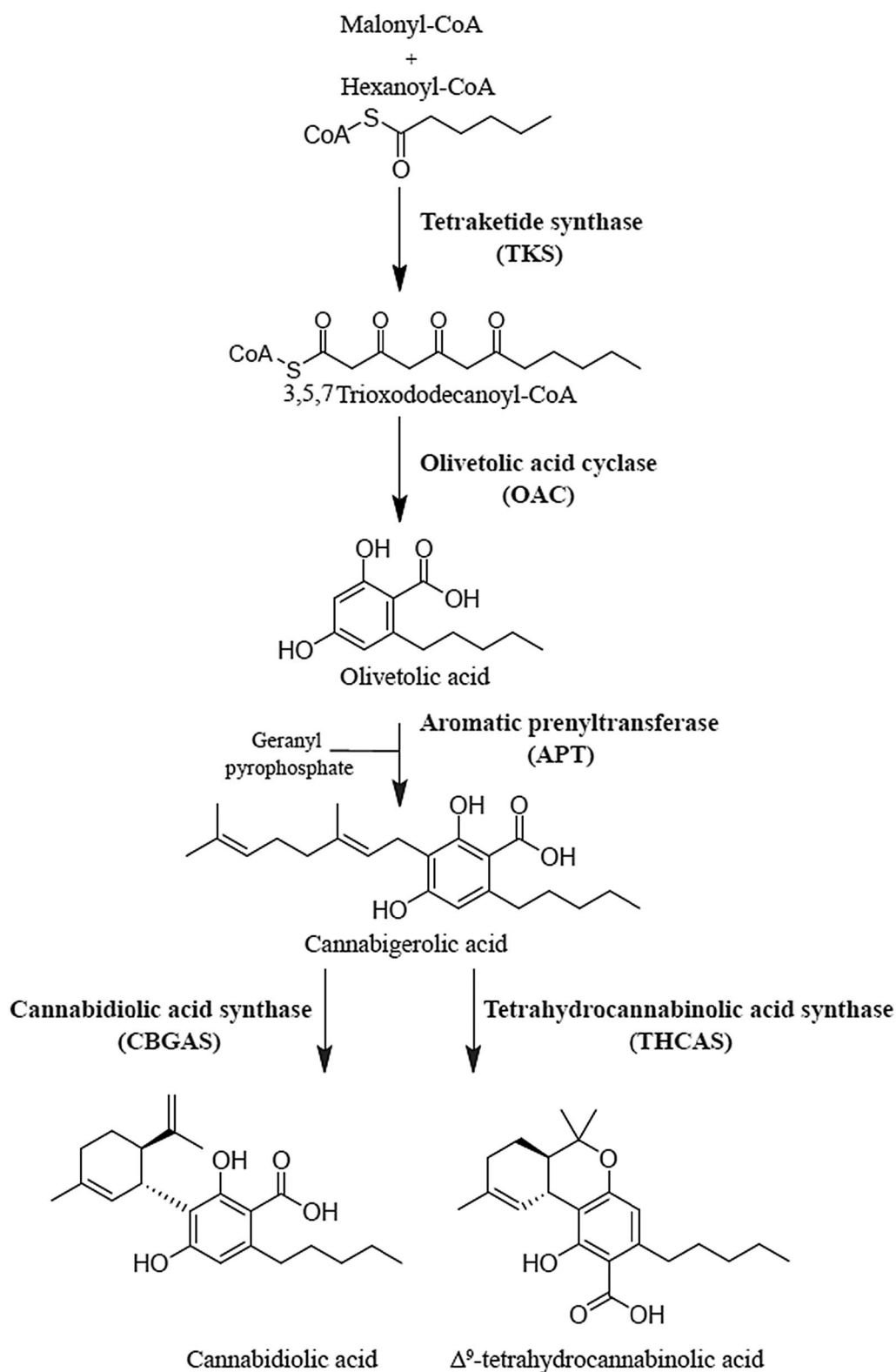


## Background

Cannabinoids are a class of molecules notorious for their recreational properties. They are naturally produced in highest abundance in the glandular trichomes of the female flowers of the *Cannabis* genus. The plant of cannabis was used for millennia for its anti-inflammatory, antiemetic, and as a pain relief. Those properties are due to the ability of cannabinoids to interact with the mammalian endocannabinoid system involved in several metabolic functions [1–3]. Today, more than a hundred terpenophenolic compounds in the glandular trichomes of *Cannabis sativa* (*C. sativa*) have been identified as cannabinoids. Among them,  $\Delta^9$ -tetrahydrocannabinol (THC) and cannabidiol (CBD) are the most well-known and abundant [4, 5]. Recently, the interest in cannabinoids focused on their potential in fighting neurodegenerative disease such as Alzheimer's and Parkinson's diseases [1, 6]. The pathway to biosynthesize THC and CBD starts from hexanoyl-CoA and malonyl-CoA moieties, the tetraketide synthase (TKS) condenses hexanoyl-CoA with three malonyl-CoA to produce 3,5,7-trioxododecanoyl-CoA (TdCoA) compound which is then cyclized in Olivetolic acid (OA) by the Olivetolic acid cyclase (OAC). An aromatic prenyltransferase (APT) fuses OA with the prenyl group of geranyl

pyrophosphate (GPP) to produce cannabigerolic acid (CBGA). Cannabidiol acid synthase (CBDAS) and tetrahydrocannabinol acid synthase (THCAS) catalyze the conversion of CBGA into cannabidiolic acid (CBDA) and  $\Delta^9$ -tetrahydrocannabinolic acid (THCA), respectively. As a result, cannabinoids are classified as terpenophenols, characterized by a cyclic monoterpene moiety derived from GPP and an alkylated diphenol backbone originating from OA [7]. Upon combustion or heat exposure, CBDA and THCA undergo decarboxylation, yielding the bioactive compounds cannabidiol (CBD) and  $\Delta^9$ -tetrahydrocannabinol (THC), respectively (Fig. 1) [8–10]. The increasing recognition of cannabinoids for their therapeutic and commercial value has led to a surge in cannabis cultivation, with the global legal market projected to exceed US\$400 billion by 2030 [11–13]. However, cannabis cultivation faces several challenges, including stringent regulatory frameworks, the limited yield of specific cannabinoids due to their restricted production in female flowers, and complex extraction and purification processes. These limitations have driven the search for alternative biotechnological production systems [14, 15].

Bioengineering of model organisms has increasingly advanced, both to confirm biological knowledge, and as an alternative way to produce compounds of interest.



**Fig. 1** Scheme of the proposed cannabinoids synthesis in *Cannabis sativa* from hexanoyl-CoA and malonyl-CoA. Arrows represent enzymatic reactions. Enzymes' names are in bold

It does not only allow to circumvent all the constraints associated with cultivation, but also has a higher organism mass volume ratio. Attempts of producing cannabinoids in *E. coli* have been performed on purified TKS using synthetic *N*-acetylcysteamine as a substrate and leading to the production of olivetol. However, attempts of producing THCA never succeeded due to the complex folding of THCAS therefore considering the bacteria unsuitable for the production of cannabinoids synthase [16]. To overcome the limits of bacteria, the production of cannabinoids has been achieved in yeast with the limit of reaching poor yields despite supplementation [9, 17]. Algae are promising heterologous hosts, as the number of available tools and data gathered have dramatically increased over recent years [18–20]. Compared to conventional platforms, algae offer additional advantages, such as the ability to fix CO<sub>2</sub> through photosynthesis. In 2019, a patent reported that cannabinoids can be heterologously produced and harvested in microalgae *Chlamydomonas reinhardtii* and *Synechococcus elongatus* [21]. Compared to other microalgae, a unique advantage of diatoms is their estimated contribution to around 20% of Earth's oxygen production [22, 23]. Among diatoms, *Phaeodactylum tricornutum* has emerged as valuable platform to produce fatty acid-derived molecules, due to its fast-growing rate, high lipid content, and natural production of omega-3 fatty acids like eicosapentaenoic acid [24, 25]. Efficient protocols have been developed for the extraction of apolar compounds such as cannabinoids [26, 27]. *P. tricornutum* also naturally produces malonyl-CoA [28], GPP [29], and is predicted to produce hexanoyl-CoA [30], all essential precursors for cannabinoid biosynthesis. Together, these features suggest that *P. tricornutum* is a suitable model organism for the heterologous biosynthesis of cannabinoids. Indeed, the introduction of *Streptomyces* sp. aromatic prenyltransferase NphB [31], and CBDAS [32] into *P. tricornutum*, has led to the in vitro production of CBGA and CBDA, respectively. Additionally, co-expression of a fused construction of the first two enzymes of the pathway—TKS and OAC from *C. sativa*—in *P. tricornutum* lead to the transient production of trace amounts of OA [33]. The reason for the loss of OA accumulation over time remains unclear. We hypothesized that expressing TKS and OAC separately in *P. tricornutum* strains, rather than as a fused construct, might enhance enzyme functionality and stability, thereby improving OA synthesis. Therefore, in this study, we aimed to produce olivetolic acid in *P. tricornutum* by inserting TKS and OAC enzymes in different transconjugants and test their activity individually or pulled together. In order to observe the extend of CsTKS impact on the metabolite content and deepen our understanding on the potential compatibility of *P. tricornutum*

to produce cannabinoids, using UPLC–qTOF-MS we assessed the impact of CsTKS accumulation on transconjugants metabolite content.

## Material and methods

### Microbial strains and growth conditions

*Escherichia coli* strains Epi300 (TransforMax Epi300 Biosearch™ Technologies, Novato, CA, USA) and 10β (New England Biolab®, ON, Canada) were grown in Luria Bertani broth (LB) at 37 °C with 220 rpm agitation when in liquid medium, supplemented with chloramphenicol (35 µg mL<sup>-1</sup>) (Thermo Fisher Scientific, Ottawa, ON, Canada) and gentamicin (50 µg mL<sup>-1</sup>) (Thermo Fisher Scientific, Ottawa, ON, Canada) when containing pTA-MOB.

*Saccharomyces cerevisiae* VL6-48 strain was grown following [33, 34]. Cultures in rich medium were done in yeast extract peptone dextrose YEPD while cultures in minimum media were done in complete minimal (CM) medium lacking histidine and uracil or tryptophan and uracil (Teknova, Hollister, CA, USA) at 30 °C.

The *P. tricornutum* strain (CCAP 1055/1, Culture Collection of Algae and Protozoa, RRID:NCBITaxon\_2850) was grown in L1 medium without silica at 18 °C under white fluorescent lights (75 µE m<sup>-2</sup> s<sup>-1</sup>) with a photoperiod of 16 h of light and 8 h of darkness under the agitation of 130 rpm for liquid cultures with zeocin (50 µg mL<sup>-1</sup>) (Invitrogen, Burlington, ON, Canada) when cultivating transconjugant cells containing the *Sh Ble* resistance gene [35].

### Plasmid constructions and insertion in *Phaeodactylum tricornutum*

Recombinant plasmids, named pTKS for CsTKS and pOAC for CsOAC, were assembled through yeast assembly using *S. cerevisiae* strain VL6-48 (ATCC MYA-3666: MATα his3-Δ200 trp1-Δ1 ura3-52 lys2 ade2-1 met14 cir<sup>0</sup>) exactly as described in [34]. The cassette containing the CsTKS under the control of *40SRPS8* promoter and *fcpA* terminator was inserted in the pPtGE30 plasmid, named EV hereafter, composed of a pCC1BAC (RRID:Addgene\_62862) backbone, containing a bleomycin resistance gene (*Sh Ble*) as a selection marker, and a *P. tricornutum* region contributing to plasmid stability.

Assembled pTKS and pOAC episomes were transformed into the *E. coli* Epi 300 strain containing pTA-MOB plasmid. The cargo plasmid was then transferred in the diatoms via conjugation as described in the literature [36]. Transconjugants of *P. tricornutum* were then selected following the protocol of [31] leading to a liquid culture of *P. tricornutum* clones in L1 supplemented with zeocin. Conjugations were verified by colony PCR.

Positive clones were selected for plasmid extraction then sent for Next Generation Sequencing (NGS).

### Growth kinetics

Optical density of the clones was measured at 730 nm ( $OD_{730nm}$ ) every two days for 14 days. For each of them, 250  $\mu$ L was added to a Microplate 96 Well, PS, F-Bottom Chimney Well  $\mu$ CLEAR<sup>®</sup>, black (Greiner Bio-One, Kremsmünster, Austria) and absorbance was recorded using Synergy H1 BioTek microplate reader (Agilent, Santa Clara, CA, USA). Curves were obtained using GraphPad Prism software (RRID:SCR\_002798). A *t*-test was performed using CGGC permutation test from Walter and Eliza Hall Institute of Medical Research with 10,000 permutations [37].

### Plasmid rescue

For each transconjugant lines, 5 mL were centrifuged at 4700 g for 10 min after 7 days of culture. After supernatant removal, 235  $\mu$ L of PDL1 Large Plasmid Extraction Kit (Geneaid<sup>™</sup>, New Taipei City, Taiwan), together with 5  $\mu$ L Lysozymes (100 mg mL<sup>-1</sup>) and 5  $\mu$ L hemicellulase (25 mg mL<sup>-1</sup>) were added to resuspend the pellet. Samples were incubated at 37 °C for 30 min. 250  $\mu$ L of PDL2 solution was added to the mix before mixing by inverting ten times, then 375  $\mu$ L of PDL3 was added before mixing. The rest of the procedure was performed according to the manufacturer protocol until elution in 50  $\mu$ L of elution buffer preheated at 80 °C. DNA concentration was measured with a N60/N50 Implen NanoPhotometer<sup>®</sup>. Purified plasmid DNA was then used to transform *E. coli* 10 $\beta$  strains from New England Biolabs (Whitby, ON, Canada) following their protocol. Bacteria were plated on LB with chloramphenicol (35 mg mL<sup>-1</sup>). After inoculating 5 mL LB with identical concentration of chloramphenicol, the liquid culture was incubated over night at 37 °C under agitation. Plasmid DNA was extracted by miniprep (Geneaid<sup>™</sup>, New Taipei City, Taiwan). Purified plasmid DNA concentration was measured with the NanoPhotometer and the quality of the DNA sample was verified by digestion with restriction enzymes. The plasmid DNA sample was then used for NGS at Massachusetts General Hospital DNA Core.

### Western blot

Approximately 160 mg of wet mass was resuspended in 1X Laemmli in PBS at a ratio of 1 mL:750 mg<sup>-1</sup>, heated at 95 °C for 10 min, and centrifuged 5 min at 14000g at 4 °C. At least 500  $\mu$ g of protein was loaded on a 10% SDS–polyacrylamide gel and migrated at 80 V for 25 min, then at 120 V for 1 h. The gel was then transferred to a 0.2  $\mu$ m polyvinylidene fluoride (PVDF) membrane, previously activated in methanol and then equilibrated in transfer

buffer (25 mM Tris, 192 mM glycine and 20% methanol). Transfer was carried out at 100 V, 400 mA, for 2 h in transfer buffer on ice. The membrane was blocked in 5% milk in Tris Buffer Saline (TBS) with 0.1% Tween 20 for 50 min at room temperature. Blots were washed three times for 10 min with TBST before being incubated with 6x-His monoclonal antibody (Invitrogen, Canada, RRID:AB\_557403) (dilution 1:1 000) in 3% bovine serum albumin overnight at 4 °C. After three washes in TBST, blots were incubated for 1 h with a 1:20000 dilution, in 5% milk of Immun-Star Goat Anti-Mouse (GAM)-HRP from Bio-Rad (Ontario, Canada, RRID:AB\_11125753). After three more washes in TBST protein detection was performed using Clarity Max, 10 ng of Multiple Tag (GenScript, Brockville, Ontario, Canada) were used as positive control. Red ponceau was applied for 2 min at room temperature and then rinsed with distilled water. Western ECL Substrate-Luminol solution from Bio-Rad. Images were acquired using ChemiDoc Imaging System (RRID:SCR\_019037) with Image Lab<sup>™</sup> Software (RRID:SCR\_014210) (Bio-Rad, Hercules, CA, USA).

### Protein extraction

For each sample, 500  $\mu$ L of cell buffer (0.75 mM SDS, 10% Glycerol, 51.4 mM, Tris HCl pH8, 0.02 mM EDTA) was added to 125 mg of pellets. PMSF and protease inhibitor were added at a ratio of 1/10 and 1/100, then proteins were extracted by sonication with the following parameters: amplitude 35%, time 3 min, ON 30 s, OFF 25 s (Fisherbrand<sup>™</sup> Model 505 Sonic Dismembrator). Extracts were then spun for 40 min at 4 °C at 14000g and the supernatant was removed and spun a second time in the same conditions. The new supernatant was flash frozen and kept at – 80 °C until use.

### Enzymatic assay

Hexanoyl-CoA and malonyl-CoA were purchased from Millipore Sigma Canada Ltd (Oakville, ON, Canada) and the olivetol and olivetolic acid used were purchased from Santa Cruz biotechnologies (Dallas, TX, United states). Protein concentration was measured by colorimetric RC DC<sup>™</sup> Protein Assay (Bio-Rad, Hercules, CA, USA). Every test was done in triplicate in 4-(2-hydroxyethyl)–1-piperazineethanesulfonic acid (HEPES) buffer 0.05 M pH 7, DTT 0.005 M, with 400  $\mu$ g of TKS and/or OAC protein extract (800  $\mu$ g total protein extract when pulled together), 0.2 mM hexanoyl-CoA, and 1 mM malonyl-CoA, completed to 100  $\mu$ L with water. The reaction was performed at 25 °C for 16 h and was stopped with trichloroacetic acid 2% v/v. 300  $\mu$ L of methanol was added to the reaction mix then was filtered on PTFE 0.2  $\mu$ m filter and transferred into new tubes. The samples were dried using a vacuum concentrator (SpeedVac Thermo Savant



SPD 2010), then resuspended in 100  $\mu$ L methanol before analyzing olivetolic acid production by high-performance liquid chromatography with diode-array detection (HPLC–DAD).

#### HPLC–DAD analysis

An InfinityLab Poroshell 120 EC-C18 (4.6  $\times$  100 mm, 2.7 mm; Agilent Technologies, Mississauga, ON, Canada) column, maintained at 30 °C, was used for separation. For each sample, 10  $\mu$ L was injected into the HPLC device. The mobile phase used was composed at 30% of (A) formic acid 0.1% v/v in milli-Q water and 70% of (B) formic acid 0.1% v/v in methanol. Samples were processed at a flow rate of 1 mL min<sup>−1</sup> following 0 min, 70% B; 1.0 min, 70% B; 6.0 min, 77% B; 15.0 min, 90% B; 15.1 min, 70% B and 18.0 min, 70% B. The total run time was 18.5 min allowing the reconditioning of the column for the next injection. D2 (deuterium) lamps acquired wavelengths from 190 to 400 nm. The analyses were done using UV wavelength at 200 nm. The maximum absorption wavelength and the retention times of each detected signal in samples were compared to standards to allow their identification.

#### Untargeted metabolomic

##### Metabolite extraction

Metabolites were extracted from three biological replicates for each clone. Each replicate was inoculated at OD<sub>730</sub> 0.1 in 200 mL, homogenized, then separated into four 50 mL volumes in 250-mL flasks to optimize the growth rate. Cultures were grown for an additional week, then for each replicate, the 50 mL cultures were pooled together. Cultures were then centrifuged for 10 min at 1500g at 4 °C and pellets were weighed. 1 mL of ethanol 95% (Fisher Chemicals, Pittsburgh, PA, USA) per 100 mg of wet biomass was added to the pellet. Samples were mixed by vortexing and incubated overnight at −20 °C. The following day, the samples were centrifuged at 4000g for 10 min at 4 °C and supernatants were filtered using polytetrafluoroethylene (PTFE) 0.2  $\mu$ m filter. All filtrates were dried using a vacuum concentrator (SpeedVac Thermo Savant SPD 2010) without heating, for 3 h at maximum ramp.

##### Data acquisition

Metabolomic extracts that were analyzed by reversed-phase chromatography (RP) were resuspended in 100  $\mu$ L of pure water, while the ones analyzed using hydrophilic interaction liquid chromatography (HILIC) were reconstituted in 100  $\mu$ L of mobile phase (ammonium formate 10 mM in acetonitrile/water (95:5), pH 3.8). Samples were vortexed and sonicated for 2 min, followed by a 5 min 16000g centrifugation. Supernatants were collected

and diluted 1:100 in their initial solvents followed by a filtration using 0.45  $\mu$ m AcroPrep filters in a 5 mm Hg vacuum. Acquisition was done with an Eksigent  $\mu$ U HPLC (Eksigent, Redwood City, CA, USA) in conjunction with an ABSciex TripleTOF 6600 (ABSciex, Foster City, CA, USA) system with an electrospray interface. For data processing, acquisition and instrument control, the Analyst TF version 1.8 software was used. Data acquisition was performed in Data Dependent Acquisition (DDA) and SWATH mode for positive ionization mode. The source voltage was set to 5.5 kV, temperature at 220 °C, curtain gas at 50 psi, ion source gas 1 at 40 psi, and ion source gas 2 at 50 psi. For all methods, the mobile phase was composed of: A (water with 0.1% formic acid and water with 10 mM ammonium fluoride) and solvent B (acetonitrile with 0.1% formic acid and acetonitrile: water (95:5) with 10 mM NH<sub>4</sub>F). Columns Luna Omega Polar 100  $\times$  2.1 1.6  $\mu$ m (RP positive) and BEH Z-HILIC 100  $\times$  2.1 Particle 1.7  $\mu$ m (HILIC) from Waters were kept at 50 °C. The gradient began at 0% B for reversed-phase, and 100% for HILIC.

#### Bioinformatics analysis

Raw mass spectra files in profile mode were aligned, deconvoluted and annotated with MS-Dial version 5.1.230912 (RRID: SCR\_023076) [38–40]. The parameters were set as follows for all three analyses: retention time range of 0–10 min; mass range of 50–1 250 Da; MS1 and MS2 tolerances, for deconvolution and alignment, of 0.01 Da and 0.025 Da, respectively; minimum peak height and width of 1000 amplitude and 5 Da, respectively; the smoothing method used was linear-weighted moving average, with smoothing level of 3 scans and mass slice of 0.1 Da; and the sigma window value was set to 0.5. For analyte identification, the following parameters were set: accurate mass tolerance 0.01 Da for MS1 and 0.05 Da for MS2; retention time tolerance was set to 0.1 min; adduct ions were set to [M+H]<sup>+</sup> and [M+NH<sub>4</sub>]<sup>+</sup> for ESI<sup>+</sup> and [M-H]<sup>−</sup> for ESI<sup>−</sup>. Deconvoluted spectra were compared to MassBank ESI<sup>+</sup> and ESI<sup>−</sup> [41] and MS-Dial's "ESI<sup>+</sup>-" and "ESI<sup>−</sup>-MS/MS from authentic standards" databases version 19 (all databases were obtained as MSP files from MS-Dial's webpage on March 2024) for annotation. These comparisons were done with the default parameters, except for MS1 and MS2 mass tolerances, set to 0.01 and 0.001 Da, respectively. Annotations marked as "reference match" by MS-Dial were validated by checking that main peaks from reference were present in the current experiment's spectra.

Peak height tables exported from MS-Dial were then used for comparative analyses in R version 4.3.3. RP ESI<sup>+</sup> and ESI<sup>−</sup> datasets included three blank samples, while

HILIC ESI<sup>+</sup> dataset included two blank samples. For each analyte, the abundance estimate was defined as

$$\text{peak\_height}_{\text{replicate}} - (3 * \max(\text{peak\_height}_{\text{blank\_samples}}))$$

If the abundance estimate was  $\leq 0$ , the analyte was considered absent. For each strain and for both quality control replicates (pooled samples), only analytes present in all replicates were considered present. Finally, only analytes considered present in at least one strain were selected for further analysis. The abundance estimation table was prepared for MetaboAnalystR version 4.0 (RRID:SCR\_016723) [42], which was used for replacing missing values, filter analytes (parameters: `qc.filter = T`, `rsd = 25`, `var.filter = "none"`, `var.cutoff = -1`, `int.filter = "mean"`, `int.cutoff = 0`) and normalize data (functions `PreparePreNormData` and `Normalization("GroupPQN", transNorm = "LogNorm", scaleNorm = "ParetoNorm", ref = "QC", ratio = F, ratioNum = 20)`). Next, MetaboAnalystR's filtered datasets were used to find significant differences in analyte abundances between pPtGE30 and AC9 strains using the function `sam` (parameters: `method = "d.stat"`, `B = 200`, `med = F`, `use.dm = T`, `var.equal = T`, `R.fold = 1`) from the R package `siggenes` version 1.76.0 [43]. Fold change was calculated as follows:

$$\log_2 \frac{\text{mean}(\text{abundance}_{\text{Ac9 clone}})}{\text{mean}(\text{abundance}_{\text{pPtGE30}})}$$

If  $p < 0.01$  and  $|\text{fold change}| > 1$ , the analyte was considered deregulated in the corresponding AC9 strain. Finally, principal component analysis (PCA) was performed using the combined normalized data from all three analyses (RP ESI<sup>+</sup>, RP ESI<sup>-</sup> and HILIC ESI<sup>+</sup>) with MetaboAnalystR function `PCA.Anal`.

## Results

### Heterologous accumulation of CsTKS and CsOAC in *P. tricornutum*

Two expression cassettes were designed to independently characterize CsTKS and CsOAC expression in *P. tricornutum*. The plasmid referred to as pTKS, contained the CsTKS coding sequence (CDS) under the control of the endogenous 40SRPS8 promoter, tagged with a 6xHis marker, and terminated with the *FcpA* terminator. The plasmid referred to as pOAC, included CsOAC CDS under the control of the *FcpD* promoter, tagged with a myc marker, and terminated with the *FcpD* terminator (Fig. 2). These plasmids along with the empty vector (EV), were introduced into *P. tricornutum* wild-type strains via interkingdom conjugation with *E. coli*. The plasmids harbored the *Sh ble* gene, conferring zeocin resistance, and

transconjugants were selected on L1 agar medium supplemented with zeocin.

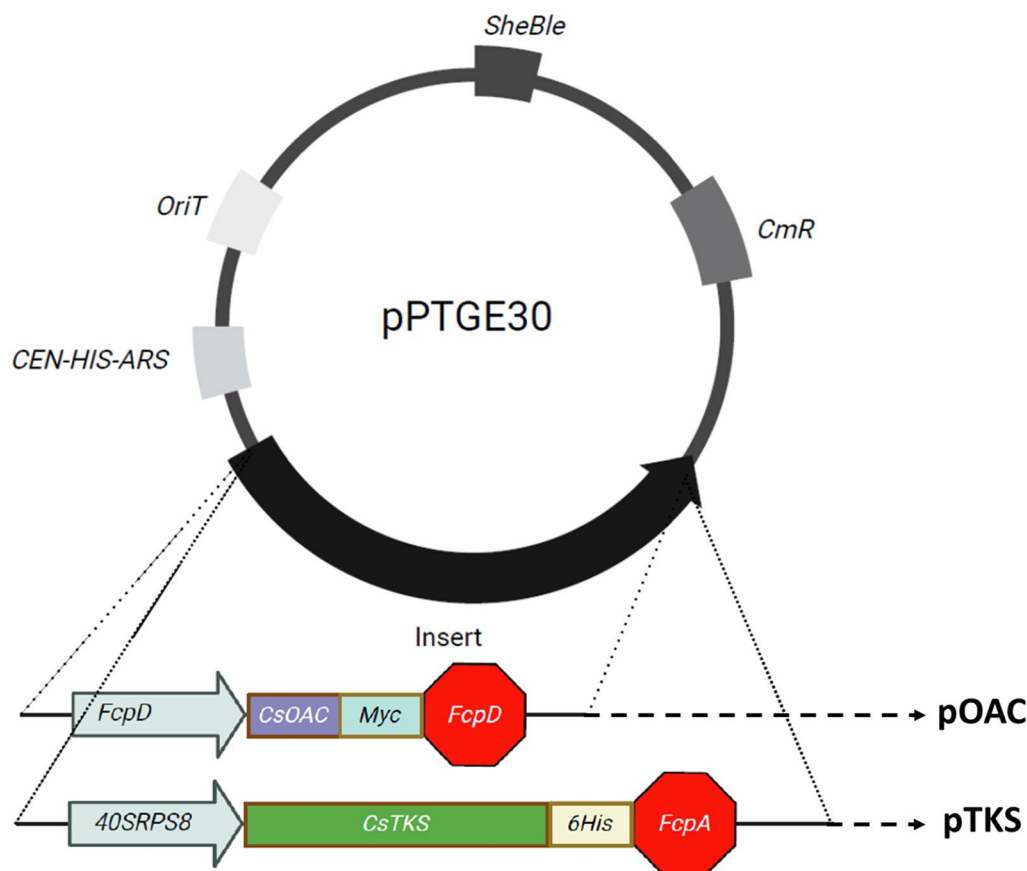
Three transconjugants of each construct were selected for plasmid rescue, and the extracted DNA was subjected to restriction digestion analysis before sequencing (Additional file 1: Fig. S1). Transconjugants showing the expected digestion profiles, were further analyzed by Next Generation Sequencing. While the pTKS sequences exhibited minor substitutions, including one in the zeocin resistance gene, two in the 40SRPS8 promoter, and one in the backbone (Additional file 1: Table S1), the CsOAC CDS remained intact. Despite the substitutions, all transconjugants displayed growth on zeocin-supplemented media, indicating that the resistance gene remained functional. Importantly, the heterologous expression of these enzymes did not impair the growth of the transconjugants (Additional file 1: Fig. S2).

Heterologous protein production was then verified by western blot. All pTKS-containing transconjugants exhibited a band at the expected size of 42 kDa, absent in EV along with an additional smaller band likely corresponding to a truncated version of the protein (Fig. 3). For the pOAC transconjugants, two (pOAC-5 and pOAC-7) displayed a band at the expected size of 12 kDa, while one transconjugants (pOAC-6) did not. Further sequencing revealed a frameshift in the *fcpD* promoter of pOAC-6, which disrupted the myc-tag and explained the absence of a signal in the western blot (Fig. 3).

To validate the added enzymatic activity from each of the recombinant proteins, total protein extracts from pTKS and pOAC were tested. Enzymatic assays were conducted using the extracts individually or by combining the CsTKS and CsOAC extracts. The resulting reactions were analyzed by HPLC-DAD, but none of the detected peaks corresponded to the OA standard (Additional file 1: Fig. S3). To rule out potential inhibitory effects from compounds present in the crude extract, protein purification was attempted. While TKS was successfully purified, purification OAC was not achieved.

### Heterologous expression of CsTKS induces metabolic changes in *P. tricornutum*

The production of a tetraketide is a critical initial step in cannabinoid biosynthesis. To assess the impact of CsTKS activity on the diatom metabolome, we compared the metabolic profiles of pTKS transconjugants with those of the EV control. Metabolite extracts were separated using either hydrophilic interaction liquid chromatography (HILIC) or reversed-phase chromatography (RP) followed by tandem mass spectrometry. HILIC-separated metabolites were analyzed in positive ionization mode, while RP-separated metabolites were analyzed in both



**Fig. 2** Schematic representation of pPTGE30 expression vector and insertion of *CsTKS* and *CsOAC* cassettes. The pTKS plasmid includes a construction that contains the 1971-bp cassette with *CsTKS* under control of native constitutive *40SRPS8* promoter, *6xHis* marker and an *FcpA* terminator. The pOAC plasmid harbors a construction with the 1265 bp cassette *CsOAC* under control of *FcpD* promoter, marked with a *Myc* tag, and ended with an *FcpD* terminator. The *Cen-His-Ars* region is necessary for episomal replication while the *OriT* permits the transfer of the plasmid in *P. tricornutum* during interkingdom conjugation. Finally, the *Sh ble* and *CmR* coding sequence allows for zeocin resistance when expressed in *P. tricornutum* and chloramphenicol resistance in *E. coli*, respectively

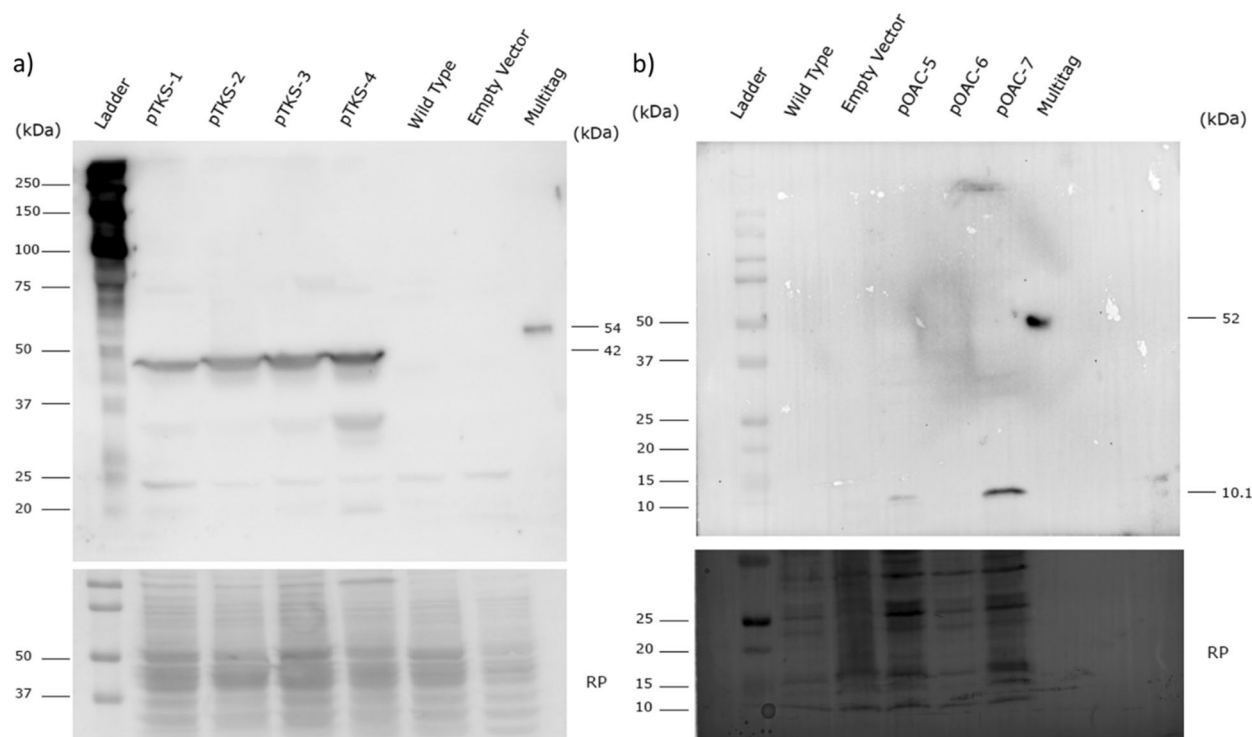
positive and negative ionization modes (referred to hereafter as HILIC positive, RP positive and RP negative).

HILIC positive and RP negative analyses detected over 3000 analytes each, with 19 and 8 metabolites annotated, respectively. RP positive analysis was more comprehensive, detecting over 7000 analytes, of which 29 were annotated (Fig. 4a). Among the metabolites detected, 12,348 compounds have been found in all replicates of each transconjugant, 160 analytes were uniquely present in *CsTKS* transconjugants and absent in the EV control. Each *CsTKS* transconjugant presented a different amount of exclusive analytes ranging from 48 to 120. A notable number of 278 analytes were present in all clones but pTKS-3 while only 66 excluded from pTKS-1 and 57 from pTKS-2. A total of 57 annotated analytes have been identified with 49 shared among all clones. No exclusive analyte from a clone was annotated. Three were exclusively shared by the pTKS clones: pyocyanin, a

phenazine pigment derived from the shikimate pathway in *Pseudomonas aeruginosa*; 2-benzyl-4-chlorophenol (chlorophene), a phenolic halogenated compound with an unidentified biosynthetic route; and abiestic acid, a diterpene biosynthesized from GPP. Notably, 18 analytes detected in the EV transconjugants were absent in all three pTKS transconjugants (Fig. 4b).

The second parameter analyzed was the quantitative impact of *CsTKS* expression on the diatom metabolome (Fig. 5a). Analytes were heterogeneously deregulated among the three pTKS transconjugants compared to the three replicates of EV transconjugants, with fold changes ranging from  $-10.3$  (analyte with retention time (RT)=3.768 min and with a  $m/z$  of 771.15744 in pTKS-3) to  $11.2$  (RT=1.449 min;  $m/z$ =431.29585 in pTKS-1). However, 48 analytes were consistently deregulated across all pTKS transconjugants triplicates (Fig. 5b).



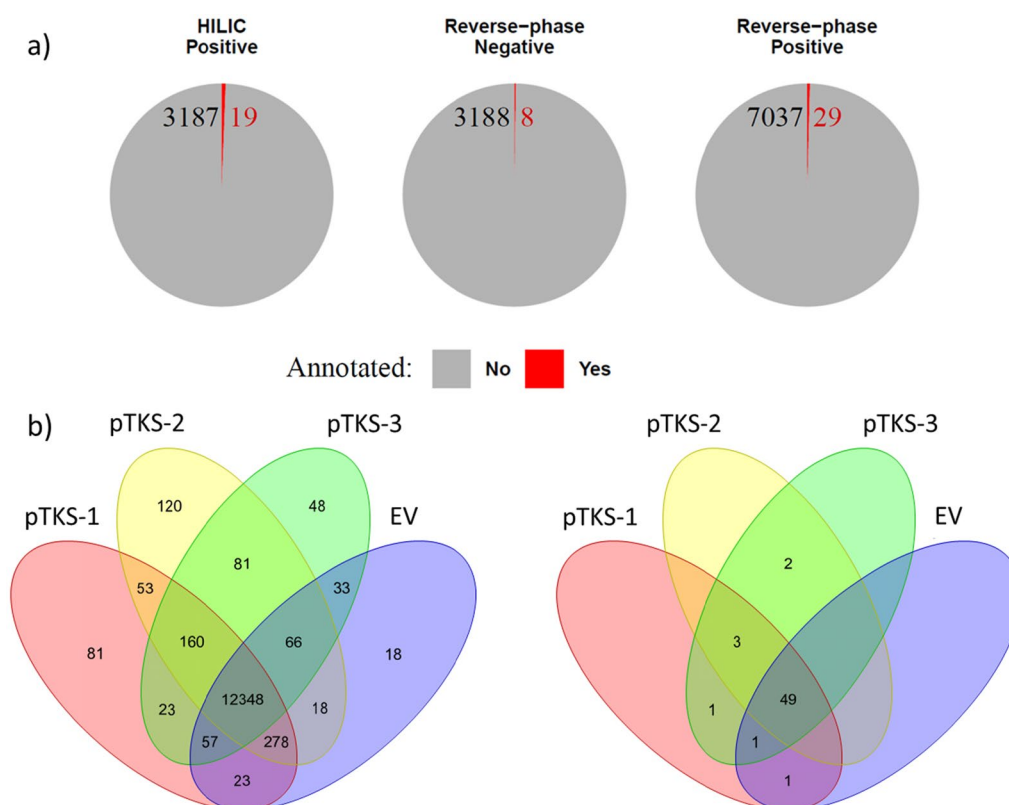


**Fig. 3** Detection of TKS and OAC enzymes in *P. tricornutum* transconjugants by western blot analysis. **a** Western blot for CsTKS detection in *P. tricornutum* pTKS transconjugant extracts using anti-6His antibodies. **b** Western blot for CsOAC detection in *P. tricornutum* pOAC transconjugant extracts using anti-Myc antibodies. In both panels, protein extracts from transconjugants are compared to those from the empty vector (EV) and wild-type *P. tricornutum*. Protein ladder sizes are indicated on the left while the expected sizes of the proteins of interest and their respective tags are shown on the right of each blot. The lower panels display red Ponceau (RP) staining to confirm protein loading

Among these, the most deregulated analytes were detected at  $RT=2.921$  min;  $m/z=577.33159$  (fold change between  $-5.1$  in pTKS-3 and  $-6.4$  in pTKS-1) and  $RT=2.842$  min;  $m/z=581.31317$  (fold change between  $3.4$  in pTKS-1 and  $4.9$  in pTKS-2). Among the 48 deregulated analytes, 26 were annotated (Fig. 6). Two of these annotations corresponded to 1-methyladenosine and two to thiamine. Among the 24 unique annotations, 13 were aromatic metabolites, including pinacidil, guanine, and worenine, the latter being downregulated in pTKS-2 transconjugants compared to the negative control. A single annotated metabolite, norvaline betaine, was deregulated in all pTKS transconjugants. Additionally, nine analytes were downregulated in all triplicates of at least one pTKS transconjugant. While analytes uniquely detected in the presence of CsTKS (Fig. 5b) were not consistently deregulated across all pTKS transconjugants due to variability in peak intensities among replicates, some showed significant upregulation. Among these, 3-hydroxy-C6-homoserine lactone, decyl dimethyl amine oxide, and proline betaine were the most upregulated analytes in the entire experiment.

#### Annotated metabolites deregulated by CsTKS in *Phaeodactylum tricornutum*

To investigate the impact of CsTKS on metabolism, we analyzed the structure and biosynthetic pathways of the nine annotated metabolites deregulated in at least two transconjugant strains (Table 1). Similar metabolites or pathways were highlighted in the table to identify shared metabolic routes. Of the nine detected and annotated analytes, eight possess aromatic structures, including two dipeptides (norvaline betaine and alanylphenylalanine), two alkaloids of quinolizine and quinolizidine nature (quinoline and thermopsine), and thiamine, which was detected as both phosphorylated and not phosphorylated. The two dipeptides originate from distinct biosynthetic pathways. While the pathway for norvaline betaine biosynthesis in *P. tricornutum* remains unresolved, norvaline is a modified valine expected to derive from pyruvate metabolism via either the (2S)-2-isopropyl malate or (S)-2-acetolactate branches, both of which converge at 2-oxoisovalerate, a precursor to valine [44]. The betaine moiety present in norvaline betaine is likely biosynthesized through the oxidation of choline to trimethyl-glycine [45, 46]. Coumarin and alanylphenylalanine



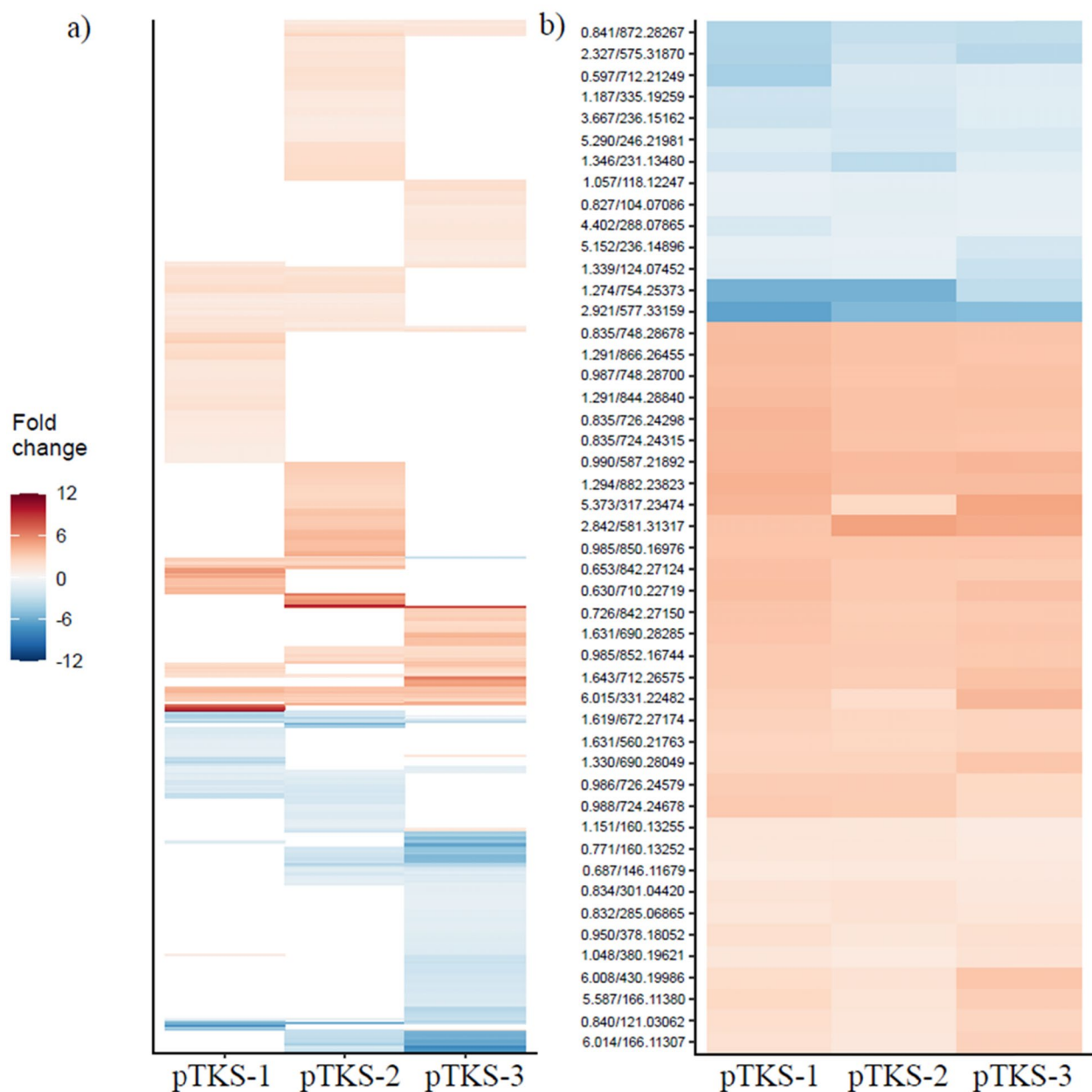
**Fig. 4** Untargeted metabolomic analysis identifies specific metabolites unique to *P. tricornutum* pTKS transconjugant lines. **a** Proportion of analytes, annotated or not, detected using three different HPLC–MS/MS analytical methods: reversed-phase chromatography with positive ionization (RP ESI<sup>+</sup>), reversed-phase chromatography with negative ionization (RP ESI<sup>-</sup>), and hydrophilic interaction chromatography with positive ionization (HILIC ESI<sup>+</sup>). **b** Total analytes detected (left) and annotated (right) for each strain, combining data from all three analytical modes. Among the annotated analytes absent in EV negative control, the following were consistently detected in pTKS transconjugant strains: pyocyanin, chlorophene and abietic acid were found in all pTKS lines; decyl dimethyl amine oxide and tri-isobutylphosphate were detected in pTKS-2 and pTKS-3; and 3-hydroxy-C6-homoserine lactone was identified in pTKS-1 and pTKS-3

are both derived from the shikimate pathway, though the specific biosynthesis of methyl 2-[(4-methyl-2-oxo-2H-chromen-7-yl)oxy] propanoate (a coumarin derivative) remains uncharacterized. 2-Hydroxyacetophenone, alanylphenylalanine, and the coumarin derivative share a benzene ring, with the latter also containing a pyran structure. Meanwhile, guanine and thiamine are derived from the nucleotide metabolism, with guanine from the purine metabolism, whereas the pyrimidine and thiazole moieties of thiamine are biosynthesized separately prior to be combined.

## Discussion

In this study, we successfully produced the *C. sativa* tetraketide synthase and olivetolic cyclase in *P. tricornutum*. Importantly, the heterologous expression of these enzymes did not impair the growth of the transconjugants (Additional file 1: Fig. S2). Due to the unavailability of 3,5,7-trioxododecanoyl-CoA, both as a substrate and as a standard, coupled with its high molecular weight

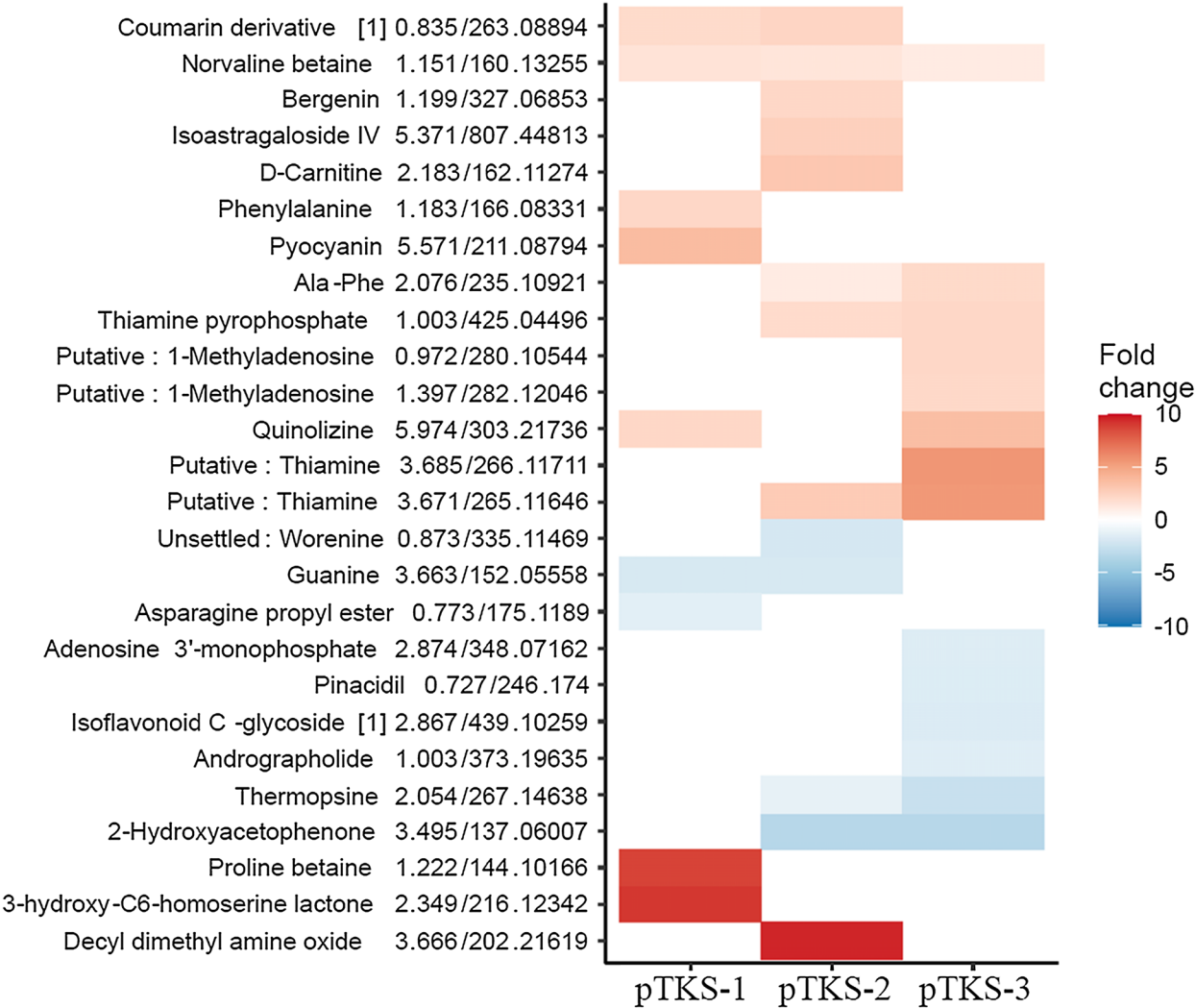
of 991.22 g mol<sup>-1</sup>, we were unable to directly test the enzymatic activities of CsTKS and CsOAC. Alternative strategies were therefore employed to evaluate their functionality. Analysis by HPLC–MS/MS (High pressure liquid chromatography with tandem mass spectrometry) of protein extracts containing CsTKS and CsOAC, either assessed alone or combined, did not yield a peak corresponding to olivetol or olivetolic acid standards. Several hypotheses could explain the absence of olivetolic acid in the extracts. Western blot analyses confirmed that both CsTKS and CsOAC were accumulating as soluble proteins in *P. tricornutum*, suggesting that protein misfolding is unlikely to account for the lack of product detection [57]. Instead, the failure to detect olivetol or olivetolic acid may be attributed to the utilization of key substrates or intermediates by competing endogenous pathways. Another plausible explanation is that CsTKS did not dimerize correctly in *P. tricornutum* cellular environment, as proper dimerization is critical for its activity [58].



**Fig. 5** Differential metabolite abundance highlights consistent deregulation in *P. tricornutum* pTKS transconjugants compared to the empty vector. **a** Overview of differentially abundant analytes detected using three analytical methods (Reversed-phased liquid chromatography coupled with ESI+/ESI- MS/MS and HILIC coupled with ESI+ MS/MS) compared to the empty vector. The y-axis lists all detected metabolites, showcasing the variability in abundance across pTKS transconjugant replicates. **b** Focused analysis of 48 analytes consistently deregulated in all pTKS transconjugants compared to the empty vector. Only analytes with a (|fold change|) greater than 1 and a significant threshold of  $p < 0.01$  were included. These analytes exhibit consistent under- or overabundance across all pTKS transconjugant replicates

To investigate the broader impact of CsTKS expression on the diatom metabolome, we performed an untargeted metabolomic analysis. This revealed significant differences in the metabolite profiles of pTKS transconjugants, with each replicate containing 50 to 120 exclusive metabolites that were absent from the control EV. A

previous research study has suggested that clone consistency is rare in *P. tricornutum* [59] supporting the variability in metabolite content and abundance observed in this study. Furthermore, 160 analytes were consistently detected in all three CsTKS-expressing replicates but were absent from the control EV. While only a fraction of



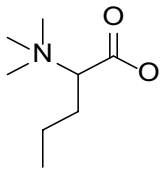
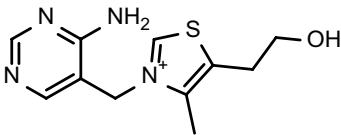
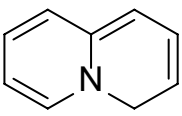
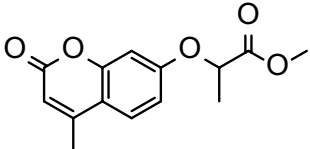
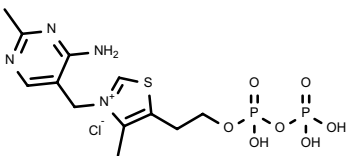
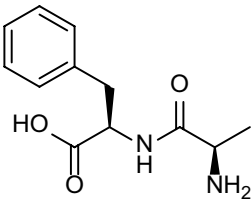
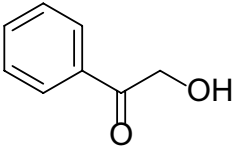
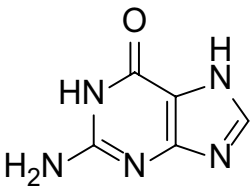
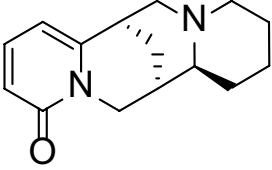
**Fig. 6** Presence of CsTKS coincides with altered abundance of cyclic metabolites in *P. tricornutum* transconjugants. Heatmap illustration the deregulation of annotated metabolites in each pTKS transconjugant replicates compared to the negative control. Only analytes with [fold change] > 1 and *p* < 0.01 were considered as deregulated. “Coumarin derivative [1]” corresponds to methyl 2-[(4-methyl-2-oxo-2H-chromen-7-yl)oxy] propanoate and “Isoflavonoid C-glycoside [1]” to Puerarin. Ala-Phe refers to alanylphenylalanine

these metabolites could be annotated, it is plausible that many of the unidentified analytes include polyketide-derived metabolites influenced CsTKS activity. Previous untargeted metabolomic studies in *P. tricornutum* using Fourier-transform ion cyclotron-resonance mass spectrometer resulted in the detection of several hundred of annotated analytes [60]. However, the lower resolution of the UPLC-qTOF-MS used in our study limited the number of annotated metabolites. Despite these limitations, clear patterns emerged in the metabolite abundance data. Specifically, 48 metabolites consistently deregulated across all *CsTKS*-expressing transconjugant replicates,

indicating a potential influence of CsTKS on intracellular metabolite content. These results suggest that CsTKS expression induces metabolic shifts in *P. tricornutum*, possibly altering the biosynthesis or accumulation of metabolites.

The structure of each annotated metabolites was analyzed to assess the potential involvement of TKS in their biosynthesis (Table 1). Polyketides are a diverse class of specialized metabolites characterized by the presence of C=O–CH<sub>2</sub> units or derived from such molecules, which complicates their structural identification. To address this challenge, we relied on structural information

**Table 1** Impact of heterologous *CsTKS* on *P. tricornutum*'s metabolism

Analyte	Chemical 2D structure	Family	Biosynthesis/metabolic route
Norvaline Betaine		Amino Acid dipeptide	Pyruvate Isoleucine pathway and choline oxidation reactions [44, 46–48]
Thiamine*		Vitamin/Cofactor	Vitamin B6, glycine, pyruvate and purine metabolism [44, 49]
Quinolizine		Alkaloid/Quinolizine	Unsolved, possible pyridine derived precursor, Malonyl-CoA and a Polyketide synthase [50]
Methyl 2-[(4-methyl-2-oxo-2H-chromen-7-yl)oxy]propanoate		Coumarin Derivative	Shikimate Pathway [44, 51]
Thiamine Pyrophosphate		Vitamin/Cofactor	Pyruvate and purine metabolism [44, 49]
Alanylphenylalanine		Amino Acid dipeptide	Shikimate Pathway [44, 52]
2-Hydroxyacetophenone		Organic aromatic alkyl-phenylketones/Benzoic Acid	Putative Shikimate and polyketide Pathway [53]
Guanine		Nucleic Acid	Nucleotide metabolism [44, 54, 55]
Thermopsine		Alkaloid/Quinolizidine	Lysine decarboxylation [50, 56]



**Table 1** (continued)

The presence of CsTKS alters the activity of key metabolic pathways, including the shikimate pathway, nucleotide metabolism, and amino acid metabolism. Metabolites highlighted in pink-orange are all upregulated, while those in blue are downregulated. Analytes are listed in descending order of fold change. The "Biosynthesis/metabolic route" column emphasizes shared pathways or biosynthetic routes among the identified analytes

corroborated by literature evidence. Among the deregulated metabolites, eight out of nine that were altered in two or more replicates and nine out of 16 altered in only one replicate were aromatic (Table 1 and Additional file 1: Table S3). In *C. sativa*, OAC catalyzes the formation of a phenolic structure from trioxododecanoyl-CoA. In *P. tricornutum*, it is plausible that a different enzyme takes over the biosynthesis of unidentified compounds from a polyketide-derived products generated by the CsTKS enzyme.

Pyocyanin, one of the three exclusive annotated metabolites, was detected in all pTKS transconjugant replicates and was significantly upregulated in pTKS-1 transconjugants (Figs. 4, 6). Pyocyanin is a well-characterized virulence factor in *Pseudomonas aeruginosa* [61, 62] and is classified as a phenazine pigment produced from chorismic acid [63]. Given that culture conditions were carefully controlled and remained stable over time and across transconjugants, the observed upregulation of pyocyanin is likely attributable to a direct or indirect modification in its biosynthetic pathway as PKS are capable of synthesizing pigments, rather than being driven by environmental factors, such as changes in light conditions [64–66].

As previously mentioned, most of the detected metabolites exhibit a phenolic structure. Among them, a coumarin derivative, methyl 2-[(4-methyl-2-oxo-2Hchromen-7-yl)oxy]propanoate, was upregulated in two pTKS transconjugant lines. While its precise biosynthetic pathway in vivo remains unclear, it likely originates from coumarin or coumaroyl-CoA, intermediates produced via the phenylpropanoid pathway. The formation of coumarin core is thought to involve condensation reactions between phenolic and carbonyl metabolites, which may also contribute to the biosynthesis of other aromatic metabolites detected in this study [67]. The potential involvement of a polyketide synthase such as CsTKS in its synthesis remains uncertain. It has been reported that 4-coumaric acid can serve as substrate for type three polyketide synthase (PKSIII), which can catalyze the production of stilbene or chalcone metabolites [68]. Although CsTKS is a PKSIII, it is not expected to produce chalcone products, and no stilbene metabolites were detected among the 16 upregulated annotated metabolite. This is unexpected, as CsTKS is hypothesized to generate stilbene

through linear polyketides synthesis [58]. Additionally, 3-hydroxy-C6-homoserine lactone, which is known to play a role in quorum sensing, was identified. While its structure resembles that of polyketides, its classification as a polyketide has not been confirmed [69, 70]. Another, metabolites, alanylphenylalanine, is a modified amino acid that could serve as a substrate for polyketide extension [71]. It is plausible that the activity of CsTKS stimulates the synthesis of alanylphenylalanine. Overall, we observed a general increase in metabolites derived from the shikimate pathway, including berberine, pyocyanin, phenylalanine and coumarin. This suggest that CsTKS may either influence the shikimate pathway or its products may affect light perception, thereby increasing the synthesis of metabolites capable of light absorption [64]. It has not been reported in literature that heterologous production of enzymes in *P. tricornutum* has by itself an impact on the activity of the shikimate pathway, although few of those study perform metabolomic analysis.

The only analyte consistently upregulated across all transconjugants in this study is norvaline betaine. The biological role and biosynthetic pathway of this modified amino acid remains largely unknown. Additionally, a significant upregulation of proline betaine was observed in TKS-1 transconjugant strains (Fig. 6), suggesting that the presence of CsTKS might enhance betaine biosynthesis. Betaine, in the form of trimethyl-glycine, along with proline betaine and carnitine (upregulated in pTKS-1 and pTKS-2 lines, respectively, as shown in Additional file 1: Table S3), are well-known quaternary ammonium compounds. These metabolites function as osmolytes and are typically upregulated in saline stress [72]. Beyond their osmoprotective properties, betaines are also involved in epigenetic regulation and can act as methyl donors, although the in vivo mechanism of amino acid betainization is still unclear [46, 73]. The observed upregulation of glycine could be associated with increased thiamine production, though the underlying causal relationship remains speculative.

Quinolizine, thermopsine, and worenine, which belong to the quinolizine family, were also identified. The biosynthesis of quinolizine in vivo is not well-characterized, though a synthetic pathway has been successfully achieved using a PKSIII enzyme from *Huperzia serrata* to

produce 2-hydroxy-4H-quinolizin-4-one [50]. This raises the possibility that CsTKS may contribute to quinolizine biosynthesis in *P. tricornutum*. However, thermopsine was downregulated in two pTKS transconjugant strains and worenine, an isoquinoline, was downregulated in one strain (Additional file 1: Table S3). This suggests that while CsTKS might be involved in the production of quinolizine-related metabolites, its precise role and the regulatory mechanism influencing these metabolites remains unclear.

Hydroxyacetophenones, such as karanjin, are suspected to be produced via the polyketide pathway [19]. If the biosynthesis of 2-hydroxyacetophenone follows a similar process in *P. tricornutum*, its reduced concentration in this study could be due to substrate channeling by the heterologous CsTKS, which may utilize precursors typically employed by native polyketide synthases.

## Conclusions

This study highlights the inherent metabolic variability among *P. tricornutum* transconjugants, reinforcing the challenge of achieving consistent metabolite production in engineered strains. Significant differences were observed among transconjugant clones, with distinct metabolic profiles and varying analyte compositions. The presence of metabolites exclusive to CsTKS-expressing transconjugants, supports the metabolic activity of the heterologous enzyme, yet no annotated analytes were unique to a single clone. Among the metabolites detected exclusively in pTKS transconjugants, compounds such as pyocyanin, abietic acid, and chlorophene were identified, though they are unlikely to originate from the polyketide pathway. These findings emphasize the complexity of *P. tricornutum* metabolism and highlight the need for a deeper understanding of its metabolic network to enable controlled and sustainable olivetolic acid production.

## Abbreviations

THC	$\Delta^9$ -Tetrahydrocannabinol
CBD	Cannabidiol
THCAS	Tetrahydrocannabinolic acid synthase
THCA	$\Delta^9$ -Tetrahydrocannabinolic acid
CBDAS	Cannabidiolic acid synthase
CBDA	Cannabidiolic acid
TKS	Tetraketide synthase
OAC	Olivetolic acid cyclase
OA	Olivetolic acid
GPP	Geranyl pyrophosphate
APT	Aromatic prenyl transferase
HILIC	Hydrophilic interaction liquid chromatography
RP	Reversed-phase
MS	Mass spectrometry
PCR	Polymerase chain reaction
EV	Empty vector

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13068-025-02638-1>.

Additional file 1. Fig. S1. *P. tricornutum* DNA plasmid digestion profiles. Agarose gel electrophoresis profile *Bam*HI-HF digestion recombinant plasmid DNA extracted from *E. coli* after plasmid rescue from *P. tricornutum* transconjugants. Fig. S2. CsTKS and empty vector clones have a similar growth kinetic. Optic density of each culture was taken at 730 nm and a t-test was performed for statistical analysis. Fig. S3. HPLC chromatogram of CsTKS and CsOAC enzymatic assays. The black line corresponds to the run with cannabinoid standards. a: Chromatogram of enzymatic assay from crude extract of AC9.1, AC9.4, AC10.5 and the combined culture made of all three clones mixed, all supplemented with hexanoyl-CoA and malonyl-CoA. Each clone has previously been tested for protein production by western blot. The two peaks visible between 11 mins and 12 mins do not correspond to any cannabinoid standard's UV profile. b: Chromatogram of pTKS.1, pTKS.4, pOAC.5 and EV enzymatic assays. All clones follow the same profile. c: Chromatogram of pPtGE30 with and without substrate. d: Chromatogram of Combined clones with and without substrate. Fig. S4. AC9.3 has a more distant metabolite profile than the 3 other clones: Principal component analysis of every clone's metabolite profile by combining HILIC, Reverse phase positive and negative mode. Fig. S5. All triplicates of each clone follow a similar profile. Heatmap of unannotated metabolite relative abundance in each clone compared to the mean in all samples. Table S1. All DNA sequences used in this study. Table S2. All primers sequences used in this study. Table S3. Annotated analytes deregulated in transconjugants. The presence of the heterologous CsTKS in *Phaeodactylum tricornutum*'s metabolism displays variations in the activity of the shikimate pathway, the nucleotide metabolism, and amino acids metabolism. Compounds in pink-orange are all upregulated while compounds in blue are deregulated. They are displayed in order of fold change from top to bottom. Biosynthesis/metabolic route section focuses on the common compounds or pathways between all analytes' biosynthesis possible routes.

## Acknowledgements

The authors would like to thank Melodie B. Plourde for her professional assistance and support and all lab members for their teamwork. *Phaeodactylum tricornutum* strain used in this study was kindly provided by Professor Bogumil Karas (Western University, London, ON, Canada), warm thanks are extended. During the preparation of this work, the authors used ChatGPT 4.0 in order to correct grammatical errors and enhance readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Author contributions

NS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing—original draft, Writing—review and editing. KCGS: Data curation, Formal analysis, Writing—review and editing. NM: Writing—review and editing. SEG: Formal analysis, Writing—review and editing. AC: Formal analysis, Validation. FA: Conceptualization, Writing—review and editing. EF: Formal analysis, Investigation, Methodology. FM: Formal analysis, Investigation, Methodology, Writing—review and editing. HG: Funding acquisition, Resources, Supervision, Writing—review and editing. IDP: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing—review and editing.

## Funding

Canada Research Chair on plant specialized metabolism Award No. CRC-2018-00137 to I.D-P. Thanks are extended to the Canadian taxpayers and to the Canadian government for supporting the Canada Research Chairs Program. Additional support in the form of scholarship to K.C.G.d.S by Mitacs-Elevate program grant #IT28769 to I.D-P.

**Data availability**

Data generated or analyzed during this study are provided in full within the published article and its supplementary materials and from the corresponding author upon reasonable request.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare no competing interests.

**Author details**

<sup>1</sup>Department of Biochemistry, Chemistry, Physics, and Forensic Science, Université du Québec à Trois-Rivières, 3351 Boulevard des Forges, Trois-Rivières, QC G9A 5H7, Canada. <sup>2</sup>Groupe de Recherche en Biologie Végétale, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada.

Received: 1 February 2025 Accepted: 14 March 2025

Published online: 04 April 2025

**References**

- Sledziński P, Nowak-Terpiłowska A, Zeyland J. Cannabinoids in medicine: cancer, immunity, and microbial diseases. *Int J Mol Sci*. 2020;22(1):263.
- Zou S, Kumar U. Cannabinoid receptors and the endocannabinoid system: signaling and function in the central nervous system. *Int J Mol Sci*. 2018;19(3):833.
- Crocq M-A. History of cannabis and the endocannabinoid system. *Dialogues Clin Neurosci*. 2020;22(3):223–8.
- Punja ZK, Sutton DB, Kim T. Glandular trichome development, morphology, and maturation are influenced by plant age and genotype in high THC-containing cannabis (*Cannabis sativa* L.) inflorescences. *J Cannabis Res*. 2023;5(1):12.
- Conneely LJ, Mauleon R, Mieog J, Barkla BJ, Kretschmar T. Characterization of the *Cannabis sativa* glandular trichome proteome. *PLoS ONE*. 2021;16(4):e0242633.
- Micale V, Mazzola C, Drago F. Endocannabinoids and neurodegenerative diseases. *Pharmacol Res*. 2007;56(5):382–92.
- Farag S, Kayser O. The cannabis plant: botanical aspects. In: *Handbook of cannabis and related pathologies*. Elsevier; 2017. pp. 3–12.
- Gagne SJ, Stout JM, Liu E, Boubakir Z, Clark SM, Page JE. Identification of olivetolic acid cyclase from *Cannabis sativa* reveals a unique catalytic route to plant polyketides. *Proc Natl Acad Sci*. 2012;109(31):12811–6.
- Luo X, Reiter MA, d'Espaux L, Wong J, Denby CM, Lechner A, Zhang Y, Grzybowski AT, Harth S, Lin W. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature*. 2019;567(7746):123–6.
- Tahir MN, Shahbazi F, Rondeau-Gagné S, Trant JF. The biosynthesis of the cannabinoids. *J Cannabis Res*. 2021;3:1–12.
- Xie Z, Mi Y, Kong L, Gao M, Chen S, Chen W, Meng X, Sun W, Chen S, Xu Z. *Cannabis sativa*: origin and history, glandular trichome development, and cannabinoid biosynthesis. *Horticult Res*. 2023;10(9):uhad150.
- Hammond D, Goodman S, Wadsworth E, Freeman TP, Kilmer B, Schauer G, Pacula RL, Hall W. Trends in the use of cannabis products in Canada and the USA, 2018–2020: findings from the International Cannabis Policy Study. *Int J Drug Policy*. 2022;105: 103716.
- Cannabis Market Size, Share & COVID-19 Impact Analysis, By Type (Flowers/Buds and Concentrates), By Application (Medical, Recreational (Edibles and Topicals), and Industrial Hemp) By Component (THC-Dominant, Balanced THC & CBD, and CBD Dominant), and Regional Forecast, 2023–2030. <https://www.fortunebusinessinsights.com/industry-reports/cannabis-marijuana-market-100219>
- Lewis MM, Yang Y, Wasilewski E, Clarke HA, Kotra LP. Chemical profiling of medical cannabis extracts. *ACS Omega*. 2017;2(9):6091–103.
- Backer R, Schwinghamer T, Rosenbaum P, McCarty V, Eichhorn Bilodeau S, Lyu D, Ahmed MB, Robinson G, Lefsrud M, Wilkins O. Closing the yield gap for cannabis: a meta-analysis of factors determining cannabis yield. *Front Plant Sci*. 2019;10:495.
- Wiles D, Shanbhag BK, O'Brien M, Doblin MS, Bacic A, Beddoe T. Heterologous production of *Cannabis sativa*-derived specialised metabolites of medicinal significance—insights into engineering strategies. *Phytochemistry*. 2022;203: 113380.
- Zirpel B, Degenhardt F, Martin C, Kayser O, Stehle F. Engineering yeasts as platform organisms for cannabinoid biosynthesis. *J Biotechnol*. 2017;259:204–12.
- Kumar G, Shekh A, Jakhu S, Sharma Y, Kapoor R, Sharma TR. Bioengineering of microalgae: recent advances, perspectives, and regulatory challenges for industrial application. *Front Bioeng Biotechnol*. 2020;8:914.
- Gao J, Jiang L, Lian J. Development of synthetic biology tools to engineer *Pichia pastoris* as a chassis for the production of natural products. *Synth Syst Biotechnol*. 2021;6(2):110–9.
- Nouemssi SB, Ghribi M, Beauchemin R, Meddeb-Mouelhi F, Germain H, Desgagné-Penix I. Rapid and efficient colony-PCR for high throughput screening of genetically transformed *Chlamydomonas reinhardtii*. *Life*. 2020;10(9):186.
- Laban A. Cannabinoid production in algae. Google Patents; 2021.
- Kassaw TK, Paton AJ, Peers G. Episome-based gene expression modulation platform in the model diatom *Phaeodactylum tricornutum*. *ACS Synth Biol*. 2022;11(1):191–204.
- George J, Kahlke T, Abbriano RM, Kuzhiumparambil U, Ralph PJ, 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.
- Pudney A, Gandini C, Economou CK, Smith R, Goddard P, Napier JA, Spicer A, Sayanova O. Multifunctionalizing the marine diatom *Phaeodactylum tricornutum* for sustainable co-production of omega-3 long chain polyunsaturated fatty acids and recombinant phytase. *Sci Rep*. 2019;9(1):11444.
- Xue J, Niu Y-F, Huang T, Yang W-D, Liu J-S, Li H-Y. Genetic improvement of the microalga *Phaeodactylum tricornutum* for boosting neutral lipid accumulation. *Metab Eng*. 2015;27:1–9.
- Lazarjani MP, Young O, Kebede L, Seyfoddin A. Processing and extraction methods of medicinal cannabis: a narrative review. *Journal Cannabis Res*. 2021;3:1–15.
- Fajardo AR, Cerdán LE, Medina AR, Fernández FGA, Moreno PAG, Grima EM. Lipid extraction from the microalga *Phaeodactylum tricornutum*. *Eur J Lipid Sci Technol*. 2007;109(2):120–6.
- Cui Y, Thomas-Hall SR, Schenk PM. *Phaeodactylum tricornutum* microalgae as a rich source of omega-3 oil: progress in lipid induction techniques towards industry adoption. *Food Chem*. 2019;297: 124937.
- Fabris M, George J, Kuzhiumparambil U, Lawson CA, Jaramillo-Madrid AC, Abbriano RM, Vickers CE, Ralph P. Extrachromosomal genetic engineering of the marine diatom *Phaeodactylum tricornutum* enables the heterologous production of monoterpenoids. *ACS Synth Biol*. 2020;9(3):598–612.
- Hexanoyl-CoA. <https://biocyc.org/compound?orgid=PHAE&id=HEXAN OYL-COA#RXNS>
- Fantino E, Awwad F, Merindol N, Garza AMD, Gélinais S-E, Robles GCG, Custeau A, Meddeb-Mouelhi F, Desgagné-Penix I. Bioengineering *Phaeodactylum tricornutum*, a marine diatom, for cannabinoid biosynthesis. *Algal Res*. 2024;77: 103379.
- Fantino E, Messaabi A, Merindol N, Awwad F, Sene N, Gélinais S-E, Custeau A, Rhéaume K-L, Meddeb-Mouelhi F, Desgagné-Penix I. Extrachromosomal expression of functional *Cannabis sativa* cannabidiolic acid synthase in *Phaeodactylum tricornutum*. *Algal Res*. 2024;103889.
- Awwad F, Fantino EI, Héneault M, Diaz-Garza AM, Merindol N, Custeau A, Gélinais S-E, Meddeb-Mouelhi F, Li J, Lemay J-F. 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(23):16624.
- Karas BJ, Molparia B, Jablanovic J, Hermann WJ, Lin Y-C, Dupont CL, Tagwerker C, Yonemoto IT, Noskov VN, Chuang R-Y, et al. Assembly of eukaryotic algal chromosomes in yeast. *J Biol Eng*. 2013;7(1):30.
- Diamond A, Diaz-Garza AM, Li J, Slattery SS, Merindol N, Fantino E, Meddeb-Mouelhi F, Karas BJ, Barnabé S, Desgagné-Penix I. Instability of extrachromosomal DNA transformed into the diatom *Phaeodactylum tricornutum*. *Algal Res*. 2023;70: 102998.

36. Karas BJ, Diner RE, Lefebvre SC, McQuaid J, Phillips APR, Noddings CM, Brunson JK, Valas RE, Deerinck TJ, Jablanovic J, et al. Designer diatom episomes delivered by bacterial conjugation. *Nat Commun.* 2015;6(1):6925.
37. Elso CM, Roberts LJ, Smyth KJ, Thomson RJ, Baldwin TM, Foote SJ, Handman E. Leishmaniasis host response loci (Imr1-3) modify disease severity through a Th1/Th2-independent pathway. *Genes Immun.* 2004;5(2):93–100.
38. Lai Z, Tsugawa H, Wohlgemuth G, Mehta S, Mueller M, Zheng Y, Ogiwara A, Meissen J, Showalter M, Takeuchi K, et al. Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics. *Nat Methods.* 2018;15(1):53–6.
39. Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, Kanazawa M, VanderGheynst J, Fiehn O, Arita M. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods.* 2015;12(6):523–6.
40. Tsugawa H, Nakabayashi R, Mori T, Yamada Y, Takahashi M, Rai A, Sugiyama R, Yamamoto H, Nakaya T, Yamazaki M, et al. A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms. *Nat Methods.* 2019;16(4):295–8.
41. Databank. <https://mona.fiehnlab.ucdavis.edu/>
42. Pang Z, Xu L, Viau C, Lu Y, Salavati R, Basu N, Xia J: MetaboAnalystR 4.0: a unified LC-MS workflow for global metabolomics. *Nat Commun.* 2024;15(1).
43. Schwender H. Siggenes: multiple testing using SAM and Efron's Empirical Bayes approaches. 2023.
44. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 2023;51(D1):D587–92.
45. Waditee R, Tanaka Y, Aoki K, Hibino T, Jikuya H, Takano J, Takabe T, Takabe T. Isolation and functional characterization of *N*-methyltransferases that catalyze betaine synthesis from glycine in a halotolerant photosynthetic organism *Aphanethece halophytica*. *J Biol Chem.* 2003;278(7):4932–42.
46. Nyyssölä A, Kerovuo J, Kaukinen P, von Weymar N, Reinikainen T. Extreme halophiles synthesize betaine from glycine by methylation. *J Biol Chem.* 2000;275(29):22196–201.
47. Soini J, Falschlehner C, Liedert C, Bernhardt J, Vuoristo J, Neubauer P. Norvaline is accumulated after a down-shift of oxygen in *Escherichia coli* W3110. *Microb Cell Fact.* 2008;7:1–14.
48. Popko J, Herrfurth C, Feussner K, Ischebeck T, Iven T, Haslam R, Hamilton M, Sayanova O, Napier J, Khozin-Goldberg I. Metabolome analysis reveals betaine lipids as major source for triglyceride formation, and the accumulation of sedoheptulose during nitrogen-starvation of *Phaeodactylum tricornutum*. *PLoS ONE.* 2016;11(10): e0164673.
49. Llaverio Pasquina M. Molecular biology of B vitamin metabolism genes and their regulation in *Phaeodactylum tricornutum*. 2020.
50. Wang J, Ding N, Wu Y, Shi X, Qi B, Liu X, Wang X, Li J, Tu P, Shi S. Enzymatic synthesis of 2-hydroxy-4H-quinolizin-4-one scaffolds by integrating coenzyme A ligases and a type III PKS from *Huperzia serrata*. *RSC Adv.* 2020;10(40):23566–72.
51. Del Mondo A, Sansone C, Brunet C. Insights into the biosynthesis pathway of phenolic compounds in microalgae. *Comput Struct Biotechnol J.* 2022;20:1901–13.
52. Silva C, da Silva FJ, Pinheiro G, Teixeira A, Freire P. Vibrational and structural properties of L-Alanyl-L-phenylalanine dipeptide by Raman spectroscopy, infrared and DFT calculations. *Vib Spectrosc.* 2018;98:128–33.
53. Singh A, Bhatt G, Gujre N, Mitra S, Swaminathan R, Limaye A, Rangan L, Karanjin. *Phytochemistry.* 2021;183: 112641.
54. Witte C-P, Herde M. Nucleotide metabolism in plants. *Plant Physiol.* 2020;182(1):63–78.
55. Mojzeš P, Gao L, Ismagulova T, Pilátová J, Moudříková Š, Gorelova O, Solovchenko A, Nedbal L, Salih A. Guanine, a high-capacity and rapid-turnover nitrogen reserve in microalgal cells. *Proc Natl Acad Sci.* 2020;117(51):32722–30.
56. Zhang P, An Q, Yi P, Cui Y, Zou J-B, Yuan C-M, Zhang Y, Gu W, Huang L-J, Zhao L-H. Thermalseedlines A–G, seven thermopsine-based alkaloids with antiviral and insecticidal activities from the seeds of *Thermopsis lanceolata* R. *Br Fitoterapia.* 2022;158: 105140.
57. Reynaud E. Protein misfolding and degenerative diseases. *Nat Educ.* 2010;3(9):28.
58. Kearsey LJ, Prandi N, Karupiah V, Yan C, Leys D, Toogood H, Takano E, Scrutton NS. Structure of the *Cannabis sativa* olivetol-producing enzyme reveals cyclization plasticity in type III polyketide synthases. *FEBS J.* 2020;287(8):1511–24.
59. Diaz-Garza AM, Merindol N, Dos Santos KCG, Lavoie-Marchand F, Ingalls B, Desgagne-Penix I. No two clones are alike: characterization of heterologous subpopulations in a transgenic cell line of the model diatom *Phaeodactylum tricornutum*. *Microb Cell Fact.* 2024;23(1):286.
60. Duarte B, Feijão E, Cruz de Carvalho R, Duarte IA, Marques AP, Maia M, Hertzog J, Matos AR, Cabrita MT, Caçador I. Untargeted metabolomics reveals antidepressant effects in a marine photosynthetic organism: the diatom *Phaeodactylum tricornutum* as a case study. *Biology.* 2022;11(12):1770.
61. Marey MA, Abozahr R, El-Nikhely NA, Kamal MF, Abdelhamid SM, El-Kholy MA. Transforming microbial pigment into therapeutic revelation: extraction and characterization of pyocyanin from *Pseudomonas aeruginosa* and its therapeutic potential as an antibacterial and anticancer agent. *Microb Cell Fact.* 2024;23(1):174.
62. Jabłońska J, Augustyniak A, Dubrowska K, Rakoczy R. The two faces of pyocyanin—why and how to steer its production? *World J Microbiol Biotechnol.* 2023;39(4):103.
63. Zhou H, Yang Y, Shang W, Rao Y, Chen J, Peng H, Huang J, Hu Z, Zhang R, Rao X. Pyocyanin biosynthesis protects *Pseudomonas aeruginosa* from nonthermal plasma inactivation. *Microb Biotechnol.* 2022;15(6):1910–21.
64. Heydarzadeh P, Boureba W, Zahedi M, Huang B, Moreau B, Lukomska E, Couzinet-Mossion A, Wielgosz-Collin G, Martin-Jézéquel V, Bougaran G. Response of CO<sub>2</sub>-starved diatom *Phaeodactylum tricornutum* to light intensity transition. *Philos Trans R Soc B Biol Sci.* 2017;372(1728):20160396.
65. Günther M, Reimer C, Herbst R, Kufs JE, Rautschek J, Ueberschaar N, Zhang S, Peschel G, Reimer L, Regestein L. Yellow polyketide pigment suppresses premature hatching in social amoeba. *Proc Natl Acad Sci.* 2022;119(43): e2116122119.
66. Lebeau J, Venkatachalam M, Fouillaud M, Petit T, Vinale F, Dufossé L, Caro Y. Production and new extraction method of polyketide red pigments produced by ascomycetous fungi from terrestrial and marine habitats. *J Fungi.* 2017;3(3):34.
67. Methyl 2-[(4-methyl-2-oxo-2H-chromen-7-yl)oxy]propanoate. <https://www.smolecule.com/products/s12619316#pos-top>
68. Watts KT, Lee PC, Schmidt-Dannert C. Biosynthesis of plant-specific stilbene polyketides in metabolically engineered *Escherichia coli*. *BMC Biotechnol.* 2006;6:1–12.
69. Weiland-Bräuer N, Kisch MJ, Pinnow N, Liese A, Schmitz RA. Highly effective inhibition of biofilm formation by the first metagenome-derived Al-2 quenching enzyme. *Front Microbiol.* 2016;7:1098.
70. Liu X, Bimerew M, Ma Y, Müller H, Ovadis M, Eberl L, Berg G, Chernin L. Quorum-sensing signaling is required for production of the antibiotic pyrrolnitrin in a rhizospheric biocontrol strain of *Serratia plymuthica*. *FEMS Microbiol Lett.* 2007;270(2):299–305.
71. Ray L, Moore BS. Recent advances in the biosynthesis of unusual polyketide synthase substrates. *Nat Prod Rep.* 2016;33(2):150–61.
72. Nikitashina V, Stettin D, Pohnert G. Metabolic adaptation of diatoms to hypersalinity. *Phytochemistry.* 2022;201: 113267.
73. Tuomainen M, Kärkkäinen O, Leppänen J, Auriola S, Lehtonen M, Savolainen MJ, Hermansen K, Riserus U, Åkesson B, Thorsdottir I. Quantitative assessment of betainized compounds and associations with dietary and metabolic biomarkers in the randomized study of the healthy Nordic diet (SYSDIET). *Am J Clin Nutr.* 2019;110(5):1108–18.

## Publisher's Note

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