Evolutionary and biochemical characterization of a *Chromochloris zofingiensis* MBOAT with wax synthase and DGAT activity

Yang Xu[a,1,2], Xue Pan[b,1], Junhao Lu[a], Juli Wang[a], Qiyuan Shan[a,c], Jake Stout[d], Guanqun Chen[a,*]

[a] Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Alberta, Canada T6G 2P5

[b] Center for Plant Cell Biology, Institute of Integrative Genome Biology, and Department of Botany and Plant Sciences, University of California, Riverside, CA, USA, 92521

[c] School of Pharmaceutical Science, Zhejiang Chinese Medical University, Hangzhou, China, 310053

[d] Department of Biological Sciences, University of Manitoba, Winnipeg, Manitoba, Canada, R3T 2N2

[1]These authors contributed equally to this work

[2] Present address: MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824, USA

*To whom correspondence should be addressed: Guanqun Chen [Phone: (+1) 780 492-3148; Fax: (+1) 780 492-4265; Email: gc24@ualberta.ca; ORCID: 0000-0001-5790-3903].

Running Title: Functional divergence of wax synthase in green lineage

Email addresses:

Yang Xu: yxu9@ualberta.ca

Xue Pan: xp2@ualberta.ca

Junhao Lu: junhao3@ualberta.ca

Juli Wang: juli1@ualberta.ca

Qiyuan Shan: qiyuan.shan@hotmail.com

Jake Stout: Jake.Stout@umanitoba.ca

Guanqun Chen: gc24@ualberta.ca

Total word count (excluding materials and methods): 4481

Number of figures: 6

Number of tables: 2

Number of support information: 1 (7 supplemental figures and 3 supplemental tables)

Highlight: Unlike jojoba WS responsible for wax biosynthesis, its homolog in *Chromochloris zofingiensis* has *in-vitro* WS and DGAT activities and affects triacylglycerol accumulation when expressed in *Chlamydomonas*, suggesting possible functional divergence.

Abstract:

Wax synthase (WS) catalyzes the last step in wax ester biosynthesis in green plants. Two unrelated subfamilies of WS including the bifunctional acyltransferase and plant-like WS have been reported, but the latter is largely uncharacterized in microalgae. Here, we functionally characterized a putative plant-like WS from the emerging model green microalga *Chromochloris zofingiensis* (CzWS1). Our results showed that plant-like WS evolved under different selection constraints in plants and microalgae, with positive selection likely contributing to functional divergence. Unlike jojoba with high amounts of wax ester in seeds and a highly active WS enzyme, *C. zofingiensis* has no detectable wax ester but high abundance of *WS* transcripts. Co-expression analysis showed that *C. zofingiensis WS* has different expression correlations with lipid biosynthetic genes from jojoba and may have a divergent function. *In vitro* characterization indicated CzWS1 had diacylglycerol acyltransferase activity along with WS activity and overexpression of *CzWS1* in yeast and *Chlamydomonas reinhardtii* affected triacylglycerol accumulation in yeast and microalgal cells. Moreover, biochemical and bioinformatic analyses revealed the relevance of the C-terminal region of CzWS1 in enzyme function. Taken together, our results indicated functional divergence of plant-like WS in plants and microalgae and the importance of C-terminal region in enzyme functional specialization.

Keywords: Acyl-CoA: fatty alcohol acyltransferase, wax synthase, DGAT, wax ester, triacylglycerol, algal lipid, enzyme kinetics, yeast, *Chromochloris zofingiensis*, *Chlamydomonas reinhardtii*

Abbreviations

CDS, Coding DNA sequences; DAG, *sn*-1,2-diacylglycerol; DGAT, Acyl-CoA:diacylglycerol acyltransferase; ER, Endoplasmic reticulum; FAE, Fatty acid elongase; FAR, Fatty Acid Reductase; MBOAT, Membrane-bound *O*-acyltransferase; $OD_{600/750}$, Optical density at 600 or 750 nm; PES, Phytyl ester synthase; S.D., Standard deviation; Sc-Ura, Synthetic complete medium lacking uracil; TAG, Triacylglycerol; TAP, Tris acetate phosphate medium; TAP-N, Nitrogen deficient TAP medium; TLC, Thin-layer chromatography; TMD, Transmembrane domain; WE, Wax ester; WS, Wax synthase (acyl-CoA:fatty alcohol acyltransferase); WSD, bifunctional WS/DGAT.

70 **Introduction**

71 Fatty acids are the most abundant forms of energy storage and reduced carbon in nature and have
72 diverse applications in food and oleochemical industries (Thelen and Ohlrogge, 2002; Xu *et al.*,
73 2018). The plant kingdom represents a vast renewable resource of fatty acids and many plants
74 and microalgae accumulate fatty acids in the form of triacylglycerol (TAG) and/or wax ester
75 (WE) as major storage neutral lipids. TAG consists of three fatty acids esterified to a glycerol
76 backbone and is stored in large amounts in plant seeds and many microalgal cells growing under
77 stress conditions (Xu *et al.*, 2018). Some plant and microalgal species, such as jojoba
78 (*Simmondsia chinensis*) and *Euglena gracilis*, can accumulate high concentrations of WE as
79 storage compounds, which is a hydrophobic ester composed of long-chain fatty acids and
80 primary long-chain fatty alcohols (Lardizabal *et al.*, 2000; Teerawanichpan and Qiu, 2010;
81 Sturtevant *et al.*, 2020). In addition to energy and carbon reservoirs, TAG and WE fill a number
82 of physiological functions in plant and microalgal architecture and development, lipid
83 homeostasis, protection and stress responses (Hu *et al.*, 2008; Rottig and Steinbuchel, 2013; Xu
84 *et al.*, 2018; Yang and Benning, 2018; Lewandowska *et al.*, 2020; Lu *et al.*, 2020).

85 In plants and microalgae, the last steps of acyl-CoA-dependent TAG and WE
86 biosynthetic pathways are functionally similar. In TAG biosynthesis, the transfer of an acyl chain
87 from acyl-CoA to an *sn*-1,2-diacylglycerol (DAG) is catalyzed by acyl-CoA:diacylglycerol
88 acyltransferase (DGAT), whereas WE biosynthesis ends with an analogous transfer of an acyl
89 chain from acyl-CoA to a fatty alcohol via the catalytic actions of acyl-CoA:fatty alcohol
90 acyltransferase (wax synthase, WS) (Lardizabal *et al.*, 2000; Xu *et al.*, 2018). At least four and
91 three different subfamilies of DGAT and WS have been identified, respectively, and these
92 subfamilies are closely related to or overlapped with each other. To date, there are three
93 membrane-bound and one soluble forms of DGAT present in plants and microalgae. DGAT1 and
94 DGAT2 are the two major membrane-bound DGATs that primarily contribute to TAG formation
95 in developing seeds and microalgae. DGAT1 belongs to the membrane-bound *O*-acyltransferase
96 (MBOAT) family and is featured by the presence of multiple transmembrane domains (TMD),
97 whereas DGAT2 is a member of the DGAT2/acyl-CoA:monoacylglycerol acyltransferase family
98 and contains fewer TMD (Xu *et al.*, 2018). Another membrane-bound form of DGAT is a
99 bifunctional acyltransferase (WSD) that exhibits both DGAT and WS activity (Liu *et al.*, 2012).
100 This enzyme was first identified in the Gram-negative bacterium *Acinetobacter calcoaceticus*

101 ADP1 and was found to contribute to both WE and TAG biosynthesis (Kalscheuer *et al.*, 2003).
102 Later, WSD homologs were also isolated and characterized in many plant and microalgal
103 species. WSD from *Arabidopsis thaliana* and *Petunia* appears to predominantly catalyze WE
104 formation (King *et al.*, 2007; Li *et al.*, 2008), whereas the microalgal WSD has either WS or
105 DGAT activity or both (Tomiyama *et al.*, 2017; Cui *et al.*, 2018). The fourth DGAT form is the
106 cytosol-localized soluble DGAT3, which is distinct from the other three DGAT subfamilies that
107 are localized in the endoplasmic reticulum (ER) (Xu *et al.*, 2018). In terms of WS, three
108 unrelated subfamilies have been identified so far, including the previously mentioned
109 bifunctional WSD, mammalian WS, and plant-like WS subfamilies (Rottig and Steinbuchel,
110 2013). The mammalian WS is a member of the DGAT2/acyl-CoA:monoacylglycerol
111 acyltransferase family. Plant-like WS including those from jojoba and *E. gracilis* belongs to the
112 same MBOAT family as DGAT1. Recently, *A. thaliana* chloroplastic phytyl ester synthases
113 (PES) and cyanobacterium acyltransferase slr2103 have been characterized to display both
114 DGAT and WS (using phytol instead of fatty alcohol) activities but are not related to any above-
115 mentioned DGAT and WS subfamilies (Lippold *et al.*, 2012; Aizouq *et al.*, 2020).

116 Up to now, proteins in the plant-like WS subfamily are the least characterized in plants
117 and microalgae especially in terms of evolution, structure-function relationships, catalytic
118 properties and physiological functions. *Chromochloris zofingiensis* is a unicellular green
119 microalga that has recently received attention as an emerging model microalga due to its
120 abundant lipid production (Roth *et al.*, 2017; Mao *et al.*, 2019). *C. zofingiensis* contains a
121 putative *WS* gene (*CzWS1*) that accumulates to high abundance under high-light stress and
122 nitrogen starvation (Roth *et al.*, 2017; Zhang *et al.*, 2019, 2020). However, no detectable level of
123 WE has been found in *C. zofingiensis* and many other microalgal species (Zhang *et al.*, 2017),
124 rendering the function of WS in these microalgae mysterious.

125 The current study aims to functionally characterize CzWS1 using *in vitro* and *in vivo*
126 assays as a means of exploring its roles in storage lipid biosynthesis and gaining insight into the
127 evolutionary, structural and catalytic features of the plant-like WS subfamily. Our findings
128 suggest that plant-like WS broadly exists in plants and microalgae but has divergent features.
129 The evolution of WS was primarily driven by strong purifying selection while selection
130 constraints, especially positive selection, contributed partially to the functional divergence. Co-
131 expression analysis showed that different from jojoba, *CzWS1* has weak expression correlation

with genes involved in the formation of WE precursors. Further *in vitro* and *in vivo* assays found that CzWS1 displays both WS and DGAT activity and affected storage lipid accumulation when expressed in yeast and the model green microalga *Chlamydomonas reinhardtii*. Overall, our results suggest the possible functional divergence of plant-like WS in plants and microalgae.

Materials and Methods

Sequence analysis of WS and DGAT related proteins

To identify the putative WS and DGAT related proteins in *C. zofingiensis*, WS, WSD1, PES1 and 2, and DGAT1 and 2 from *A. thaliana*, WS from *E. gracilis* and *S. chinensis*, diacylglycerol acetyltransferase from *Euonymus alatus* (EaDAcT) and WSD from *A. baylyi* were used as queries to BLAST against the *C. zofingiensis* genomic database (v.5.2.3.2; https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Czofingiensis_er, accessed on June 20, 2020). Putative WS, WSD, and DGAT proteins from different animal, plant, bacteria and microalgal species (Supplemental Table S1) were subjected to multiple sequence alignments using the L-INS-i method implemented in MAFFT v7 web server (https://mafft.cbrc.jp/alignment/server/; accessed on October 12, 2020; Katoh *et al.*, 2017). The resulting alignments were tested for the best fit protein alignment model using the IQ-TREE web server (http://iqtree.cibiv.univie.ac.at/, accessed on October 12, 2020; Trifinopoulos *et al.*, 2016) and a maximum likelihood phylogenetic tree was then constructed with the best fit model (VT+F+I+G4) using IQ-TREE and visualized with iTOL v4 (Letunic and Bork, 2016). The sequence logo of the conserved motif of six representative MBOAT enzymes was generated using the WebLogo server (https://weblogo.berkeley.edu/logo.cgi, accessed on June 16, 2020). The topology organization of WS and DGAT related proteins was predicted using Phobius (Käll *et al.*, 2007). The structures of CzWS1, EgWS and EaDAcT were predicted using Phyre2 software (Kelley *et al.*, 2015) with *Streptococcus thermophilus* DltB (c6buhH) as a template (Ma *et al.*, 2018).

Evolution and functional motif analyses of the WS gene family across green plants

To identify *WS* genes and their homologs across green plants, the *WS* sequence from *A. thaliana* was used as a query to BLAST against the whole genome sequences of 51 species, which cover

161 the main lineages of green plants including green algae, bryophytes, lycophytes and
162 angiosperms. The identified sequences were downloaded from the Phytozome database
163 (http://www. phytozome.net/, accessed on August 2, 2020). Sequences with obvious annotation
164 errors, such as lack of start or stop codon, improper bases, and short length (<500bp), were
165 deleted. We then built a phylogeny with the remaining sequences and the sequences with outlier
166 long branches were further removed. In the end, 276 sequences were selected for phylogenetic
167 tree construction. For phylogeny construction, protein-coding DNA sequences (CDS) were first
168 aligned using the MAFFT sequence alignment program with default settings
169 (https://www.ebi.ac.uk/Tools/msa/mafft/, accessed on August 2, 2020). Then, jModelTest 0.1.1
170 analysis (Posada, 2008) was carried out to select the best-fit model under the Akaike Information
171 Criterion framework (Akaike, 1974). The result of jModelTest indicates that the best-fit
172 substitution model to determine the evolution of the *WS* family is the General Time Reversible
173 model with the shape of the gamma distribution plus the proportion of invariable sites.
174 According to the best-fit model, the maximum likelihood phylogenetic tree was generated via the
175 CIPRES Web Portal http://www.phylo.org (accessed on August 2, 2020) using MrBayes 3.2.7a
176 (Huelsenbeck and Ronquist, 2001) with 1,000,000 generations, four Markov chains, and two
177 runs. To identify the functional motifs of WS proteins, we performed MEME analysis (Bailey
178 and Elkan, 1994) (http://meme.nbcr.net/meme/cgi-bin/meme.cgi, accessed on October 20, 2020)
179 with the following parameter settings: distribution of motifs = zero or one occurrence per
180 sequence, maximum number of motifs = 15, and optimum motif width = six to 50 residues.

181 ***Selection pressure analysis on the WS gene family across green plants***

182 The selection pressure operating on each of four subfamilies was estimated using the
183 nonsynonymous-to-synonymous substitution rate ratio (ω) as an indicator (Yang and Bielawski,
184 2000; Anisimova and Kosiol, 2009) with $0 < \omega < 1$, $\omega = 1$ and $\omega > 1$ corresponding to purifying
185 selection, neutral evolution and positive selection, respectively. The estimation of ω ratio was
186 performed in the CodeML program of the PAML (Phylogenetic Aanlysis by Maximum
187 Likelihood) package (version 4.9j), with the equilibrium frequency for each sense codon set to
188 F3x4. We first used the simple one-ratio site-model M0 (model = 0 and nonsynonymous site = 0)
189 to estimate branch lengths, which were then used as initial values for more complicated models.

The Clade model C (CmC) specified by model = 3 and nonsynonymous site = 2 in the CodeML program was used to test divergent selective pressures among *WS* subfamilies. CmC assumes that the phylogeny can be divided into foreground and background partitions. For each analysis, the clade of interest was set as the foreground partition while the remining phylogeny was set as the background partition. The CmC contains three site classes: 0, 1 and 2. Site classes 0 and 1 correspond to codons that experiencing either purifying selection ($0 < \omega_0 < 1$) or neutral evolution ($\omega_1 = 1$), while site class 2 accounts for codons that experience divergent selection pressure in which independent ω is estimated to the background ($\omega_2 > 0$) and foreground ($\omega_3 > 0$) partitions. The model M2a_rel which assumes the same ω between the foreground and background partitions is used as a null model for CmC analyses to test the presence of divergent selection. Each dataset was run three times from different starting ω to avoid local optima in the likelihood surface. To identify potential positive selection, we used the branch-site model A to construct branch-site test 2, which is also known as the branch-site test of positive selection. The branch-site test has four site classes: 0, 1, 2a and 2b. The first two classes are the same as the ones in the CmC. For sites classes 2a and 2b, positive selection is allowed on codons in the foreground branches but codons in background branches are under either purifying selection or neutral evolution, respectively. The null hypothesis of the branch-site model A is the same model but with ω_2 fixed to 1. The posterior probability (PP) for positively selected amino acid sites was calculated using the Bayes empirical Bayes (BEB) method implemented for the branch-site model A.

For both CmC and branch-site model analyses, significant differences between the alternative and null models were evaluated by comparing twice the difference of log-likelihood scores between the models against a χ^2 distribution with the degree of freedom equal to the difference in the numbers of free parameters between two models.

Expression correlation of WS with lipid biosynthetic genes in C. zofingiensis and S. chinensis

Expression data of lipid biosynthetic genes were searched against the previously released transcriptome databases of *C. zofingiensis* (Roth *et al.*, 2017) and *S. chinensis* (Sturtevant *et al.*, 2020). The fold changes of gene expression in *C. zofingiensis* at 0.5, 1, 3, 6 and 12 h after high-light stress relative to control condition and gene expression at different development stages of *S. chinensis* seeds were used for cluster analysis using the ClustVis software (Metsalu and Vilo,

2015) with average linkage and correlation distance method. Pearson's correlation coefficient was calculated between *WS* and each of lipid biosynthetic genes in *C. zofingiensis* and *S. chinensis* using Excel.

Construct preparation

The coding sequence of *CzWS1* (Phytozome accession number: Cz02g29020; Roth *et al.*, 2017) was chemically synthesized (General Biosystems, Morrisville, NC). For yeast heterologous expression, the coding sequence of *CzWS1* and its N/C-terminal truncation mutants were amplified by PCR and inserted downstream of the galactose-inducible *GAL1* promoter in the pYES2.1 yeast expression vector (Invitrogen, Burlington, Canada). For membrane yeast two-hybrid assay, the coding sequence of *CzWS1* was cloned into the pBT3N bait vector and pPR3N prey vector (kindly provided by Dr. Igor Stagljar, University of Toronto; Snider *et al.*, 2010) to yield the pBT3N: CzWS1 (bait) and pPR3N: CzWS1 (prey), respectively. For over-expression of *CzWS1* in *C. reinhardtii*, the coding sequence of *CzWS1* was cloned into the pOpt2_mVenus_Paro vector (Wichmann *et al.*, 2018; obtained from the Chlamydomonas Resource Center, http://www.chlamycollection.org/; accessed on 20 April 2020) using the ClonExpress II One Step Cloning Kit (Vazyme, Nanjing, China). The primers used in the current study are listed in Supplemental Table S2. The integrity of all vector sequences was confirmed by sequencing.

Yeast transformation, heterologous expression, and membrane yeast two-hybrid assay

For yeast heterologous expression, the constructs were transformed into the quadruple mutant strain *Saccharomyces cerevisiae* H1246 (*MATα are1-Δ::HIS3, are2- Δ::LEU2, dga1-Δ::KanMX4, lro1-Δ::TRP1 ADE2*), which is devoid of storage lipid (TAG and steryl esters) biosynthesis ability (Sandager *et al.*, 2002), using the lithium acetate mediated method (Gietz and Schiestl, 2007). The yeast transformants were grown at 30ºC with shaking at 220 rpm in liquid minimal medium [0.67% (w/v) yeast nitrogen base, 0.2% (w/v) synthetic complete medium lacking uracil (SC-Ura) and 2% (w/v) raffinose] till the mid-log phase before inoculated the induction medium [0.67% (w/v) yeast nitrogen base, 0.2% (w/v) SC-Ura, 2% (w/v) galactose, and 1% (w/ v) raffinose] at an initial optical density of 0.4 at 600 nm (OD_{600}). For the fatty acid and alcohol feeding experiment, yeast cells were cultured in the induction medium

249 with the supplementation of 200 mM of oleic acid (C18:1Δ^{9cis}, 18:1), linoleic acid
250 (C18:2$\Delta^{9cis,12cis}$, 18:2), α-linolenic acid (18:3$\Delta^{9cis,12cis,15cis}$, 18:3) or cetyl alcohol (C16:0-OH).

251 The membrane yeast two-hybrid system was performed as described previously (Snider
252 *et al.*, 2010; Xu *et al.*, 2019*a*). In brief, the pBT3N:CzWS1 (bait) was then co-transformed with
253 the pPR3N:CzWS1 (prey), Ost-N_{ub}I 'positive' control prey or Ost-N_{ub}G 'negative' control prey
254 into the yeast strain NMY51 [*MATa, his3Δ200, trp1-901, leu2- 3,112, ade2, LYS2::(lexAop)4-*
255 *HIS3,ura3::(lexAop)8-lacZ,ade2::(lexAop)8-ADE2, GAL4*] using the protocol described above.
256 The successful transformants were spotted on synthetic drop-out agar plates lacking Ade, His,
257 Leu and Trp for probing possible interaction.

258 ***In vitro enzyme assays***

259 Yeast microsomal fractions containing recombinant CzWS1 or variant enzymes were isolated
260 from yeast cells at mid-log growth stage as described previously (Xu *et al.*, 2020). *In vitro* WS
261 and DGAT assays were performed according to the procedures described previously with
262 modification (Arne *et al.*, 2017; Xu *et al.*, 2020). In brief, WS or DGAT activity was measured
263 in a 60-μL reaction mixture at 30°C for 4-30 min with shaking, which contains 200 mM HEPES-
264 NaOH (pH 7.4), 3.2 mM $MgCl_2$, 15 μM [1-^{14}C] oleoyl-CoA (56 μCi/μmol; American
265 Radiolabeled Chemicals, St. Louis, MO, USA), 333 μM *sn*-1,2-diolein [dispersed in 0.2% (v/v)
266 Tween 20; for DGAT assay] or 167 μM cetyl alcohol [dispersed in 0.5 μL dimethyl sulfoxide
267 (DMSO) for WS assay], and 5-20 μg of microsomal protein. The reaction was initiated and
268 terminated with the addition of 10 μL of microsomal fractions containing recombinant enzyme
269 and 10 μL of 10% (w/v) SDS, respectively. The reaction products were separated by a thin-layer
270 chromatography (TLC) plate (0.25 mm Silica gel, DC-Fertigplatten, Macherey-Nagel, Germany)
271 with hexane/diethyl ether/acetic acid (80:20:1, v/v/v) and the resolved lipids were visualized by
272 phosphorimaging (Typhoon Trio Variable Mode Imager, GE Healthcare, Mississauga, ON,
273 Canada). The corresponding WE or TAG spots were then scraped from the plates and quantified
274 for radioactivity using a LS 6500 multi-purpose scintillation counter (Beckman-Coulter,
275 Mississauga, ON, Canada).

276 For substrate specificity assay, 15 μM [1-^{14}C] oleoyl-CoA (56 μCi/μmol) or palmitoyl-
277 CoA (60 μCi/mmol) and 167 μM cetyl alcohol or stearyl alcohol (C18:0-OH) were used in the
278 WS assay. For kinetic assay, the reaction was performed using 6.4 μg of microsomal protein for

279 8 min with the concentration of [1-^{14}C] oleoyl-CoA varying from 0.1 to 30 μM while cetyl
280 alcohol keeping constant at 167 μM. The kinetic parameters were calculated by fitting the data to
281 the Michaelis-Menten or allosteric sigmoidal equation using the GraphPad Prism program
282 (version 6.0; GraphPad Software, La Jolla, CA, USA).

283 ***Algal strain, growth conditions, C. reinhardtii transformation and quantitative RT-PCR***
284 ***analysis***

285 *C. reinhardtii* CC-5325 (cw15, mt-) was obtained from the Chlamydomonas Resource Center
286 (http://www.chlamycollection.org/; accessed on 20 April 2020). *C. reinhardtii* cells were grown
287 in Tris acetate phosphate (TAP) medium (Harris, 1989) under continuous illumination (50 µmol
288 photons $m^{-2}s^{-1}$) at 22°C with continuous shaking at 110 rpm. For nitrogen starvation, mid-log
289 phase cells were collected by centrifugation (1500 x g, 5 min), washed twice with nitrogen
290 deficient TAP medium (TAP-N, without NH_4Cl), and resuspended in TAP-N medium at an
291 OD_{750} of 0.25. *C. reinhardtii* transformation was performed using the glass bead mediated
292 transformations as described previously (Kindle, 1990). The vectors were linearized by *Eco31*
293 and transformed and the positive transformants were selected on TAP agar plates containing 10
294 mg/L paromomycin.

295 The expression levels of *CzWS1* in the transgenic *C. reinhardtii* cells were analyzed by
296 quantitative RT-PCR (qPCR) on a StepOnePlus Real-Time PCR System (Applied Biosystems,
297 USA) with the Platinum SYBR Green qPCR Master Mix (Invitrogen) with the primers shown in
298 Table S2 as described previously (Xu *et al*., 2017). In brief, total RNA was isolated from
299 microalgal cells treated with nitrogen starvation using the Sigma Spectrum Plant Total RNA kit
300 (Sigma-Aldrich Canada Co., Oakville, ON), which was used to synthesize first-strand cDNA
301 using the SuperScript IV first-strand cDNA synthesis kit (Invitrogen). *CBLP* was used as the
302 internal reference gene.

303 ***Lipid analysis***

304 Total lipids were extracted from yeast and microalgal cells using the Bligh and Dyer method
305 (Bligh and Dyer, 1959). The extracted lipids were further separated on a TLC plate (0.25 mm
306 Silica gel, DC-Fertigplatten) using the solvent system hexane/diethyl ether/glacial acetic acid
307 (80:20:1, v/v/v). Lipid classes were visualized with 10% cupric acetate (w/v) and 10%
308 phosphoric acid (v/v) followed by charring at 180°C for 5-10 min. The TAG and WE bands were

identified according to triheptadecanoin and jojoba oil standards, respectively. For TAG quantification, triheptadecanoin (C17:0 TAG) was added as an internal standard. TAG binds were visualized by primulin and transmethylated, and the resulting fatty-acid methyl esters were analyzed using GC as described previously (Xu *et al*., 2020).

Results

Plants and microalgae contain different WS and DGAT related enzymes for storage lipid biosynthesis

To explore the possible players involved in storage lipid biosynthesis in plants and microalgae, the putative WS and DGAT related enzymes catalyzing the last step of storage lipid biosynthesis were analyzed. Alignment and phylogenetic analysis of 79 putative WS and DGAT related enzymes from microalgae, plants, animals and bacteria divide these enzymes into five groups: (i) the DGAT1 group, (ii) the plant-like WS group, (iii) the DGAT2 group, (iv) the PES group and (v) the WSD group (Fig. 1A). Jojoba WS (ScWS) clusters with WS from microalgae *E. gracilis* (EgWS) and CzWS1 (the plant-like WS group), mammal WS clusters with DGAT2 (the DGAT2 group), while bifunctional WSD, DGAT1 and PES form three separate groups. The plant-like WS group enzymes share an origin with the DGAT1 group enzymes (Fig. 1B), where they both belong to the MBOAT superfamily and show a typical MBOAT membrane topology feature (Figs. 1B and C and Supplemental Table S1). The structures of three members from the plant-like WS group, including CzWS1, EgWS and EaDAcT were predicted using the Phyre2 software with bacteria MBOAT *S. thermophilus* DltB (c6buhH) as a template (Ma *et al.*, 2018). Structural alignment shows that these acyltransferases appear to share a certain level of conservation in structural features, although the catalytic histidine exhibited subtle differences in position (Figs. 1D and E).

WS broadly exists in plants and microalgae but has divergent features

To further study the evolutionary relationship of *WS* genes among the green plants, we constructed the maximum likelihood tree using 276 *WS* coding sequences from 51 species across green plants. As shown in Fig. 2, the *WS* gene family can be divided into four major clades with high bootstrap values. Agal *WS* genes are phylogenetically divergent from the land-plant *WS* and

group together (Clade I) at the base of the tree. Inside the land plants, *WS* genes of basal linages forms the monophyletic clade (Clade II), while those from monocots and dicots further diverged from each other and form clades III to IV. We further performed MEME analysis to identify the conserved and divergent signature motifs of WS proteins. As shown in Supplemental Figure S1, the motif patterns match well with cladding patterns in the phylogenetic tree. The motif composition of algal WS is very different from that in land plants. Among land plants, significant patterns of motif conservation are present within and between WS clades, indicating that certain motifs have been conserved throughout WS evolution.

WS evolution has been driven primarily by purifying selection and partially by positive selection

In order to test possible differences in evolutionary patterns between different subfamilies, we estimated the selective pressure acting on each subfamily using nonsynonymous-to-synonymous substitution rate ratio (ω) as an indicator. We first applied Clade model C (CmC) with either the entire algal, bryophyte/lycophyte, monocot or dicot clade set as the foreground partition; all other branches comprised the background partition. As shown in Table 1, in all cases, CmC fits the data significantly better than the M2a_rel null model ($P < 0.05$). Parameter estimates indicate that the majority of sites (83%-85.4%) evolved under stronger purifying selection and a small set of sites (4.6-9.3%) evolved under divergent selective pressure, indicating that strong purifying selection plays a central role in the evolution of WS family to maintain its function. In addition, the ω ratio for the divergent selection site class decreases from algae to land plants, suggesting that selective constraint has been strengthened after the divergence from algae to land plants, and may indicate the action of stronger positive selection in land plants as well. We next used the branch-site model A to test the possibility of positive selection on algal and land plant clades. The test showed a highly significant support of positive selection for both algae and land plant WS subfamilies (Table 2), suggesting that the evolution of WS has been partly driven by positive selection.

Co-expression analysis of WS with lipid biosynthetic genes in C. zofingiensis and jojoba suggests potential functional divergence of CzWS from ScWS

The green microalga *C. zofingiensis* can accumulate storage lipids in response to stress conditions (Roth *et al.*, 2017; Mao *et al.*, 2019) and contains multiple WS and DGAT related

368 enzymes that may contribute to this process, including the previously characterized CzDGAT1
369 and CzDGAT2 (Mao *et al.*, 2019; Xu *et al.*, 2019*b*, 2020), 4 putative CzWS isoforms (CzWS1-
370 4), 1 putative CzWSD isoform and 2 putative CzPES isoforms (CzPES1 and 2) (Figs. 1 and 3).
371 The transcriptional response of these acyltransferases to high-light stress was searched against a
372 transcriptome database (Roth *et al.*, 2017). *CzWS1* and *CzDGAT2B* were found to be largely up-
373 regulated during high light with *CzWS1* displaying the highest transcript abundance (Fig. 3 and
374 Supplemental Figure S2). Based on the fold change of the expression levels at 5 different time
375 points under high-light stress compared to control condition (Roth *et al.*, 2017), these genes can
376 be clustered into at least two clusters. CzWS1 falls into the same cluster with CzDGAT1A and
377 CzDGAT2A and 2B, which contribute largely to TAG biosynthesis in *C. zofingiensis* (Mao *et*
378 *al.*, 2019). Expression correlation of *CzWS* with each of lipid biosynthetic genes in *C.*
379 *zofingiensis* was determined by calculating the corresponding Pearson correlation coefficients
380 and was compared with those from jojoba (Sturtevant *et al.*, 2020) (Fig. 3). Both *CzWS1* and
381 *ScWS* showed positive correlations with many genes involved in plastid fatty acid biosynthesis
382 and oil body formation. Interesting, positive correlation between *WS* and other genes in WE
383 formation, including *Fatty Acid Elongase* (*FAE*) and *Fatty Acid Reductase* (*FAR*) were found in
384 jojoba but not in *C. zofingiensis*. Considering the transcript abundance of *CzWS1* during stress, it
385 is possible that CzWS1 may have a divergent function from jojoba WS.

386 ***CzWS1 encodes an active WS***

387 To explore the functionality of CzWS1, its coding sequence was expressed in the yeast mutant
388 H1246, which is devoid of storage lipid biosynthesis ability (Sandager *et al.*, 2002). Considering
389 CzWS1 is closely related to EgWS, which catalyzes esterification of medium-chain fatty acyl-
390 CoAs and medium-chain fatty alcohols in *E. gracilis* (Teerawanichpan and Qiu, 2010), the WS
391 activity of CzWS1 was tested by culturing the H1246 yeast expressing *CzWS1* in the presence or
392 absence of fatty alcohol. CzWS1 was able to produce a considerable amount of WE in yeast
393 when the yeast cells were fed with fatty alcohol (Fig. 4A and Supplemental Figure S3),
394 suggesting that *CzWS1* encodes an active WS. Furthermore, *in vitro* enzyme assay using yeast
395 microsomal fractions confirmed that CzWS1 displayed a strong WS activity (Fig. 4B). To
396 explore the enzyme kinetic behavior, the activity of CzWS1 was analyzed over increasing
397 concentrations of oleoyl-CoA (Fig. 4C). CzWS1 activity increased gradually with the increase of
398 oleoyl-CoA concentration and reached the maximum activity at 25-30 μM oleoyl-CoA (Fig. 4C).

399 The substrate saturation curve was further fitted to different kinetic equations and the allosteric
400 sigmoidal equation gave a better fit with a R^2 of 0.9981. The Hill coefficient of CzWS1 was
401 1.66±0.04, suggesting it exhibited positive cooperativity. Usually, allosteric modulation of
402 enzymes is also accompanied with the formation of functional oligomers (Liu *et al.*, 2012).
403 Therefore, the possible self-interaction of CzWS1 was probed using membrane yeast two-hybrid
404 assay. The results, however, showed no protein-protein interaction between CzWS1 with itself
405 under our experimental conditions (Fig. 4D), suggesting that CzWS1 may function in the form of
406 a monomer. To characterize the substrate specificity of CzWS1, 16:0-OH or 18:0-OH and 18:1-
407 CoA or 16:0-CoA were supplied in separate assays (Fig. 4E and F). CzWS1 displayed a large
408 preference towards 16:0-OH over 18:0-OH by 4.5 and 4.3 folds when 16:0-CoA (Fig. 4E) and
409 18:1-CoA (Fig. 4F) were used as acyl donors, respectively. Moreover, CzWS1 showed a higher
410 substrate specificity for 18:1-CoA than 16:0-CoA, with CzWS1 displaying 2.2 and 2.3-fold
411 higher activity with 18:1-CoA than 16:0-CoA when using 16:0-OH and 18:0-OH as acyl
412 acceptors, respectively (Fig. 4E and F).

413 ***CzWS1 displays DGAT activity and affects storage lipid accumulation in yeast and microalgae***

414 The ability of CzWS1 in utilizing different substrates as acyl acceptors was further explored by
415 yeast complementation and *in vitro* enzyme activity assays. H1246 yeast expressing *CzWS1* was
416 able to accumulate a small amount of TAG (Figs. 5A and B and Supplemental Figure S4),
417 suggesting that CzWS1 displays a weak DGAT activity. The TAG accumulation of H1246 yeast
418 expressing *CzWS1* was further accessed by feeding the yeast cells with exogenous fatty acids,
419 including oleic acid (18:1), linoleic acid (18:2), and α-linolenic acid (18:3). Interesting, CzWS1
420 was able to incorporate 18:1 and 18:2 rather than 18:3 into yeast TAG (Figs. 5A and B and
421 Supplemental Figure S4). Further *in vitro* enzyme assay using yeast microsomal fractions
422 confirmed that CzWS1 displayed a weak DGAT activity (Fig. 5C). Together, these results
423 suggested that CzWS1 is a bifunctional enzyme with both WS and DGAT activities.

424 To explore the possible physiological role of CzWS1, the coding sequence of *CzWS1* was
425 expressed in the model green alga *C. reinhardtii* and the resulting transgenic lines were subjected
426 to nitrogen starvation. The expression of *CzWS1* in the transgenic *C. reinhardtii* lines was
427 confirmed using qPCR (Fig. 5D) and no significant difference was observed in cell growth and
428 neutral lipid production between the lines expressing CzWS1 and empty vector (Supplemental

Figure S5). After 3 days of nitrogen starvation, no detectable amount of WE was observed in *CzWS1* expressing lines. Interestingly, CzWS1 appears to affect the fatty acid composition of TAG by increasing 18:1 and decreasing 16:0 (Fig. 5E), which is consistent with the substrate preference as shown in Figs. 4E and 4F.

C-terminal domain is crucial for WS activity and may be important for functional specialization

To gain more insight on WS from both structure and evolutionary perspectives, the predicted positive selection sites were mapped on the structure model of CzWS1 (Figs. 6A and B). As a typical MBOAT enzyme, CzWS1 is predicted to contain 9 TMDs with 26 and 49 amino-acid long N-terminal and C-terminal hydrophilic tails, respectively (Fig. 6C). The structure homology model of CzWS1 in general agrees on the multiple TMDs feature and the presence of a N-terminal hydrophilic tail (Fig. 6C). The C-terminal domain of CzWS1 appears to be embedded inside the proposed substrate binding tunnel (Fig. 6A and Supplemental Figure S6) and may be sensitive to modification. Among the 30 predicted positive selection sites occurring throughout the whole CzWS1 peptide, 2 and 6 sites were located at the N- and C-termini, which account for 7.7% and 12.3% of the amino acid residues of these regions, respectively, in contrast to 6.8% in TMDs (Fig. 6B and Supplemental Table S3). Considering positive selection may facilitate functional divergence, more positively selected sites in the C-terminal domain might indicate its importance for functional specialization.

To explore the possible role of the N- and C-termini in CzWS1 activity, the full-length CzWS1 (1-399) and its N- and C-terminal truncated variants were produced in yeast H1246 and the resulting microsomal fractions containing the recombinant enzymes were used for *in vitro* WS activity. Removal of the first 20 amino acid residues (21-399), which represents the majority of the N-terminal hydrophilic tail (Fig. 6D), led to a 63% reduction in enzyme specificity activity (Fig. 6D), suggesting that the N-terminal tail is not dispensable for the enzyme activity. Further removal of the first (44-399) and second (71-399) TMDs, however, totally abolished the enzyme activity (Fig. 6D). C-terminal truncation appears to affect the enzyme activity more pronounced. Truncation of the last 29 amino acid residues (1-371), which is partial of the C-terminal tail (Fig. 6C), completely inactivated the enzyme, whereas removal of the last 5 amino acid residues (1-394) reduced the enzyme activity by 66% (Fig. 6D). Taken together, these results indicated that

459 the C-terminus of CzWS1 is crucial for WS activity and may have roles in functional
460 specialization.

461

462 **Discussion**

463 Plant and microalgal WS and DGAT related enzymes form five phylogenetically different
464 groups (Fig. 1) and are responsible for storage lipids biosynthesis. Among them, the plant-like
465 WS group is the least explored especially in terms of evolution, structure-function features and
466 enzyme characterization. Our phylogenetic analysis on *WS* gene family across the green plants
467 showed that *WS* broadly exists in plants and algae but algal *WS* is phylogenetically divergent
468 from that of bryophytes/lycophytes, monocots and dicots (Fig. 2 and Supplemental Figure S1).
469 Strong purifying selection appears to have contributed primarily to the evolution of *WS* family to
470 main enzyme function, while selective constraints have been strengthened after the divergence
471 from algae to land plants, suggesting the action of positive selection may act stronger in land
472 plants than algae and partially drive functional divergence (Tables 1 and 2).

473 Plant-like WS group acyltransferases belong to the MBOAT family (Fig. 1C), which are
474 integral membrane-bound enzymes containing 7-10 TMDs (Supplemental Table S1). By aligning
475 the predicted structures of CzWS1, EgWS and EaDAcT with the recently solved crystal structure
476 of DltB, a Gram-positive bacteria MBOAT, it appears that these acyltransferases share certain
477 structural conservation with each other, although the proposed substrate binding tunnel and
478 funnel in DltB had spatial variations (Supplemental Figures S6). In addition, the N-terminal
479 regions of CzWS1, EgWS and EaDAcT had less structural conservation and led to different
480 funnel opening configurations (Figs. 1 and 6 and Supplemental Figure S6). Indeed, the first two
481 small α-helixes of CzWS1 (corresponding to amino acid residue 1-20) that cover the funnel
482 opening are absent in EgWS and EaDAcT. This region appears to be not pivotal to enzyme
483 function, since CzWS1 still retained 37% enzyme activity when the first two α-helixes were
484 removed (Figs. 1 and 6 and Supplemental Figure S6). The C-terminal domain of CzWS1, on the
485 other hand, is embedded in the proposed substrate binding tunnel and sensitive to modification,
486 where the removal of the last 29 amino acid residues completely inactivated the enzyme (Fig. 6).
487 Considering the tunnel is crucial for substrate binding and catalytic action in MBOAT (Ma *et al.*,
488 2018), the C-terminal domain of CzWS1 might be important for both WS and DGAT activity.

489 This region might also be importance for functional specialization since it contains multiple
490 positively selected sites (Fig. 6 and Supplemental Table S3). Moreover, it is interesting to note
491 that the removal of the last 5 amino acid residues (395QLNFL399) led to a 66% decrease in
492 enzyme activity (1-394, Fig. 6), though these residues are less conserved among the plant-like
493 WS family (Supplemental Figure S1). Interestingly, the predicted structure homology model
494 (Supplemental Figure S6) suggests these residues might form a part of a small α-helix in the
495 tunnel opening region which is to some extent structurally related to the C-terminus of EaDAcT
496 but is absent in other plant-like WS such as EgWS. Therefore, the C-terminus of CzWS1 might
497 affect its substrate specificity. Further structure resolution and biochemical analysis of the
498 individual enzymes are certainly warranted to explain the importance of the C-terminal region of
499 CzWS1 and the catalytic mechanisms.

500 Many microalgae such as *C. zofingiensis* have no detectable WE but abundant expression
501 of *WS* genes. Expression correlation of *CzWS1* and *ScWS* with lipid biosynthetic genes (Fig. 3)
502 showed that unlike *ScWS* which displays positive correlation with genes (*FAR* and *FAE*)
503 involved the formation of precursors for WE formation, expression of *CzWS1* has weak
504 correlation with the putative *FAR* and *FAE* genes with an exception to *CzFAR1*, suggesting that
505 CzWS1 may have divergent functions from ScWS. Indeed, although yeast complementation
506 assay and *in vitro* enzyme characterization showed that *CzWS1* encodes an active WS (Fig. 4),
507 we were not able to detect the formation of WE in *C. reinhardtii* transgenic lines expressing
508 *CzWS1* alone or together with *FAR* from *E. gracilis* or jojoba, in *N. benthamiana* leaves
509 expressing *CzWS1* and in high-light stressed or nitrogen starved *C. zofingiensis* cells. It may be
510 possible that fatty alcohol is not the physiological substrate for CzWS1 and the *in vitro* assay
511 results may not be able to represent what happens in physiological conditions, despite that *C.*
512 *zofingiensis* contains at least five *FAR* homologs in its genome with considerable expression
513 levels (Supplemental Figure S7). Since plants can accumulate free phytol, a primary C20
514 isoprenoid alcohol, under stress conditions as a result of chlorophyll degradation (Lippold *et al.*,
515 2012), we also checked whether CzWS1 can use phytol as an acyl acceptor by yeast feeding
516 assay, but no obvious formation of phytol ester was detected. Furthermore, the H1246 yeast
517 expressing *CzWS1* was only able to accumulate a small amount of TAG (Fig. 5) and there was no
518 detectable formation of fatty acid ethyl ester or sterol ester, suggesting that CzWS1 may not be
519 able to use ethanol and sterol as acyl acceptors. *C. zofingiensis* accumulates a substantial amount

520 of astaxanthin in the form of astaxanthin ester but the identity of acyltransferase catalyzing the
521 acylation of astaxanthin is still unknown. CzWS1 has been proposed to be responsible for this
522 step (Zhang *et al*., 2019; Zhang *et al*., 2020; Roth *et al*., 2017). We, therefore, tested whether
523 CzWS1 is also able to acylate astaxanthin using in vitro enzyme assay with the recombinant
524 enzyme in yeast microsomal fractions, but no detectable astaxanthin ester was formed under our
525 experimental conditions.

526 It is interesting to note that CzWS1 displays weak DGAT activity, slightly restores TAG
527 biosynthesis in yeast mutant H1246 and affects TAG formation in *C. reinhardtii* (Fig. 5).
528 Although many acyltransferases from the DGAT1, DGAT2, WSD and PES groups have
529 previously been found to exhibit both WS and DGAT activities and are able to catalyze the
530 esterification of acyl-CoA to both fatty alcohol and DAG molecules (Kalscheuer *et al.*, 2003;
531 Turkish *et al.*, 2005; Yen *et al.*, 2005; Du *et al.*, 2014; Aizouq *et al.*, 2020), the members from
532 the plant-like WS group appear to have rigid substrate specificity (Teerawanichpan and Qiu,
533 2010). Currently, only a couple of plant WSs (EgWS and jojoba WS) from the plant-like WS
534 group have been characterized *in vitro* (Lardizabal *et al.*, 2000; Teerawanichpan and Qiu, 2010).
535 EgWS was found to be merely capable of WS synthesis rather than TAG synthesis, whereas
536 jojoba WS displays a predominant WS activity but appears to be able to produce a slight amount
537 of TAG *in vitro* using 16:0-CoA (Teerawanichpan and Qiu, 2010; Miklaszewska and Banaś,
538 2016).

539 While the physiological function of CzWS1 in *C. zofingiensis* remains further
540 exploration, heterologous expression of *CzWS1* in *C. reinhardtii* may have shed some light on
541 the role of CzWS1 in green microalgae. It appears that CzWS1 may have a profound effect in
542 TAG biosynthesis rather that WE formation in *C. reinhardtii* (Fig. 5), as no detectable WE was
543 found. This is different from the *in vitro* enzyme activity results where recombinant CzWS1 in
544 yeast microsomal fractions displayed an extremely low DGAT activity but a high WS activity
545 (Figs. 4 and 5). This might be explained by the limited availability of fatty alcohol substrate for
546 WE formation in *C. reinhardtii* and the differences between *in vitro* and physiological
547 conditions. Therefore, knocking out and overexpression of *CzWS1* in *C. zofingiensis*, along with
548 comprehensive physiological characterizations of the mutants, would expand our understanding
549 of its functions in this green microalgal species.

550 The effect of CzWS1 on TAG formation in yeast and microalgae and its preference of
551 CzWS1 for 18:1-CoA over 16:0-CoA (Fig. 4 and 5) may have physiological and applicable
552 importance. Indeed, under high-light and nitrogen starvation conditions, the fatty acid profile of
553 *C. zofingiensis* total lipids and TAG was shifted with a slight decrease in 16:0 and a large
554 increase in 18:1 (Zhu *et al.*, 2015; Liu *et al.*, 2016; Mao *et al.*, 2018). Considering the up-
555 regulation of *CzWS1* under stress conditions, it is possible that the substrate preference of
556 CzWS1 may contribute to the altered fatty acid profile. In line with this, expression of *CzWS1* in
557 *C. reinhardtii* affected the fatty acid profile of TAG by decreasing 16:0 and increasing 18:1 (Fig.
558 5E). Furthermore, the ability of CzWS1 in the production of WE in H1246 yeast (Fig. 4)
559 suggests that CzWS1 may have potential in engineering WE production. Especially considering
560 that most of the WSs identified so far have low preference towards 18:1-CoA (Lardizabal *et al.*,
561 2000; Stöveken *et al.*, 2005; Teerawanichpan and Qiu, 2010; Cui *et al.*, 2018), CzWS1 with
562 18:1-CoA preference (Fig. 4) may represent a promising target for producing 18:1-enriched WE,
563 such as oleyl oleate, which has favorite properties for lubrication (Yu *et al.*, 2018).

564 In conclusion, our results showed that the plant-like WS is broadly existent in algae and
565 plants but has divergent features. Evolution of plant-like WS across the green plants underwent
566 strong purifying selection while positive selection contributed partially to functional divergence
567 but to a weaker level in algae than land plants. *CzWS1* had weak expression correlation with
568 genes involved in the formation of WE precursors and may have divergent function. CzWS1 was
569 *in vitro* characterized as a bifunctional enzyme with strong WS and weak DGAT activities and
570 can affect storage lipid formation when expressed in yeast and *Chlamydomonas*. The C-terminal
571 region of CzWS1 with multiple predicted positive selection sites is crucial for enzyme function
572 and may have possible importance for functional specialization. Overall, the findings in this
573 study provide information on the evolution, enzyme function and catalytic properties of plant-
574 like WS and may shed insight into engineering wax biosynthesis in oleaginous organisms.

575 **Supplementary data**

576 Supplemental Table S1. Protein sequences of WS and DGAT related enzymes used for multiple
577 sequence alignment.

578 Supplemental Table S2. Primers used in the current study.

579 Supplemental Table S3. Predicted transmembrane domain (TMD) and positive selection sites of
580 CzWS1.

581 Supplemental Figure S1. Conserved motifs on wax synthases among green plants.

582 Supplemental Figure S2. Transcriptional change of *C. zofingiensis* genes encoding WS and
583 diacylglycerol acyltransferase (DGAT) related enzymes in response to high-light stress (data
584 from Roth et al., 2017).

585 Supplemental Figure S3. TLC analysis of wax ester (WE) from H1246 yeast expressing *CzWS1*
586 cultured in the presence (+) or absence (-) of 16:0-OH.

587 Supplemental Figure S4. TLC analysis of TAG from H1246 yeast expressing *CzWS1* cultured in
588 the presence or absence of various fatty acids.

589 Supplemental Figure S5. Overexpression of *CzWS1* in *Chlamydomonas reinhardtii*.

590 Supplemental Figure S6. Overlay of the predicted three-dimensional structures of bacteria DltB,
591 CzWS1, *Euonymus alatus* diacylglycerol acetyltransferase (EaDAcT), and *Euglena gracilis* WS
592 (EgWS).

593 Supplemental Figure S7. Identification of putative fatty acid reductase (*FAR*) genes in *C.*
594 *zofingiensis* and their transcriptional changes in response to high-light stress (data from Roth et
595 al., 2017).

596

597

598

599

600

601 **Acknowledgements**

602 We acknowledge the support provided by the University of Alberta Start-up Research Grant,
603 Natural Sciences and Engineering Research Council of Canada Discovery Grant (RGPIN-2016-
604 05926) and the Canada Research Chairs Program. We are grateful to Drs. Xiao Qiu (University
605 of Saskatchewan) and Edgar B. Cahoon (University of Nebraska-Lincoln) for providing the
606 constructs containing *EgFAR* and *SchFAR*, respectively, and Dr. Igor Stagljar (University of
607 Toronto) for providing the membrane yeast two-hybrid system.

608

609 **Data availability statement**

610 All relevant data can be found within the manuscript and its supplementary data.

611

612 **Author contributions**

613 GC conceived and supervised the experiment; YX designed and performed most of the
614 experiments, analyzed the data and prepared the initial draft of the manuscript. XP performed
615 phylogenetic and selection pressure analyses of WS from green plants and wrote the related
616 parts. JL and JW helped with *Chlamydomonas* transformation and yeast lipid analysis. JL, JW
617 and QS contributed valuable discussion during this study. JS contributed to discussion and
618 manuscript editing. All authors contributed to the preparation of the final article.

619

620 **Conflict of interest**

621 The authors declare that they have no conflicts of interest with the content of this article.

References:

Aizouq M, Peisker H, Gutbrod K, Melzer M, Hölzl G, Dörmann P. 2020. Triacylglycerol and phytyl ester synthesis in *Synechocystis* sp. PCC6803. Proceedings of the National Academy of Sciences of the United States of America **117**, 6216–6222.

Akaike H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control **19**, 716–723.

Anisimova M, Kosiol C. 2009. Investigating protein-coding sequence evolution with probabilistic codon substitution models. Molecular Biology and Evolution **26**, 255–271.

Arne JM, Widjaja-adhi MAK, Hughes T, Huynh KW, Silvaroli JA, Chelstowska S, Moiseenkova-bell VY, Golczak M. 2017. Allosteric modulation of the substrate specificity of acyl-CoA wax alcohol acyltransferase 2. Journal of Lipid Research **58**, 719–730.

Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the International Conference on Intelligent Systems for Molecular Biology **2**, 28–36.

Bligh EG, Dyer WJ. 1959. A rapid method of total lipid extraction and purification. Canadian Journal of Biochemistry and Physiology **37**, 911–917.

Cui Y, Zhao J, Wang Y, Qin S, Lu Y. 2018. Characterization and engineering of a dual-function diacylglycerol acyltransferase in the oleaginous marine diatom *Phaeodactylum tricornutum*. Biotechnology for Biofuels **11**, 32.

Du X, Herrfurth C, Gottlieb T, Kawelke S, Feussner K, Rühling H, Feussner I, Maniak M. 2014. *Dictyostelium discoideum* Dgat2 can substitute for the essential function of Dgat1 in triglyceride production but not in ether lipid synthesis. Eukaryotic Cell **13**, 517–526.

Gietz RD, Schiestl RH. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nature Protocols **2**, 31–34.

Harris EH. 1989. *The chlamydomonas sourcebook*. Academic Press: San Diego.

Hu Q, Sommerfeld M, Jarvis E, Ghirardi M, Posewitz M, Seibert M, Darzins A. 2008. Microalgal triacylglycerols as feedstocks for biofuel production: Perspectives and advances. The Plant Journal **54**, 621–639.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17**, 754–755.

Käll L, Krogh A, Sonnhammer ELL. 2007. Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server. Nucleic Acids Research **35**, 429–432.

Kalscheuer R, Steinbuchel A, Steinbüchel A. 2003. A novel bifunctional wax ester synthase/acyl-CoA:diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in Acinetobacter calcoaceticus ADP1. Journal of Biological Chemistry **278**, 8075–8082.

Katoh K, Rozewicki J, Yamada KD. 2017. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics **20**, bbx108–bbx108.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols **10**, 845–858.

Kindle KL. 1990. High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. Proceedings of the National Academy of Sciences of the United States of America **87**, 1228–1232.

King A, Nam JW, Han JX, Hilliard J, Jaworski JG. 2007. Cuticular wax biosynthesis in Petunia petals: cloning and characterization of an alcohol-acyltransferase that synthesizes wax-esters. Planta **226**, 381–394.

Lardizabal KD, Metz JG, Sakamoto T, Hutton WC, Pollard MR, Lassner MW. 2000. Purification of a jojoba embryo wax synthase, cloning of its cDNA, and production of high levels of wax in seeds of transgenic Arabidopsis. Plant Physiology **122**, 645–655.

Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Research **44**, W242–W245.

Lewandowska M, Keyl A, Feussner I. 2020. Wax biosynthesis upon danger: its regulation upon abiotic and biotic stress. New Phytologist **227**, 698–713.

Li F, Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R, Kunst L. 2008. Identification of the wax ester synthase/acyl-coenzyme A:diacylglycerol acyltransferase WSD1 required for stem

wax ester biosynthesis in Arabidopsis. Plant Physiology **148**, 97–107.

Lippold F, vom Dorp K, Abraham M, *et al.* 2012. Fatty acid phytyl ester synthesis in chloroplasts of Arabidopsis. The Plant Cell **24**, 2001–2014.

Liu J, Mao X, Zhou W, Guarnieri MT. 2016. Simultaneous production of triacylglycerol and high-value carotenoids by the astaxanthin-producing oleaginous green microalga *Chlorella zofingiensis*. Bioresource Technology **214**, 319–327.

Liu Q, Siloto RMP, Lehner R, Stone SJ, Weselake RJ. 2012. Acyl-CoA:diacylglycerol acyltransferase: molecular biology, biochemistry and biotechnology. Progress in Lipid Research **51**, 350–377.

Lu J, Xu Y, Wang J, Singer SD, Chen G. 2020. The role of triacylglycerol in plant stress response. Plants **9**, 1–13.

Ma D, Wang Z, Merrikh CN, Lang KS, Lu P, Li X, Merrikh H, Rao Z, Xu W. 2018. Crystal structure of a membrane-bound O-acyltransferase. Nature **562**, 286–290.

Mao X, Wu T, Kou Y, Shi Y, Zhang Y, Liu J. 2019. Characterization of type I and type II diacylglycerol acyltransferases from the emerging model alga *Chlorella zofingiensis* reveals their functional complementarity and engineering potential. Biotechnology for Biofuels **12**, 28.

Mao X, Wu T, Sun D, Zhang Z, Chen F. 2018. Differential responses of the green microalga Chlorella zofingiensis to the starvation of various nutrients for oil and astaxanthin production. Bioresource Technology **249**, 791–798.

Metsalu T, Vilo J. 2015. ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Research **43**, W566–W570.

Miklaszewska M, Banaś A. 2016. Biochemical characterization and substrate specificity of jojoba fatty acyl-CoA reductase and jojoba wax synthase. Plant Science **249**, 84–92.

Posada D. 2008. jModelTest: Phylogenetic model averaging. Molecular Biology and Evolution **25**, 1253–1256.

Roth MS, Cokus SJ, Gallaher SD, *et al.* 2017. Chromosome-level genome assembly and transcriptome of the green alga *Chromochloris zofingiensis* illuminates astaxanthin production. Proceedings of the National Academy of Sciences of the United States of America **114**, E4296–

E4305.

Rottig A, Steinbuchel A. 2013. Acyltransferases in bacteria. Microbiology and Molecular Biology Reviews **77**, 277–321.

Sandager L, Gustavsson MH, Stahl U, Dahlqvist A, Wiberg E, Banas A, Lenman M, Ronne H, Stymne S. 2002. Storage lipid synthesis is non-essential in yeast. Journal of Biological Chemistry **277**, 6478–6482.

Snider J, Kittanakom S, Damjanovic D, Curak J, Wong V, Stagljar I. 2010. Detecting interactions with membrane proteins using a membrane two-hybrid assay in yeast. Nature Protocols **5**, 1281–1293.

Stöveken T, Kalscheuer R, Malkus U, Reichelt R, Steinbüchel A. 2005. The wax ester synthase/acyl coenzyme A:diacylglycerol acyltransferase from Acinetobacter sp. strain ADP1: Characterization of a novel type of acyltransferase. Journal of Bacteriology **187**, 1369–1376.

Sturtevant D, Lu S, Zhou ZW, *et al.* 2020. The genome of jojoba (*Simmondsia chinensis*): A taxonomically isolated species that directs wax ester accumulation in its seeds. Science Advances **6**, 1–14.

Teerawanichpan P, Qiu X. 2010. Fatty acyl-coA reductase and wax synthase from euglena gracilis in the biosynthesis of medium-chain wax esters. Lipids **45**, 263–273.

Thelen JJ, Ohlrogge JB. 2002. Metabolic engineering of fatty acid biosynthesis in plants. Metabolic Engineering **4**, 12–21.

Tomiyama T, Kurihara K, Ogawa T, Maruta T, Ogawa T, Ohta D, Sawa Y, Ishikawa T. 2017. Wax ester synthase/diacylglycerol acyltransferase isoenzymes play a pivotal role in wax ester biosynthesis in *Euglena gracilis*. Scientific Reports **7**, 13504.

Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic acids research **44**, W232–W235.

Turkish AR, Henneberry AL, Cromley D, Padamsee M, Oelkers P, Bazzi H, Christiano AM, Billheimer JT, Sturley SL. 2005. Identification of two novel human acyl-CoA wax alcohol acyltransferases: Members of the diacylglycerol acyltransferase 2 (DGAT2) gene superfamily. Journal of Biological Chemistry **280**, 14755–14764.

Wichmann J, Baier T, Wentnagel E, Lauersen KJ, Kruse O. 2018. Tailored carbon partitioning for phototrophic production of (*E*)-α-bisabolene from the green microalga *Chlamydomonas reinhardtii*. Metabolic Engineering **45**, 211–222.

Xu Y, Caldo KMP, Falarz L, Jayawardhane K, Chen G. 2020. Kinetic improvement of an algal diacylglycerol acyltransferase 1 via fusion with an acyl-CoA binding protein. The Plant Journal **102**, 856–871.

Xu Y, Caldo KMP, Jayawardhane K, Ozga JA, Weselake RJ, Chen G. 2019*a*. A transferase interactome that may facilitate channeling of polyunsaturated fatty acid moieties from phosphatidylcholine to triacylglycerol. Journal of Biological Chemistry **294**, 14838–14844.

Xu Y, Caldo KMP, Pal-Nath D, Ozga J, Lemieux MJ, Weselake RJ, Chen G. 2018. Properties and biotechnological applications of acyl-CoA:diacylglycerol acyltransferase and phospholipid:diacylglycerol acyltransferase from terrestrial plants and microalgae. Lipids **53**, 663–688.

Xu Y, Falarz L, Chen G. 2019*b*. Characterization of type-2 diacylglycerol acyltransferases in the green microalga *Chromochloris zofingiensis*. Journal of Agricultural and Food Chemistry **67**, 291–298.

Xu Y, Chen G, Greer MS, Caldo KMP, Ramakrishnan G, Shah S, Wu L, Lemieux MJ, Ozga J, Weselake RJ. 2017. Multiple mechanisms contribute to increased neutral lipid accumulation in yeast producing recombinant variants of plant diacylglycerol acyltransferase 1. Journal of Biological Chemistry **292**, 17819-17831.

Yang Y, Benning C. 2018. Functions of triacylglycerols during plant development and stress. Current Opinion in Biotechnology **49**, 191–198.

Yang Z, Bielawski JP. 2000. Statistical methods for detecting molecular adaptation. Trends Ecol Evol **15**, 496–503.

Yen C-LE, Brown CH, Monetti M, Farese R V. 2005. A human skin multifunctional O-acyltransferase that catalyzes the synthesis of acylglycerols, waxes, and retinyl esters. Journal of Lipid Research **46**, 2388–2397.

Yu D, Hornung E, Iven T, Feussner I. 2018. High-level accumulation of oleyl oleate in plant seed oil by abundant supply of oleic acid substrates to efficient wax ester synthesis enzymes.

Biotechnology for Biofuels **11**, 53.

Zhang N, Mao Z, Luo L, Wan X, Huang F, Gong Y. 2017. Two bifunctional enzymes from the marine protist *Thraustochytrium roseum*: biochemical characterization of wax ester synthase/acyl-CoA:diacylglycerol acyltransferase activity catalyzing wax ester and triacylglycerol synthesis. Biotechnology for Biofuels **10**, 185.

Zhang Y, Shi M, Mao X, Kou Y, Liu J. 2019. Time-resolved carotenoid profiling and transcriptomic analysis reveal mechanism of carotenogenesis for astaxanthin synthesis in the oleaginous green alga *Chromochloris zofingiensis*. Biotechnology for biofuels **12**, 287.

Zhang Y, Ye Y, Ding W, Mao X, Li Y, Gerken H, Liu J. 2020. Astaxanthin is ketolated from zeaxanthin independent of fatty acid synthesis in *Chromochloris zofingiensis*. Plant Physiology **183**, 883–897.

Zhu S, Wang Y, Shang C, Wang Z, Xu J, Yuan Z. 2015. Characterization of lipid and fatty acids composition of *Chlorella zofingiensis* in response to nitrogen starvation. Journal of Bioscience and Bioengineering **120**, 205–209.

Tables

Table 1 Results from Clade model C analyses of the *WS* gene family across the green plants

Model	Foreground	*In*L	0: Purifying		1: Neutral		2: Divergent		2Δℓ	p-values
			ω_0	p_0	ω_1	p_1	ω_2, ω_3	p_2		
M2a_ref		-122935.277772	0.01009	0.85496	1.00000	0.09940	ω_2:0.43005	0.04564		
CmC I	Algae	-122835.608685	0.00916	0.82962	1.00000	0.07781	ω_2:0.34578 ω_3:0.99760	0.09257	199.33817	<0.001
CmC II	Bryophytes/ Lycophytes	-122931.028100	0.01023	0.85420	1.00000	0.09845	ω_2: 0.42263 ω_3: 0.60028	0.04735	8.49934	0.00355
CmC III	Monocots	-122932.226140	0.01001	0.85426	1.00000	0.09953	ω_2: 0.43863 ω_3: 0.35473	0.04620	6.10326	0.01349
CmC V+VI	Dicots	-122879.753084	0.00879	0.83719	1.00000	0.08820	ω_2: 0.60621 ω_3: 0.32014	0.07461	111.04938	<0.001

Table 2 Branch-site model A analyses of the *WS* gene family from algae and land plants

Foreground	Model	*In*L	Parameters	Site Class				2Δℓ	p-values
				0	1	2a	2b		
Algae	Null	-123050.783583	Proportion	0.80346	0.09226	0.09354	0.01074		
			Background ω	0.03379	1.00000	0.03379	1.00000		
			Foreground ω	0.03379	1.00000	1.00000	1.00000		
	Branch-Site	-123019.425088	Proportion	0.80850	0.09119	0.09015	0.01017	62.71699	<0.001
			Background ω	0.03636	1.00000	0.03636	1.00000		
			Foreground ω	0.03636	1.00000	1.65783	1.65783		
			Background ω	0.03542	1.00000	0.03542	1.00000		
			Foreground ω	0.03542	1.00000	2.73679	2.73679		
			Background ω	0.04356	1.00000	0.04356	1.00000		
			Foreground ω	0.04356	1.00000	2.33856	2.33856		
Land Plants	Null	-123232.662443	Proportion	0.8776	0.12168	0.00064	0.00009		
			Background ω	0.03681	1.00000	0.03681	1.00000		
			Foreground ω	0.03681	1.00000	1.00000	1.00000		
	Branch-Site	-122549.687569	Proportion	0.81945	0.15448	1.00000	0.00414	1365.94975	<0.001
			Background ω	0.04917	1.00000	0.04917	1.00000		
			Foreground ω	0.04917	1.00000	2.22392	2.22392		

Figure legends

Figure 1. Plants and microalgae contain multiple wax synthase (WS) and diacylglycerol acyltransferase (DGAT) related enzymes. A. Phylogenetic tree of WS and DGAT related enzymes from different organisms. WS, DGAT1, DGAT2, bifunctional WS/DGAT (WSD) and phytyl ester synthase (PES) are shown in blue, pink, grey, green and yellow, respectively. B. A closer look at the phylogenetic relationship of members from the membrane-bound *O*-acyl transferase (MBOAT) family. C. Sequence logo for the conserved MBOAT signature motifs. D and E. Overlay of the predicted three-dimensional structures (D) and catalytic histidine (E) of *Chromochloris zofingiensis* WS1 (CzWS1), *Euonymus alatus* diacylglycerol acetyltransferase (EaDAcT), and *Euglena gracilis* WS (EgWS). The predicted structures and catalytic histidine residues of CzWS1, EaDAcT and EgWS are shown in purple, blue and yellow, respectively. For A and B, the branches of enzymes from *Arabidopsis thaliana*, *Simmondsia chinensis* and *C. zofingiensis* are shown in green, blue and orange, respectively.

Figure 2. The maximum-likelihood phylogeny of wax synthase (WS) proteins from 51 green plant and algal species. The *WS* gene family can be divided into four major clades, representing *WS* genes from algae (Clade I), bryophytes/lycophytes (Clade II), monocots (Clade III) and eudicots (Clade IV). The maximum-likelihood tree was generated using the MrBayes 3.2.7a program with 1,000,000 generations, four Markov chains, and two runs. Numbers above branches represent the support values (Bayesian posterior probabilities). The scale bar represents the number of nucleotide replacements per site.

Figure 3. Cluster analysis of *WS* and *DGAT* related genes in *Chromochloris zofingiensis* and jojoba (*Simmondsia chinensis*) and their expression correlation with genes in lipid biosynthetic pathways. Expression of *C. zofingiensis* genes in response to high-light stress was searched against a transcriptome database (Roth *et al.*, 2017) and the fold changes of the gene expression levels at 0.5, 1, 3, 6 and 12 h after high-light stress compared to control condition are indicated as stages I to V, respectively. Gene expression at different development stages of *S. chinensis* seeds (early, mid and late developing seed and dry seed) was search against a transcriptome database (Sturtevant *et al.*, 2020) and is shown as stages I to IV, respectively. Cluster analysis was performed using ClustVis software with average linkage and correlation distance method. Expression correlation of *WS* and *DGAT* related genes with each of lipid

biosynthetic genes in *C. zofingiensis* and *S. chinensis* was determined by calculating the corresponding Pearson correlation coefficients.

Figure 4. *CzWS1* encodes an active wax ester synthase. A. TLC analysis of wax ester (WE) from H1246 yeast expressing *CzWS1* cultured in the presence (+) or absence (-) of 16:0-OH. B. *In vitro* wax ester synthase assay using yeast microsomal fractions containing recombinant CzWS1. C. Wax ester synthase activity of CzWS1 at oleoyl-CoA concentration from 0.1 to 30 μM. Data were fitted to the allosteric sigmoidal equation using the GraphPad Prism. D. CzWS1 cannot interact with itself as shown by membrane yeast two-hybrid assay. CzWS1 were ligated to the Lex A- C-terminal fragment of ubiquitin (C_{ub}) and the N-terminal fragment of ubiquitin containing an Ile/Gly point mutation ($N_{ub}G$), yielding C_{ub}-CzWS and $N_{ub}G$-CzWS, respectively. Serial dilutions of yeast cells producing each combination were spotted on synthetic drop-out (SD) agar plates lacking Ade, His, Leu and Trp (SD-A-H-L-T). E and F. Substrate specificity of CzWS1 towards 16:0-OH and 18:0-OH with 16:0-CoA (E) or 18:1-CoA (F) as an acyl donor. For B, C, E and F, data represent means ± SD (n = 3).

Figure 5. CzWS1 displays diacylglycerol acyltransferase (DGAT) activity and affects triacylglycerol (TAG) formation in yeast and *Chlamydomonas reinhardtii*. A. TLC analysis of TAG from H1246 yeast expressing *CzWS1* cultured in the presence or absence of various fatty acids (FAs). B. Yeast TAG content (w/w, dry weight). C. *In vitro* DGAT assay using yeast microsomal fractions containing recombinant CzWS1. D. Relative expression level of *CzWS1* in the transgenic *C. reinhardtii* line cultivated under 48 and 72 h nitrogen starvation. E. TAG fatty acid composition of the *C. reinhardtii* lines expressing *CzWS1* and the empty vector cultivated under 72 h nitrogen starvation. Data represent means ± SD of three biological replicates.

Figure 6. Predicted locations of positive selection sites in CzWS1 and truncation analysis of the N and C-terminal domains. A. Predicted structure homology model of CzWS1 showing the location of predicted positive selection sites in blue (a Bayes Empirical Bayes posterior probability >99%) and green (a Bayes Empirical Bayes posterior probability >95%). B. A closer look at the predicted positive selection sites at the N- and C- termini. C. Predicted topology of CzWS1. D. *In vitro* wax ester synthase activity of different CzWS1 truncated mutants. Data represent means ± SD (n = 3). For A to C, the numbers and color code indicate the different

truncation points. Amino acid residues 1-20, 21-43, 44-70, 71-370, 371-394, and 395-399 are shown in light grey, dark grey, black, purple, olive green, and tan yellow, respectively.

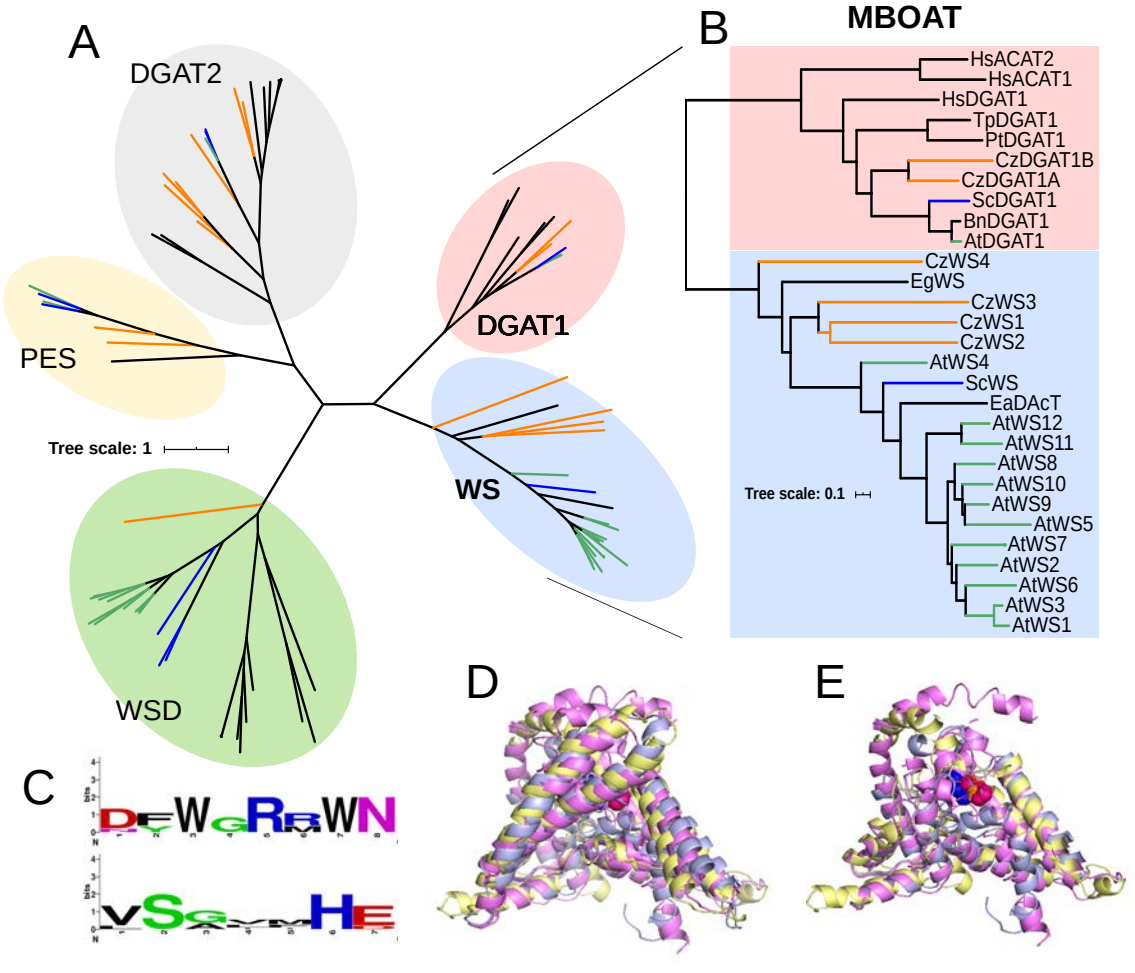


Figure 1. Plants and microalgae contain multiple wax synthase (WS) and diacylglycerol acyltransferase (DGAT) related enzymes.

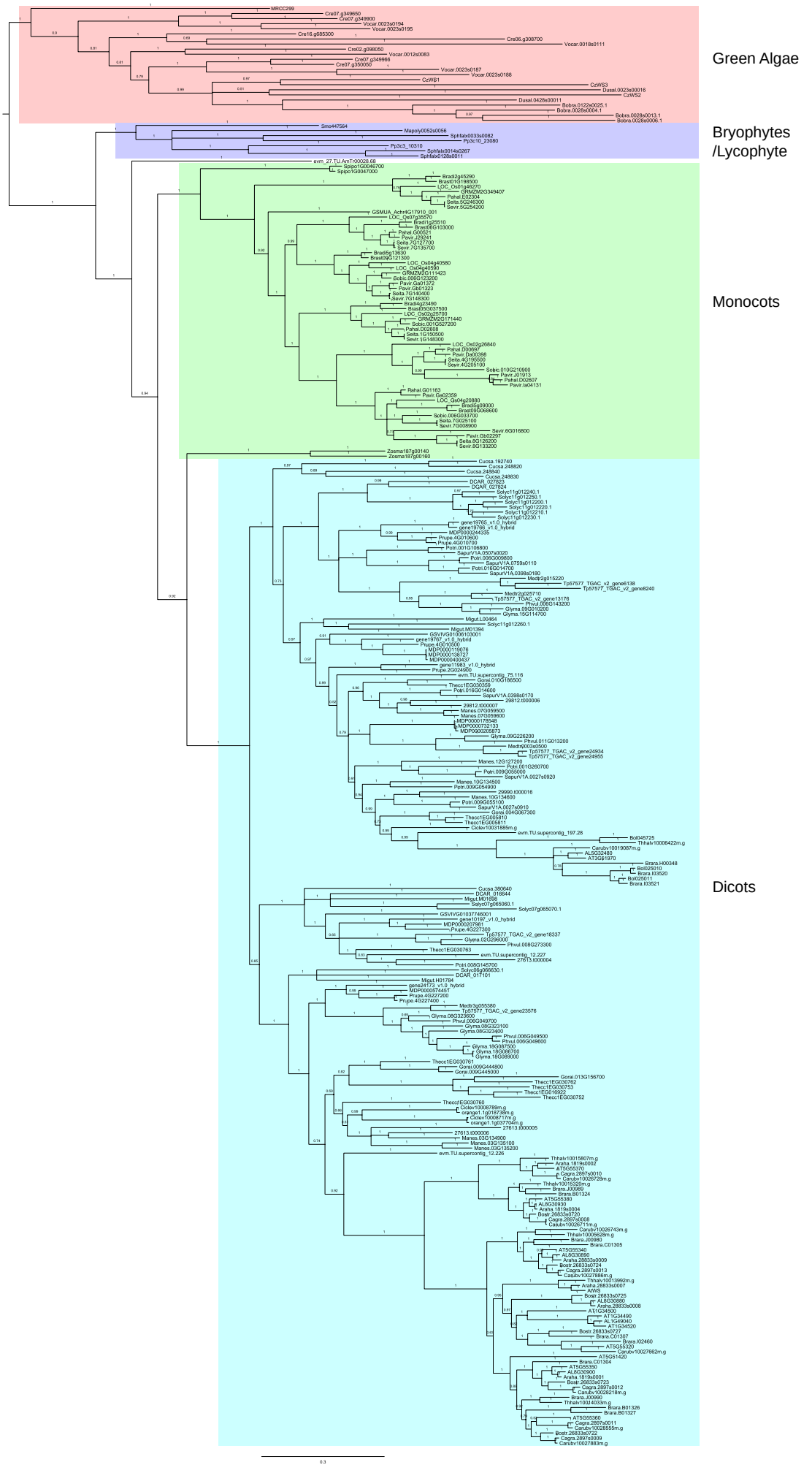


Figure 2. The maximum-likelihood phylogeny of wax synthase (WS) proteins from 1 51 green plant and algal species.

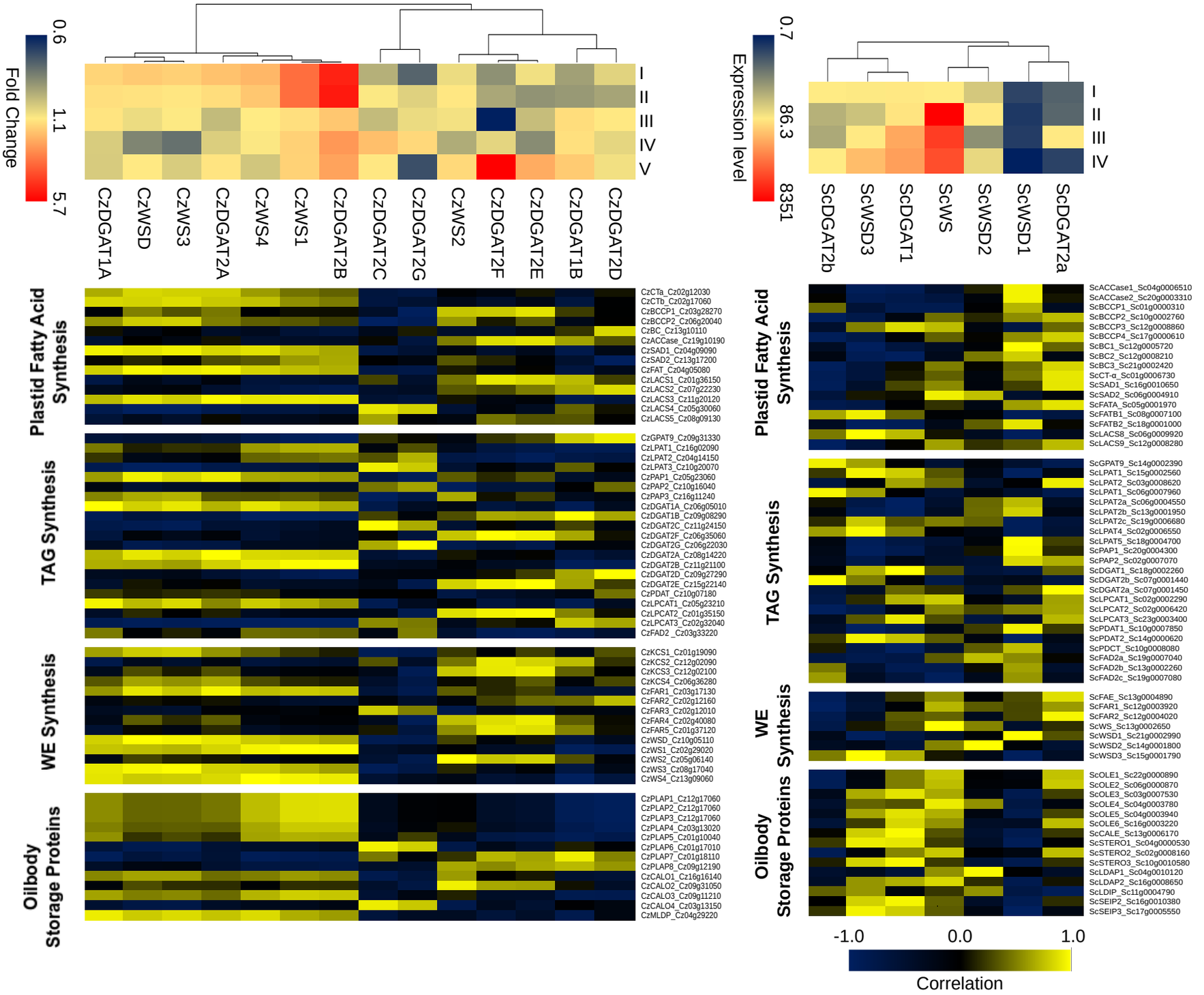


Figure 3. Cluster analysis of *WS* and *DGAT* related genes in *Chromochloris zofingiensis* and jojoba (*Simmondsia chinensis*) and their expression correlation with genes in lipid biosynthetic pathways.

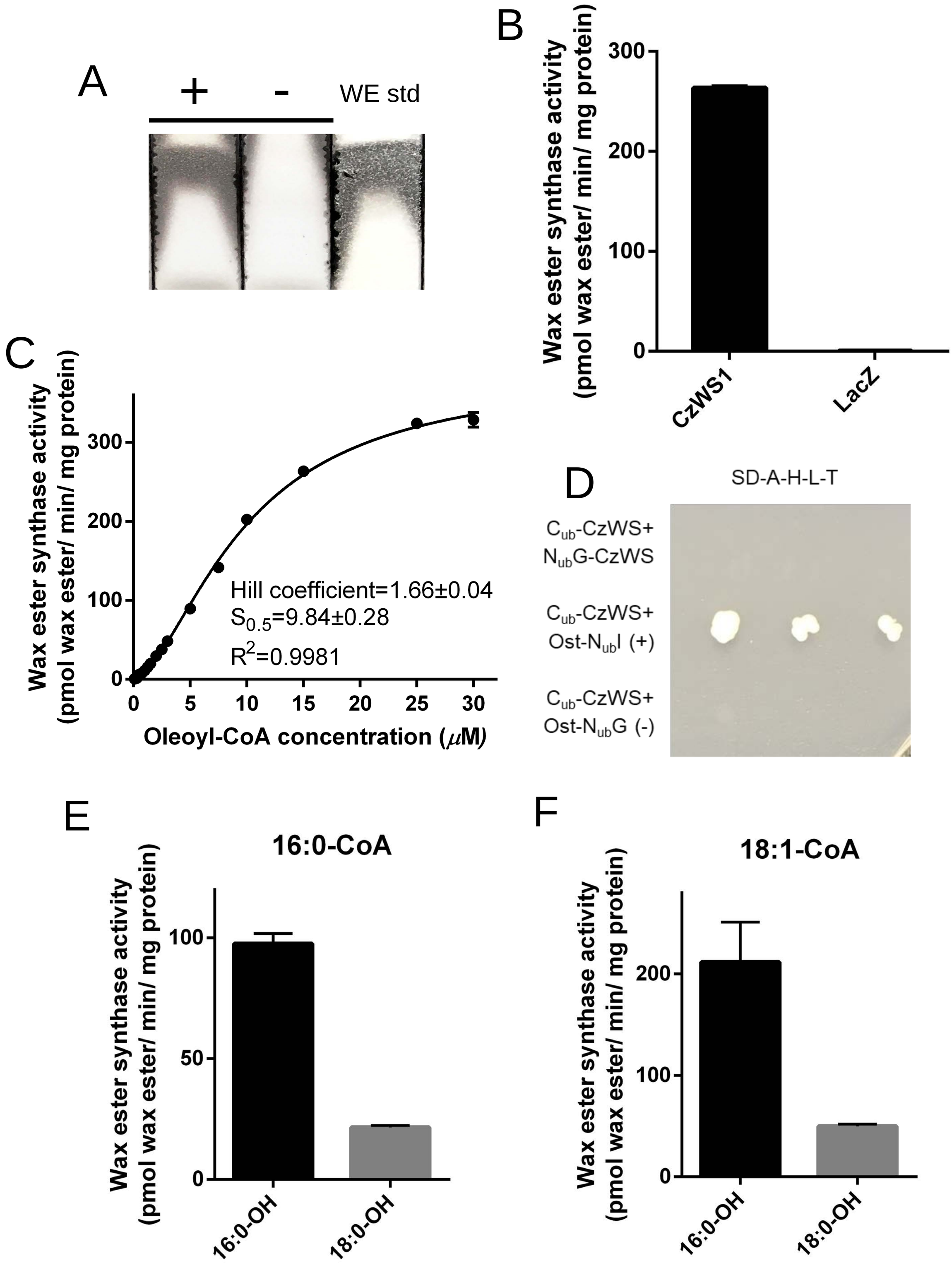


Figure 4. *CzWS1* encodes an active wax ester synthase.

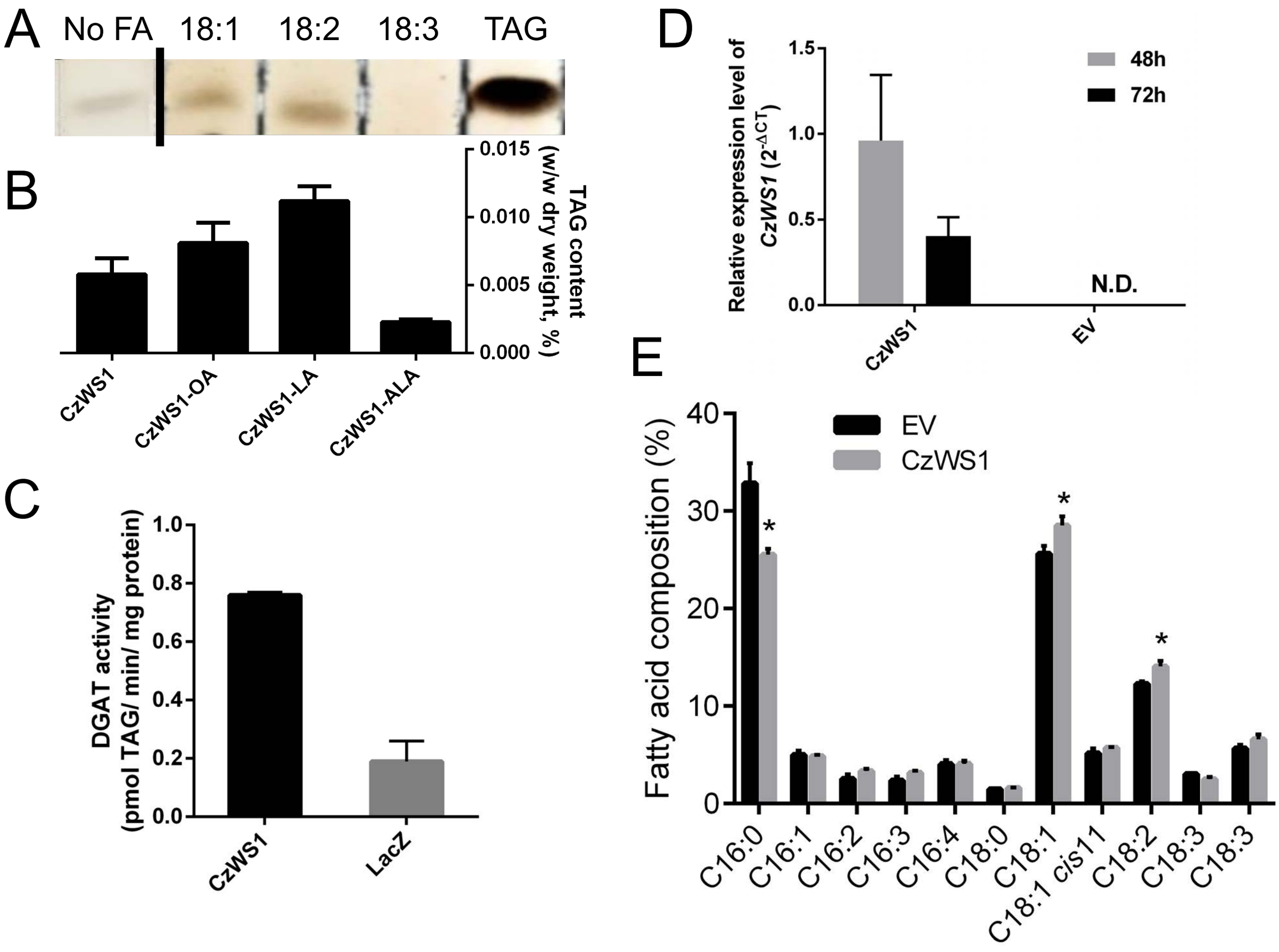


Figure 5. CzWS1 displays diacylglycerol acyltransferase (DGAT) activity and affects triacylglycerol (TAG) formation in yeast and *Chlamydomonas reinhardtii.*

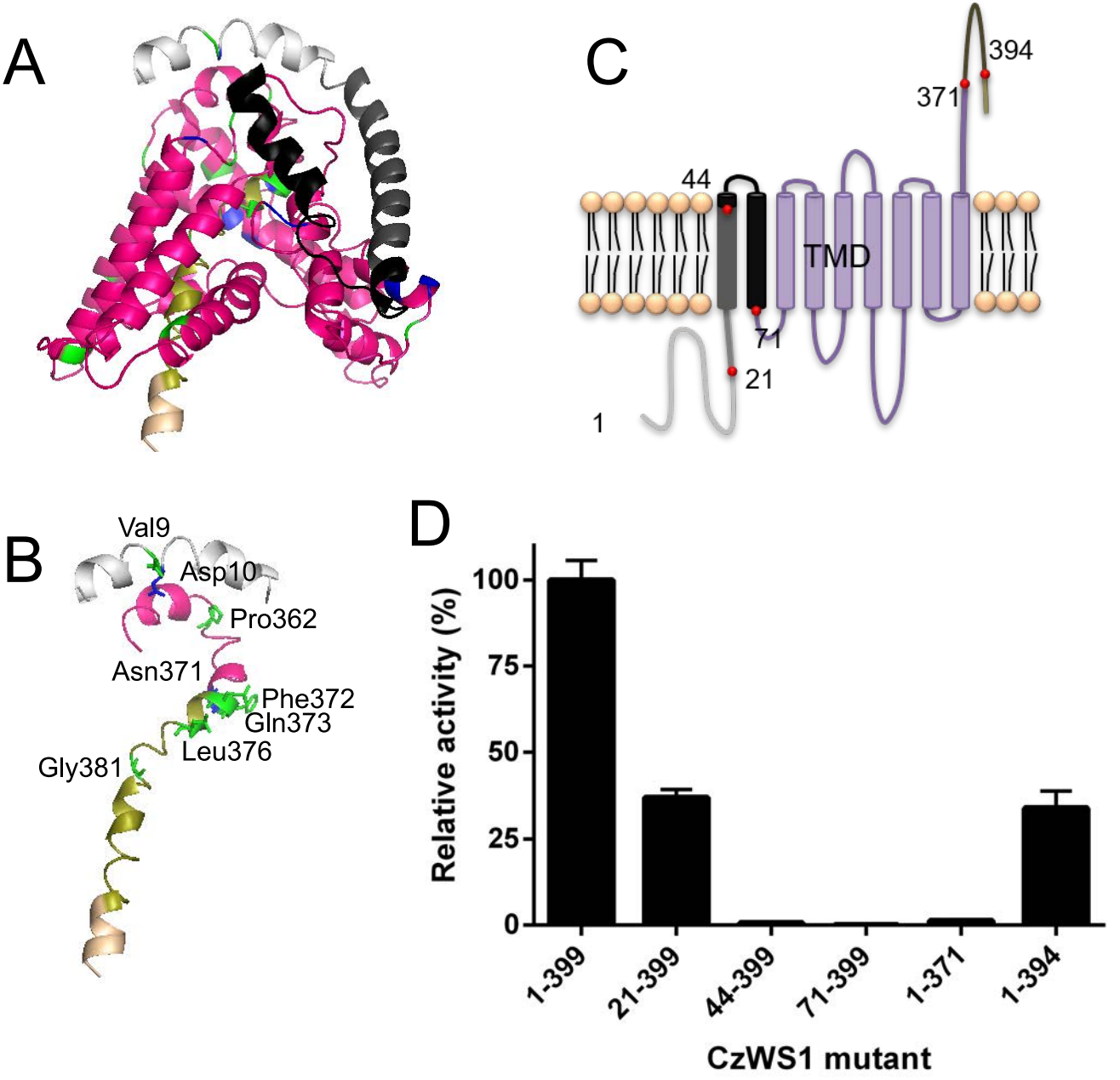


Figure 6. Predicted locations of positive selection sites in CzWS1 and truncation analysis of the N and C-terminal domains.

SUPPORTING INFORMATION

Supplemental Method S1. Subcellular localization of CzWS1 in *Nicotiana benthamiana* leaves.

Supplemental Table S1. Protein sequences of WS and DGAT related enzymes used for multiple sequence alignment. The topology organization was predicted using Phobius (Käll *et al.*, 2007). TMD, transmembrane domain.

Supplemental Table S2. Primers used in the current study. A restriction site is shown in **bold and underlined** and a Kozak translation initiation sequence for yeast expression is shown in *italic*.

Supplemental Table S3. Predicted transmembrane domain (TMD) and positive selection sites of CzWS1. TMD was predicted by Phobius. Positive selection sites with a Bayes Empirical Bayes posterior probability ≥ 99%, or ≥ 95%, are shown in **bold** or normal font, respectively.

Supplemental Figure S1. Conserved motifs on wax synthases among green plants.

Supplemental Figure S2. **Transcriptional change of *C. zofingiensis* genes encoding WS and diacylglycerol acyltransferase (DGAT) related enzymes in response to high-light stress (data from Roth *et al.*, 2017).**

Supplemental Figure S3. TLC analysis of wax ester (WE) from H1246 yeast expressing *CzWS1* cultured in the presence (+) or absence (-) of 16:0-OH.

Supplemental Figure S4. TLC analysis of TAG from H1246 yeast expressing *CzWS1* cultured in the presence or absence of various fatty acids.

Supplemental Figure S5. Overexpression of *CzWS1* in *Chlamydomonas reinhardtii*. Growth plots (A) and neutral lipid content (represented by Nile red value; B) of the *C. reinhardtii* lines expressing *CzWS1* and the empty vector cultivated under nitrogen starvation. Nile red value is calculated by dividing Nile red fluorescence of TAG (ΔF_{TAG}) by Nile red fluorescence of polar lipids (ΔF_{PL}). Data represent means ± SD of three biological replicates.

Supplemental Figure S6. Overlay of the predicted three-dimensional structures of bacteria DltB, CzWS1, *Euonymus alatus* diacylglycerol acetyltransferase (EaDAcT), and *Euglena gracilis* WS (EgWS). A. The substrate binding funnel and tunnel of DltB. B. Overlay of the

substrate binding funnel and tunnel of DltB, CzWS, EaDAcT and EgWS. The C-termini of CzWS1 and EaDAcT are circled in red. A closer view of the funnel and tunnel of DltB (C, H), CzWS (D, I), EaDAcT (E, J), EgWS (F, K) and overlay (G, L), respectively. Dltb structure, the predicted structures of CzWS, EaDAcT and EgWS, and their corresponding catalytic histidine are shown in red, purple, blue and yellow, respectively. As for A, C, and H, the residues that form the funnel and tunnel in DltB are shown in blue and red, respectively.

Supplemental Figure S7. Identification of putative *fatty acid reductase* (*FAR*) genes in *C. zofingiensis* and their transcriptional changes in response to high-light stress (data from Roth *et al.*, 2017).

Supplemental Figure S8. Subcellular localization of CzWS1 in *Nicotiana benthamiana* leaf cells. Venus::CzWS1 was co-localized with a SCFP3A-tagged *Arabidopsis thaliana* glycerol-3-phosphate acyltransferase (AtGPAT9), a known endoplasmic reticulum (ER) localized protein. Scale bars represent 20 μm.

Supplemental Method S1. Subcellular localization of CzWS1 in *Nicotiana benthamiana* leaves

For the subcellular localization, the coding sequences of *CzWS1* and *AtGPAT9* (used as an ER marker) were fused in frame to the downstream of *Venus* and upstream of *SCFP3C*, respectively, in the modified pGPTVII vector (kindly provided by Dr. Jörg Kudla, University of Münster) (Becker *et al.*, 1992; Gehl *et al.*, 2009), as described previously (Xu *et al.*, 2020). Subcellular localization of CzWS1 was examined by transient expression of *CzWS1* in *Nicotiana benthamiana* leaves as described previously (Xu *et al.*, 2019, 2020). Individual construct (*Venus::CzWS1*, *AtGPAT9::SCFP3C* and *p19* vector) was transformed to *Agrobacterium tumefaciens* GV3101 cells. *A. tumefaciens* cultures containing *Venus::CzWS1* or *AtGPAT9::SCFP3C* construct and the *p19* vector encoding a viral suppressor protein were mixed in a transformation medium, containing 50 mM MES, 2 mM Na_3PO_4, 0.5% (w/v) glucose and 0.1 mM acetosyringone. The final OD_{600} of each culture at 0.125 was used to infiltrate *N. benthamiana* leaves as described by Vanhercke *et al.*, 2013. The fluorescence of the lower epidermis of leaves after 2-3 days of infiltration was visualized using a fluorescent microscope with the excitation wavelengths for Venus and SCFP at 488 and 543 nm, and the emission filter wavelengths at 505–530 nm and 560–615 nm, respectively.

Supplemental Table S1. Protein sequences of WS and DGAT related enzymes used for multiple sequence alignment. The topology organization was predicted using Phobius (Käll *et al.*, 2007). TMD, transmembrane domain.

Name	Organism	Phytozome/GenBank accession number	# of predicted TMD
AbWSD1	*Acinetobacter baylyi (ADP1)*	AF529086	0
AbWSD2	*Acinetobacter baylyi*	WP_004922247	0
AtDGAT1	*Arabidopsis thaliana*	NM_127503	10
AtDGAT2	*Arabidopsis thaliana*	NM_115011	2
AtPES1	*Arabidopsis thaliana*	At1g54570	0
AtPES2	*Arabidopsis thaliana*	AT3G26840	0
AtWS1	*Arabidopsis thaliana*	At1g34490	8
AtWS2	*Arabidopsis thaliana*	At1g34500	8
AtWS3	*Arabidopsis thaliana*	At1g34520	8
AtWS4	*Arabidopsis thaliana*	At3g51970	9
AtWS5	*Arabidopsis thaliana*	At5g51420	11
AtWS6	*Arabidopsis thaliana*	At5g55320	8
AtWS7	*Arabidopsis thaliana*	At5g55330	9
AtWS8	*Arabidopsis thaliana*	At5g55340	9
AtWS9	*Arabidopsis thaliana*	At5g55350	7
AtWS10	*Arabidopsis thaliana*	At5g55360	9
AtWS11	*Arabidopsis thaliana*	At5g55370	8
AtWS12	*Arabidopsis thaliana*	At5g55380	9
AtWSD1	*Arabidopsis thaliana*	AT5G37300	1
AtWSD2	*Arabidopsis thaliana*	AT1G72110	1
AtWSD4	*Arabidopsis thaliana*	At3g49190	0
AtWSD5	*Arabidopsis thaliana*	At3g49200	1
AtWSD6	*Arabidopsis thaliana*	At3g49210	1
AtWSD7	*Arabidopsis thaliana*	At5g12420	3
AtWSD8	*Arabidopsis thaliana*	At5g16350	1
AtWSD9	*Arabidopsis thaliana*	At5g22490	3
AtWSD10	*Arabidopsis thaliana*	At5g53380	0
AtWSD11	*Arabidopsis thaliana*	At5g53390	1
BnDGAT1	*Brassica napus*	JN224473	10
CzDGAT1A	*Chromochloris zofingiensis*	MH523419	9
CzDGAT1B	*Chromochloris zofingiensis*	Cz09g08290	9
CzDGAT2A	*Chromochloris zofingiensis*	Cz08g14220	2
CzDGAT2B	*Chromochloris zofingiensis*	Cz11g21100	2

CzDGAT2C	*Chromochloris zofingiensis*	Cz11g24150	1
CzDGAT2D	*Chromochloris zofingiensis*	Cz09g27290	1
CzDGAT2E	*Chromochloris zofingiensis*	Cz15g22140	2
CzDGAT2F	*Chromochloris zofingiensis*	Cz06g35060	0
CzDGAT2G	*Chromochloris zofingiensis*	Cz06g22030	0
CzPES1	*Chromochloris zofingiensis*	Cz02g16070	0
CzPES2	*Chromochloris zofingiensis*	Cz07g16210	0
CzWS1	*Chromochloris zofingiensis*	Cz02g29020	9
CzWS2	*Chromochloris zofingiensis*	Cz05g06140	6
CzWS3	*Chromochloris zofingiensis*	Cz08g17040	5
CzWS4	*Chromochloris zofingiensis*	Cz13g09060	6
CzWSD	*Chromochloris zofingiensis*	Cz10g05110	1
EaDAcT	*Euonymus alatus*	ADF57327	7
EgWS	*Euglena gracilis*	GU733920.1	7
HsACAT1	*Homo sapiens*	NP_003092.4	9
HsACAT2	*Homo sapiens*	NP_003569.1	9
HsAWAT1	*Homo sapiens*	NP_001013597.1	2
HsAWAT2	*Homo sapiens*	NP_001002254.1	2
HsDGAT1	*Homo sapiens*	NP_036211.2	9
HsDGAT2	*Homo sapiens*	NP_115953.2	2
HsWS	*Homo sapiens*	AY605053	2
MaWSD1	*Marinobacter aquaeolei*	WP_011783747	0
MaWSD2	*Marinobacter aquaeolei VT8*	WP_011786509	0
MhWS1	*Marinobacter hydrocarbonoclasticus*	ABO21021	0
MhWS2	*Marinobacter hydrocarbonoclasticus*	EF219377	0
MmAWAT2	*Mus musculus*	NM_177746	2
MmWS	*Mus musculus*	AY611032	2
NgWSD	*Nannochloropsis gaditana*	EWM29694	1
PcWSD	*Psychrobacter cryohalolentis K5*	WP_011512619	0
PtDGAT1	*Phaeodactylum tricornutum*	HQ589265	9
PtDGAT2	*Phaeodactylum tricornutum*	AFQ23659	5
PtWSD	*Phaeodactylum tricornutum*	XP_002184474	1
RjWSD	*Rhodococcus jostii RHA1*	WP_011594556	0
SceDGAT2	*Saccharomyces cerevisiae*	NP_014888	1
ScDGAT1	*Simmondsia chinensis*	Sc18g0002260.01	9
ScDGAT2a	*Simmondsia chinensis*	Sc07g0001450.01	0
ScDGAT2b	*Simmondsia chinensis*	Sc07g0001440.01	1
ScPES1	*Simmondsia chinensis*	Sc03g0004810.01	0
ScPES2	*Simmondsia chinensis*	Sc09g0003660.01	0
ScWS	*Simmondsia chinensis*	AF149919; Sc13g0002650.01	9
ScWSD1	*Simmondsia chinensis*	Sc21g0002990.01	1

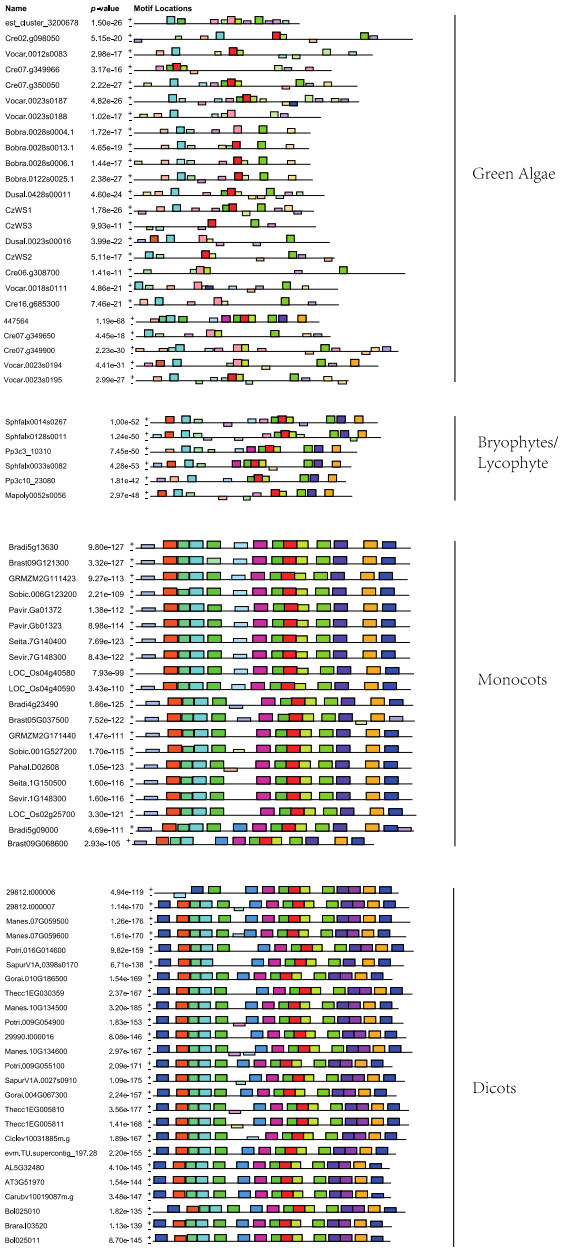
ScWSD2	*Simmondsia chinensis*	Sc14g0001800	1
ScWSD3	*Simmondsia chinensis*	Sc15g0001790	0
slr2103	*cyanobacterium Synechocystis sp. PCC6803*	WP_010872622	0
TpDGAT1	*Thalassiosira pseudonana*	XP_002287215	9
TpDGAT2	*Thalassiosira pseudonana*	XP_002286252	2
TrWSD	*Thraustochytrium roseum*	MF037228	0

Supplemental Table S2. Primers used in the current study. A restriction site is shown in **bold and underlined** and a Kozak translation initiation sequence for yeast expression is shown in *italic*.

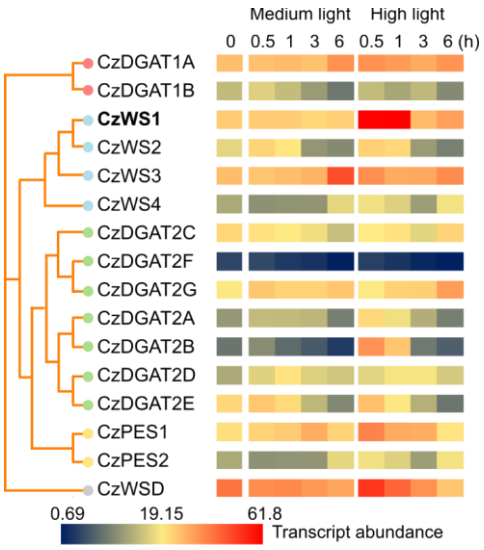
Primer name	Oligonucleotide sequence
CzWS1$_{1\text{-}400}$-F (to pYES2.1)	GCAGA**GCGGCCGC***GAA*ATGGAAGCAGCACTGACCA
CzWS1$_{1\text{-}400}$-R (to pYES2.1)	TAT**GTCGAC**AATGGCTGACTGTTGCACC
CzWS1$_{21\text{-}400}$-F (to pYES2.1)	GCAGA**GCGGCCGC***GAA*ATGTTCTTGCAAAGAAGCCAGATA
CzWS1$_{44\text{-}400}$-F (to pYES2.1)	GCAGA**GCGGCCGC***GAA*ATGTCAAGGATAGCACCTGGCTG
CzWS1$_{71\text{-}400}$-F (to pYES2.1)	GCAGA**GCGGCCGC***GAA*ATGCATATGAGAGAGGAGGTCGTCAG
CzWS1$_{1\text{-}394}$-R (to pYES2.1)	TAT**GTCGAC**GCACCAACCCAGCTGCTAC
CzWS1$_{1\text{-}371}$-R (to pYES2.1)	TAT**GTCGAC**GCTGGTGACTGCTTGAGCTA
CzWS1-F (to MYTH vector)	AC**GGCCATTACGGCC**ATGGAAGCAGCACTGACCA
CzWS1-R (to MYTH vector)	TAT**GGCCGAGGCGGCC**TCAAATGGCTGACTGTTGCACC
CzWS1-R (to pGreen vector)	CGC**TCTAGA**TCAAATGGCTGACTGTTGCACC
CzWS1-F (to pOpt2_mVenus_Paro vector)	TGCAGGATGCATATGGGATCCATGGAAGCAGCACTGACCACTG
CzWS1-R (to pOpt2_mVenus_Paro vector)	GCCCTCGATGACGTCAGATCTTCAAATGGCTGACTGTTGCACC
CzWS1-QF (for qRT-PCR)	GGTTGGTGAAGAAGGCCAAA
CzWS1-QR (for qRT-PCR)	CATGGATCAGCCCACTCAATG
CBLP-QF (for qRT-PCR)	CTTCTCGCCCATGACCAC
CBLP-QR (for qRT-PCR)	CCCACCAGGTTGTTCTTCAG

Supplemental Table S3. Predicted transmembrane domain (TMD) and positive selection sites of CzWS1. TMD was predicted by Phobius. Positive selection sites with a Bayes Empirical Bayes posterior probability ≥ 99%, or ≥ 95%, are shown in **bold** or normal font, respectively.

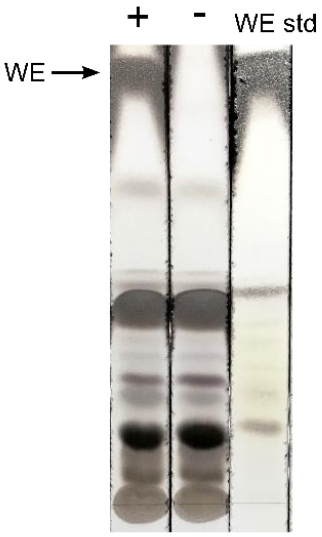
Region	Amino acid	Length (Amino Acid)	Predicted positive selection sites	Number of positive selection sites	Ratio of positive selection site (%)
N-terminus (Inside)	1-26	26	V9, D10	2	7.69
TMD1	27-47	21	K40	1	
Outside	48-52	5		0	
TMD2	53-70	18	**L66**	1	
Inside	71-81	11		0	
TMD3	82-99	18	**L85, I86, S87**	3	
Outside	100-104	5		0	
TMD4	105-122	18	W108, **V110, M113**, A119	4	
Inside	123-142	20		0	
TMD5	143-164	22	L163, L164	2	6.79
Outside	165-175	11	K174	1	
TMD6	176-201	26	**Y183, G184, S187**	3	
Inside	202-270	69	R246	1	
TMD7	271-292	22		0	
Outside	293-297	5	R294, Y296	2	
TMD8	298-318	21	**P298**	1	
Inside	319-329	11	K320	1	
TMD9	330-350	21	T336, **F349**	2	
C-terminus (Outside)	351-399	49	P362, **N371**, F372, Q373, L376, G381	6	12.24


Name
p-value
Motif Locations
est_cluster_3200678 1.50e-26
Cre02.g098050 5.15e-20
Vocar.0012s0083 2.98e-17
Cre07.g349966 3.17e-16
Cre07.g350050 2.22e-27
Vocar.0023s0187 4.82e-26
Vocar.0023s0188 1.02e-17
Bobra.0028s0004.1 1.72e-17
Bobra.0028s0013.1 4.65e-19
Bobra.0028s0006.1 1.44e-17
Bobra.0122s0025.1 2.38e-27
Dusal.0428s00011 4.60e-24
CzWS1 1.78e-26
CzWS3 9.93e-11
Dusal.0023s00016 3.99e-22
CzWS2 5.11e-17
Cre06.g308700 1.41e-11
Vocar.0018s0111 4.86e-21
Cre16.g685300 7.46e-21
447564 1.19e-68
Cre07.g349650 4.45e-18
Cre07.g349900 2.23e-30
Vocar.0023s0194 4.41e-31
Vocar.0023s0195 2.99e-27
Green Algae
Sphfalx0014s0267 1.00e-52
Sphfalx0128s0011 1.24e-50
Pp3c3_10310 7.45e-50
Sphfalx0033s0082 4.28e-53
Pp3c10_23080 1.81e-42
Mapoly0052s0056 2.97e-48
Bryophytes/
Lycophyte
Bradi5g13630 9.80e-127
Brast09G121300 3.32e-127
GRMZM2G111423 9.27e-113
Sobic.006G123200 2.21e-109
Pavir.Ga01372 1.38e-112
Pavir.Gb01323 8.98e-114
Seita.7G140400 7.69e-123
Sevir.7G148300 8.43e-122
LOC_Os04g40580 7.93e-99
LOC_Os04g40590 3.43e-110
Bradi4g23490 1.86e-125
Brast05G037500 7.52e-122
GRMZM2G171440 1.47e-111
Sobic.001G527200 1.70e-115
Pahal.D02608 1.05e-123
Seita.1G150500 1.60e-116
Sevir.1G148300 1.60e-116
LOC_Os02g25700 3.30e-121
Bradi5g09000 4.69e-111
Brast09G068600 2.93e-105
Monocots
29812.t000006 4.94e-119
29812.t000007 1.14e-170
Manes.07G059500 1.26e-176
Manes.07G059600 1.61e-170
Potri.016G014600 9.82e-159
SapurV1A.0398s0170 6.71e-138
Gorai.010G186500 1.54e-169
Thecc1EG030359 2.37e-167
Manes.10G134500 3.20e-185
Potri.009G054900 1.83e-153
29990.t000016 8.08e-146
Manes.10G134600 2.97e-167
Potri.009G055100 2.09e-171
SapurV1A.0027s0910 1.09e-175
Gorai.004G067300 2.24e-157
Thecc1EG005810 3.56e-177
Thecc1EG005811 1.41e-168
Ciclev10031885m.g 1.89e-167
evm.TU.supercontig_197.28 2.20e-155
AL5G32480 4.10e-145
AT3G51970 1.54e-144
Carubv10019087m.g 3.48e-147
Bol025010 1.82e-135
Brara.I03520 1.13e-139
Bol025011 8.70e-145
Dicots

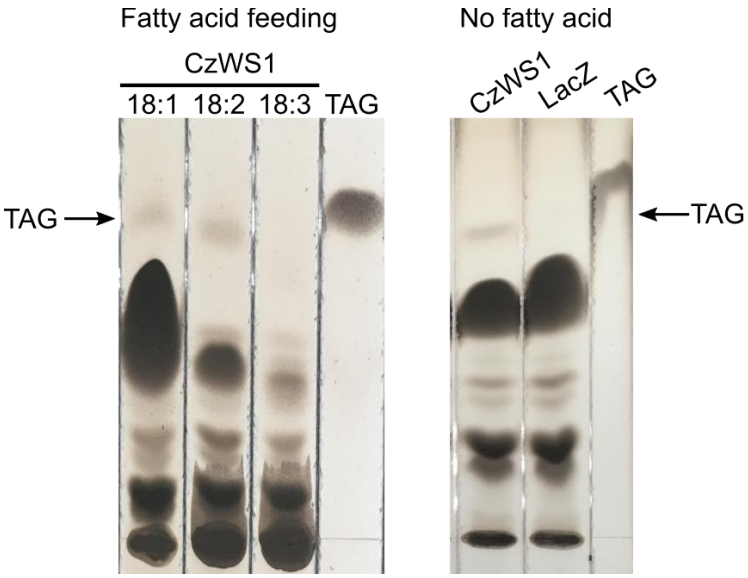
Supplemental Figure S1. Conserved motifs on wax synthases among green plants.



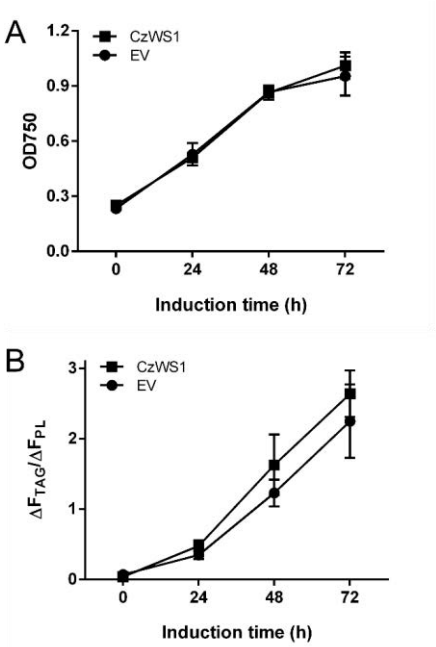
Supplemental Figure S2. **Transcriptional change of *C. zofingiensis* genes encoding WS and diacylglycerol acyltransferase (DGAT) related enzymes in response to high-light stress (data from Roth et al., 2017).**



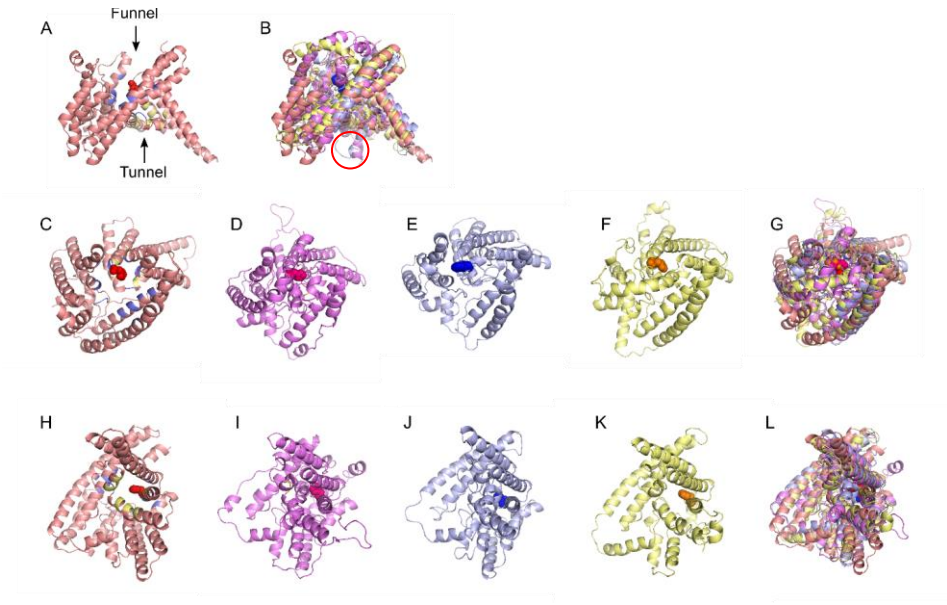
Supplemental Figure S3. TLC analysis of wax ester (WE) from H1246 yeast expressing *CzWS1* cultured in the presence (+) or absence (-) of 16:0-OH.



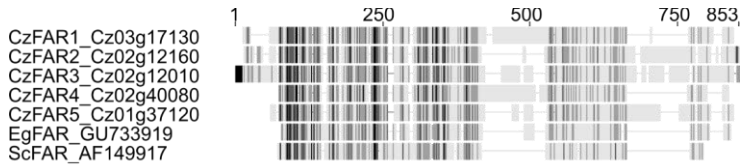
Supplemental Figure S4. TLC analysis of TAG from H1246 yeast expressing *CzWS1* cultured in the presence or absence of various fatty acids.



Supplemental Figure S5. Overexpression of *CzWS1* in *Chlamydomonas reinhardtii*. Growth plots (A) and neutral lipid content (represented by Nile red value; B) of the *C. reinhardtii* lines expressing *CzWS1* and the empty vector cultivated under nitrogen starvation. Nile red value is calculated by dividing Nile red fluorescence of TAG (ΔF_{TAG}) by Nile red fluorescence of polar lipids (ΔF_{PL}). Data represent means ± SD of three biological replicates.

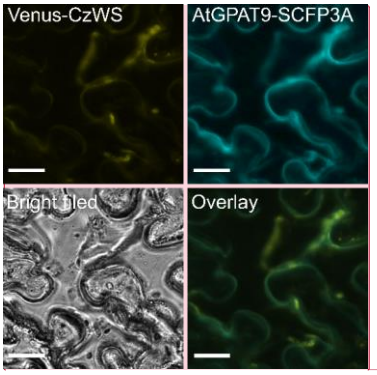


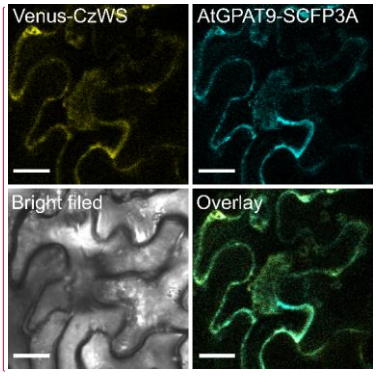
Supplemental Figure S6. Overlay of the predicted three-dimensional structures of bacteria DltB, CzWS1, *Euonymus alatus* diacylglycerol acetyltransferase (EaDAcT), and *Euglena gracilis* WS (EgWS). A. The substrate binding funnel and tunnel of DltB. B. Overlay of the substrate binding funnel and tunnel of DltB, CzWS, EaDAcT and EgWS. The C-termini of CzWS1 and EaDAcT are circled in red. A closer view of the funnel and tunnel of DltB (C, H), CzWS (D, I), EaDAcT (E, J), EgWS (F, K) and overlay (G, L), respectively. Dltb structure, the predicted structures of CzWS, EaDAcT and EgWS, and their corresponding catalytic histidine are shown in red, purple, blue and yellow, respectively. As for A, C, and H, the residues that form the funnel and tunnel in DltB are shown in blue and red, respectively.



	Medium Light (h)						High Light (h)				
	0	0.5	1	3	6	12	0.5	1	3	6	12
CzWS1	25.0	24.8	24.9	23.1	24.2	21.0	61.5	61.8	27.9	33.0	28.6
CzFAR1	90	93	85	71	29	19	116	92	61	25	20
CzFAR2	68	72	74	68	49	38	77	59	68	48	40
CzFAR2	12	9.2	11	12	9.6	10	9.6	11	12	10	10
CzFAR4	65	70	61	49	17	5.6	80	51	37	13	8.7
CzFAR5	26	28	27	15	8.2	6.4	28	20	10	7.9	9.5

Supplemental Figure S7. Identification of putative *fatty acid reductase* (*FAR*) genes in *C. zofingiensis* and their transcriptional changes in response to high-light stress (data from Roth *et al.*, 2017).





Commented [XY1]: This is a previous version which is obtained using fluorescent microscope. I used a different color to show the cyan fluo but didn't change intensity or anything else.

Commented [XY2]: I repeated this using confocal but the intensity is too low and this is the best image I could get. But it looks weird to me.

Supplemental Figure S8. Subcellular localization of CzWS1 in *Nicotiana benthamiana* leaf cells. Venus::CzWS1 was co-localized with a SCFP3A-tagged *Arabidopsis thaliana* glycerol-3-phosphate acyltransferase (AtGPAT9), a known endoplasmic reticulum (ER) localized protein. Scale bars represent 20 μm.

References:

Becker D, Kemper E, Schell J, Masterson R. 1992. New plant binary vectors with selectable markers located proximal to the left T-DNA border. Plant Molecular Biology **20**, 1195–1197.

Gehl C, Waadt R, Kudla J, Mendel RR, Hänsch R. 2009. New GATEWAY vectors for high throughput analyses of protein-protein interactions by bimolecular fluorescence complementation. Molecular Plant **2**, 1051–1058.

Käll L, Krogh A, Sonnhammer ELL. 2007. Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server. Nucleic Acids Research **35**, 429–432.

Roth MS, Cokus SJ, Gallaher SD, *et al.* 2017. Chromosome-level genome assembly and transcriptome of the green alga *Chromochloris zofingiensis* illuminates astaxanthin production. Proceedings of the National Academy of Sciences of the United States of America **114**, E4296–E4305.

Vanhercke T, El Tahchy A, Shrestha P, Zhou X-RR, Singh SP, Petrie JR. 2013. Synergistic effect of *WRI1* and *DGAT1* coexpression on triacylglycerol biosynthesis in plants. FEBS letters **587**, 364–369.

Xu Y, Caldo KMP, Falarz L, Jayawardhane K, Chen G. 2020. Kinetic improvement of an algal diacylglycerol acyltransferase 1 via fusion with an acyl-CoA binding protein. The Plant Journal **102**, 856–871.

Xu Y, Caldo KMP, Jayawardhane K, Ozga JA, Weselake RJ, Chen G. 2019. A transferase interactome that may facilitate channeling of polyunsaturated fatty acid moieties from phosphatidylcholine to triacylglycerol. Journal of Biological Chemistry **294**, 14838–14844.