Diversity and engineering of diatom metabolism for new and improved sterol products

Ana Cristina Jaramillo-Madrid

PhD by research

Submitted in fulfilment of the requirements for

The degree of Doctor of Philosophy

Climate Change Cluster

School of Life Sciences

University of Technology Sydney
2020

Certificate of original authorship

I, Ana Cristina Jaramillo Madrid declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctorate of Philosophy, in the Faculty of Science at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

Signature:

Production Note:
Signature removed prior to publication.

Ana Cristina Jaramillo Madrid
Date: 20 of December 2019

Acknowledgements

Si el Señor no construye la casa, en vano se cansan los constructores

Salmo 126

I would like to express my sincerest gratitude to my supervisors Prof. Peter Ralph and Dr. Justin Ashworth. Peter has been always a kind and supportive mentor, he believed in me from day one and encouraged me to cultivate my skills as a scientist. Thank you for providing and promoting the amazing C3 culture we enjoy today. I want to especially thank Justin for his mentorship; I barely have words to express the immense contributions he has made to the development of my scientific capabilities. He pushed me to learn and perform rigorous science. He always led by example, providing guidance and direction. His passion for science is contagious and helped me to do my job with enthusiasm. Thanks, Justin, for always keeping your door open to give me advice and encouragement. Without your expertise I would not have gotten half as far.

I would also like to thank Raffaela Abriano and Michele Fabris, the godfather of my project, for their thoughtful insights that have greatly expanded the results obtained in this thesis. A special thanks to Raffie for always having a kind and constructive attitude to guide me through the PhD journey.

Thank you to my fellow PhD students for being an amazing group of people to work with. Jestin George and Lorenzo Barolo for making the long hours in the lab more enjoyable and for the lovely dinners during game nights. To the other members of Algae Biosystems and Biotechnology group, especially Mathieu Pernice and Audrey Commault, thanks for your support and feedback, for listening to my struggles and for being there when I needed support.

Thank you to the tech staff at UTS for all the help provided. Thanks to Unnikrishnan Kuzhiumparambil, Taya Lapshina and Kun Xiao for methodological guidance in GC/MS analysis and Flow Cytometry.

A special thank you to Angelina Hormaza, my undergrad and master's supervisor back in Colombia, who always recognised and cultivated my potential. She has been a very close mentor throughout my scientific career and encouraged me to pursue my PhD studies abroad. She is one of the great examples of women in science and an inspiration to me.

To the German, Julian Kofler, for his patience and support throughout this PhD process. You were always there and helped me to keep going during the stressful times. To my parents, Gabriel Jaramillo and Gladis Madrid, for being the people I can always count on for anything. They have constantly sustained and encouraged me to pursue my dreams. They taught me to be hard working and to stay true to myself. Thanks for believing in me and setting a great example for me. A special thanks to my dad for the amazing job editing of this thesis in LaTeX, it certainly looks better than my Word version.

I am very grateful to you all for your input in helping me deliver this thesis and supporting my love of science.

Contents

List of Figures

List of Tables

Abbreviation List

AACCO	Aceto-acetyl-CoA
AACT	Acetoacetyl-CoA thiolase
ACCOA	Acetyl-CoA
AltSQE	Alternative squalene epoxidase enzyme
CDP-ME	4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol
CDP-MEP	2-phospho-4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol
CMK	4-(cytidine 5′-diphospho)-2-Cmethyl-d-erythritol kinase
DMAPP	Dimethylallyl diphosphate
DXP	1-deoxy-D-xylulose-5-phosphate
DXR	1-deoxy-d-xylulose 5-phosphate reductoisomerase
DXS	1-deoxy-dxylulose 5-phosphate synthase
EPA	Eicosapentaenoic acid
ER	Endoplasmic reticulum
FPP	Farnesyl diphosphate
FPPS	Farnesyl diphosphate synthase
GC-MS	Gas chromatography-mass spectrometry
GPP	Geranyl diphosphate
GPPS	Geranyl pyrophosphate synthase
HDR	4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase
HDS	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
HMBPP	1-hydroxy-2-methyl-2butenyl-4-diphosphate
HMG-CoA	3-hydroxy-3-methylglutaryl-CoA
HMGR	3- hydroxy-3-methylglutaryl-CoA reductase
HMGS	Hydroxymethylglutaryl-CoA synthase
HMMs	Hidden Markov Models
IDI	Isopentenyl diphosphate
IDI-SQS	Isopentenyl diphosphate isomerase-squalene synthase
IPK	Isopentenyl phosphate kinase
IPP	Isopentenyl diphosphate
MCT	2-C-methyl-d-erythritol 4-phosphate cytidylyltransferase

MDS	2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase
ME-cPP	2-C-methyl-D-erythritol2,4-cyclodiphosphate
MEP	Methylerithriol phosphate pathway
MMETSP	Marine Microbial Eukaryote Transcriptome Sequencing Project
MVA	Mevalonate pathway
MVAP	Mevalonate phosphate
MVAPP	Mevalonate diphosphate
MVK	Mevalonate kinase
OSC	Oxidosqualene cyclase
SCAP	SREBP cleavage activating protein
Si	Silicon
SQE	Squalene epoxydase
SQS	Squalene synthase
SRE	Sterol Regulatory Element
SREBPs	Sterol Regulatory Element Binding Proteins
SSD	Sterol sensing domain
tHMGR	Truncated 3- hydroxy-3-methylglutaryl-CoA reductase

Thesis Abstract

Diatoms are a large group of eukaryotic microalgae that arose through secondary endosymbiosis and are renowned for their wide ecological distribution. Diatoms have genetically diversified their physiology, metabolism and natural products, while adapting to dynamic environments. Among these metabolic products are an expanded repertoire of phytosterols, a class of essential terpenoids that are involved in the regulation of membrane dynamics, signalling, and membrane-bound protein functions in higher plants, algae, fungi, and vertebrates. Phytosterols are considered a marker of eukaryotic life and have been used to identify and date evolutionary events. They are also useful natural products due to their wide range of biological applications. The principal therapeutic and nutraceutical properties of phytosterols include cholesterol-lowering, anti-inflammatory and anti-diabetic activities. The global phytosterol market by 2013 was US$ 300 million and it is growing at about 7-9% per annum. In order to meet this demand, diatom microalgae are proposed as an alternative source of natural products.

The function, distribution and biosynthesis of sterols is well characterised and conserved in model animal, plant and fungal organisms. However, the biological role and metabolism of the high diversity of sterols produced by diatoms is not well understood. To establish diatoms as a suitable platform for phytosterols production, in this PhD project we provide insight into key aspects of sterol compounds from diatoms: i) The response of sterol levels to changes in environmental conditions, ii) The reconstruction of the sterol biosynthesis pathways of multiple diatom species, and iii) Genetic investigations and engineering of diatoms to alter sterol product profiles.

In Chapter 1, I provide an updated review of the phytosterol repertoire in diatoms, including the biology and regulation of sterol biosynthesis according to the latest primary studies, and new genetic approaches by which the productive metabolisms of these organisms can be further optimised.

In the first data chapter, Chapter 2, I investigated the occurrence of different sterol types in twelve different diatom species, as well as the effect of temperature reduction and changes in salinity on the sterol contents of three model diatoms. In Chapter 3, I experimentally examined the sterol biosynthesis pathways of three divergent diatom species, using empirical biochemical profiling and comparative 'omics. This Chapter experimentally explored hypotheses with regard to what extent the sterol biosynthesis pathways

of three diatom species are conserved, and where each of these has diverged to produce different phytosterols. This study introduces in-depth multi-species analyses in order to compare and contrast the biosynthesis pathways of distantly related species. The results expand our understanding of sterol biosynthesis in diatoms, including a new model for cholesterol synthesis in diatoms.

Finally, in Chapter 4, I implemented and performed genetic engineering technologies to test the extent to which natural sterol levels can be rationally manipulated in the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*. Three different genetic targets were chosen, including i) The overexpression of a rate-limiting enzyme in sterol biosynthesis, HMGR, and ii) the expression of a N-terminal truncated HMGR and introduction of a heterologous squalene epoxidase enzyme from the microalgae *Nannochloropsis oceanica*.

This thesis is structured with one introduction chapter (Chapter 1), currently published as a review, three data Chapters (Chapters 2 to 4), each written in the form of a journal manuscript for peer-review and a conclusion Chapter (Chapter 5). At the time of thesis submission, all chapters, except conclusion chapter, have been either published, are under peer-review, or in final draft for submission.

The overarching aim of this research project was to investigate and optimise the production of bioactive sterols in diatoms for commercial applications. This project first investigated the diversity and differential production of sterols by several diatom species under different growth conditions. Inhibitors of the key enzymes in the sterol metabolic pathway will then be used to identify relevant intermediate compounds in the biosynthesis of sterols. Finally, enzymes responsible for the synthesis of sterol compounds will be genetically targeted for metabolic engineering of the selected diatom species. The specific aims of this project were:

Aim 1: Characterise the sterols produced under different environmental conditions by several diatoms strains.

Objectives:

- Identify the most abundant sterols produced by diatoms growing in enriched medium.

- Determine the sterols produced under different culture conditions.

H_0: Sterol production does not vary according to diatom strain and culture conditions.
H_a: Sterol production varies according to diatom strain and culture conditions.

Aim 2: Identify enzymes and genes putatively involved in the sterol biosynthesis pathway

Objectives:

- Identify target intermediate compounds in the metabolic pathway of sterols using chemical inhibitors.

- Assemble a general sterol biosynthesis pathway to identify genetic targets for the enhanced production of sterol compounds.

H_0: Inhibition of enzymes involved in the sterol metabolic pathway of diatoms does not result in the production of phytosterol intermediate compounds

H_0: Sterol biosynthesis inhibitors will not differently affect the sterol profiles of different diatom species

Aim 3: Genetically engineer diatoms to probe and optimise the production of sterols

Objectives:

- Genetically over-express biosynthetic enzymes to increase production of sterols.

- Genetically up-regulate and/or disrupt native enzymes and/or regulatory genes to alter or increase production of sterols.

H_0: It is not possible to transgenically alter the sterol products or amounts of diatoms; the natural levels are strictly balanced and regulated

H_a: Genetic modification of enzymes participating in the sterol pathway of diatoms leads to alteration of sterol profiles.

In summary, this project addressed the following research questions:

- What is the effect of different growth conditions on the sterols produced by diatoms?

- Which are the principal intermediate compounds in the sterol metabolic pathway, and which enzymes participate in their formation?

- Does genetic engineering or disruption of genes involved in sterol biosynthesis alter the production of phytosterols by diatoms?

Chapter 1

Phytosterol biosynthesis and production by diatoms (Bacillariophyceae)

Author Contributions

A.C.J.M. wrote the manuscript. J.A. advised and assisted in writing. M.F. advised and assisted in writing P.J.R. advised, materially supported and assisted in writing.

Published in Phytochemistry on July 2019 and featured on front cover:

Abstract

Diatoms are abundant unicellular marine photosynthetic algae that have genetically diversified their physiology and metabolism while adapting to numerous environments. The metabolic repertoire of diatoms presents opportunities to characterise the biosynthesis and production of new and potentially valuable microalgal compounds, including sterols. Sterols of plant origin, known as phytosterols, have been studied for health benefits including demonstrated cholesterol-lowering properties. In this review we summarise sterol diversity, the unique metabolic features of sterol biosynthesis in diatoms, and prospects for the extraction of diatom phytosterols in comparison to existing sources. We also review biotechnological efforts to manipulate diatom biosynthesis, including culture conditions and avenues for the rational engineering of metabolism and cellular regulation.

Keywords: diatoms, phytosterols, biosynthesis, secondary metabolism, metabolites.

1.1 Introduction

Sterols are essential membrane components of all eukaryotic organisms, serving roles as regulators of membrane dynamics, signalling molecules, and the modulation of membrane-bound protein functions (Dufourc, 2008). Sterols are a class of terpenoids, a large class of natural products synthesised from isoprene subunits (Fahy et al., 2005). Sterols are triterpenoids with a steroid skeleton structure, tetracyclic cyclopenta-(α)-phenanthrene (Fig. 1.1), that is derived from squalene. These compounds are involved in the regulation of vital processes in higher plants, algae, most fungi and vertebrates (Desmond & Gribaldo, 2009; Hannich et al., 2011).
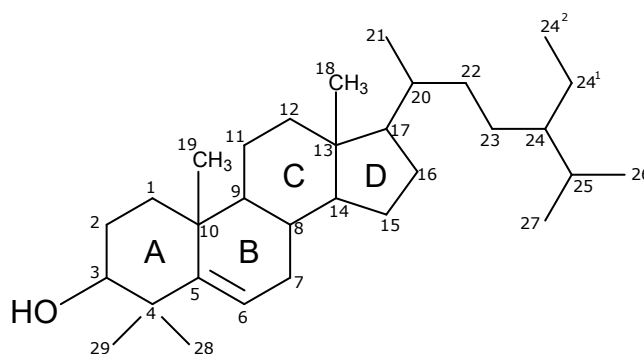
Figure 1.1: Steroid skeleton structure according to IUPAC 1989 recommendations (Moss, 1989). Note: previous guidelines (1976) numbered the atoms 24^1 and 24^2 as 28 and 29. 24^1 and 24^2 can be also found in the literature as 24′ and 24″.

Several sterols of plant origin, known as phytosterols, have been validated or explored for therapeutic and nutraceutical properties that include cholesterol-lowering (Ras et al.,

2014), anti-inflammatory (Aldini et al., 2014) and anti-diabetic activities (Wang et al., 2017). The global phytosterol market by 2013 was US$ 300 million and it is growing at about 7-9% per annum (Borowitzka 2013). In order to meet this demand, microalgae are proposed as an alternative source of natural products (Luo et al., 2015).

Diatoms are eukaryotic microalgae that produce an array of sterol compounds (Stonik & Stonik, 2015), including some in higher diversity than in edible plants (Rampen et al., 2010). Phytosterols typically occur at 0.025% to 0.41% by dry weight of plants material (Piironen et al., 2000). The phytosterol content of the diatom *Phaeodactylum tricornutum* has been reported to be 0.34% dry weight (Ahmed et al., 2015). Diatoms grow readily in culture and do not require organic carbon sources or long growth seasons for cultivation (Fu et al., 2015; Gong et al., 2011). Diatoms have been the focus of research and development for the production of a range of sustainable products including fine chemicals, plastics, biofuels, and biologically active compounds as pharmaceutical, cosmetic and food products (Bozarth et al., 2009; Lebeau & Robert, 2003*a*).

In this review we describe the natural diversity and biosynthesis of sterols by diatoms, as well as prospects for the production of new phytosterol compounds from diatom sources. The topics addressed include: i) an overview of the diversity of sterols present in diatom species and their biosynthesis, ii) prospects for the production of phytosterol compounds from diatoms, and iii) new methods for the optimisation and engineering of diatoms that may be applicable to the investigation and improvement of sterol and related tri-terpenoid compounds in diatoms.

1.2 Sterol biosynthesis and diversity in diatoms

Diatoms (Bacillariophyceae) are unicellular phytoplankton that serve as primary photosynthetic engines in the global carbon cycle. They are present in both seawater and freshwater systems, with a great diversity of habitats, products, shapes, symmetries, and structures (Beyene et al., 2014; Round et al., 1990). Diatoms belong to the Chromalveolate lineage, which originated through secondary endosymbiosis between a red alga, a cyanobacterium and a phagotrophic eukaryote (Armbrust, 2009). Depending on the symmetry of the silica cell wall (frustule), diatom cells are divided into centric (radially symmetric) and pennate (bilaterally symmetric) forms (Round et al., 1990).

The ecological success and diversity of diatom species is attributed to their evolutionary history, genetic diversity, and interaction with other organisms (Moustafa et al., 2009). The nuclear genome of the pennate diatom *P. tricornutum*, which contains 12,233 predicted protein coding genes, was found to possess 584 genes that appear to originate from prokaryotes (Rastogi et al., 2018); more than 300 of these genes were also found in the centric diatom *T. pseudonana* (Armbrust et al., 2004). These two diatoms share only 57% of their genes (Bowler et al., 2008), while the rest bear no similarity to genes of known function.

The gene sets of dozens of additional species follow this trend, highlighting the novelty of diatom and phytoplankton genomes with respect to protein diversity and putative new gene functions (Keeling et al., 2014). Preliminary investigations of sterol metabolism in *P. tricornutum* have revealed unique diatom-specific enzymes that catalyse fundamental biosynthetic reactions (Fabris et al., 2014).

1.2.1 Sterols found in diatoms

The varieties of sterols produced by more than one hundred diatom species have been reported to date, including over 25 different sterols (Rampen et al., 2010) (Table 1.1). The distribution of sterols in microalgae is diverse, and covers a large range of chemical species. Sterols containing 27, 28 and 29 carbon atoms are all found in diatoms. Alkylation at C-24 and a lack of methyl groups at C-4 can also be found in diatom sterols (Volkman, 2003). Double bonds at position 5 and 7 are also common; from which $C_{28}\Delta^5$ is the most abundant. The most common among these sterols were 24-methylcholesta-5,24(24^1)-dien-3β-ol ($C_{28}\Delta^{5,24}$); cholest-5-en-3β-ol (cholesterol, $C_{27}\Delta^5$); 24-methylcholest-5-en-3β-ol (campesterol, $C_{28}\Delta^5$); 24-ethylcholest-5-en-3β-ol (sitosterol, $C_{29}\Delta^5$) and 24-methylcholesta-5,22-dien-3β-ol (brassicasterol, $C_{28}\Delta^{5,22E}$); these sterols were present in 40% of analysed species (Rampen et al., 2010). Additional studies confirm that 24-methylcholesta-5,24(24^1)-dien-3β-ol and 'diatomsterol' (a C-24 epimer of brassicasterol) are the most abundant sterols found in diatoms (Volkman, 2003).

Many of the sterols in diatoms are present in free forms. However, sterols can also exist as conjugates, wherein the hydroxyl group in ring "A" is covalently bound to fatty acids (esterified) or sugars (glycosylated) (Moreau et al., 2002). Most studies of sterols employ extraction techniques that convert steryl esters to their free forms (Volkman, 2016), however P. tricornutum produced significant amounts (40-50%) of conjugated sterols, including steryl esters, steryl glycosides, and acylated steryl glycosides (Véron et al., 1996). *Haslea ostrearia* was also shown to synthesise large amounts of steryl glycosides (Véron et al., 1998). The marine diatom *Skeletonema marinoi*, often present in blooms in temperate coastal waters, produces sterol sulphates as cellular signals to induce cell death (Gallo et al., 2017). Varying distributions of phytosterol types are presumably defined by species-specific and conditional variations in biosynthetic pathways and metabolic regulation.

Table 1.1: The occurrence of sterols in diatom species. Reported biological activities are noted.

Sterol common name and other nomenclatures	Diatom sources	Bio-functionality
Brassicasterol Brassicasterin (22E)-Ergosta-5,22-dien-3β-ol 24-Methylcholesta-5,22-dien-3β-ol ($\Delta^{5,22E}$) (24β=24R) CAS # 474-67-9 $C_{28}H_{46}O$	*Amphiprora alata (17), Achnanthes brevipes (72), Achnanthes sp. (76), Achnathes cf. longipes (85), Amphiprora paludosa (1), Arcocellulus mammifer (8), Biddulphia sp. (32), Brockmanniella brockmannii (11), Cymatosira belgica (3), Delphineis sp. (62), Dickieia ulvacea (97), Ditylum brightwellii (2), Entomoneis cf. alata (27), Extubocellulus cribiger (12), Extubocellus spinifer (13), Fragilaria pinnata (21), Grammatophora oceanica (61), Helicotheca thamensis T (32), Hyalosira sp. (67), Leynella arenaria (3), Lithodesmium undulatum (4), Minutocellus cf. sp. (2), Minutocellus polymorphus (5), Nanofrustulum shiloi (11), Odontella aurita (14), Odontella longicruris (42), Papiliocellulus sp. (5), Pauliella taeniata (86), Plagiogrammopsis vanheurckii (21), Stauroneis constricta (100), Stauroneis simulans (1), Synedra fragilaroides (85), Synedra hyperborea (21), Synedropsis cf. recta (10), Talaroneis sp. (88), Thalassionema sp. (1), Thalassiosira stellaris (tr)*	Atherosclerosis prevention (Izar et al., 2011) Hypocholesterolemic (de Jesus Raposo et al., 2013)
Campesterol Campesterin (24R)-Ergost-5-en-3β-ol Δ^5- 24α-Methyl-cholesten-3β-ol (24R)-24-Methylcholest-5-en-3β-ol Campest-5-en-3β-ol (Δ^5) (24α=24R) CAS # 474-62-4 $C_{28}H_{48}O$	*Achnanthes brevipes (8), Achnanthes sp. (3), Achnathes cf. longipes (6), Amphipora hyaline (0.3), Amphiprora alata (8), Amphora sp. (47), Arcocellulus mammifer (1), Attheya ussurensis (1), Attheya longicornis (44), Attheya septentrionalis (17), Attheya septentrionalis (19), Attheya septentrionalis (36), Aulacoseira granulate var. Angustissima (61), Bacteriastrum hyalinum (29), Biddulphia sp. (7), Brockmanniella brockmannii (2), Coscinodiscus granii (17), (44.4), Coscinodiscus sp. (6), Cyclotella cryptica (7), Delphineis sp. (1), (1.6), Dickieia ulvacea (2), Entomoneis cf. Alata (14), Extubocellulus cribiger (1), Extubocellulus cribiger (2), Fragilaria pinnata (2.4), Fragilaria striatula (5), Halassiosira punctigera (18), Hyalodiscus sp. (13), Hyalodiscus stelliger (10), Hyalosira sp. (25), Leynella arenaria (3), Melosira cf. Octogona (9), Minidiscus trioculatus (41), Minutocellus cf. Sp. (tr), Minutocellus polymorphus (1), Navicula phyllepta (11), Navicula sp.(9), Nitzschia thermalis (46), Odontella aurita (9), Odontella longicruris (6), Papiliocellulus sp. (3), Paralia sp. (14), Paralia sulcata (9), Pauliella taeniata (10), Phaeodactylum tricornutum (0.5-1) Plagiogrammopsis vanheurckii (1), Proboscia eumorphis (3), Skeletonema costatum (28.2), Skeletonema sp. (13.6), Skeletonema subsalsum (5), Stellarima microtrias (10), Stephanopyxis palmeriana (5), Stephanopyxix turris (8), Synedra fragilaroides (2), Synedra hyperborea (9), Synedropsis cf. Recta (8), Thalassiosira pseudonana (6), Thalassiosira rotula (32.9), Thalassiosira stellaris (16.3), Thalassiosira weissflogii (0.7-11)*	Hypocholesterolemic (De Jesus Raposo et al. 2013) Protective in cancer models (Shahzad et al., 2017)
Fucosterol [24(24^1)E]-Stigmasta-5,24(24^1)-dien-3β-ol [24(24′)E]-Stigmasta-5,24(24′)-dien-3β-ol [24(28)E]-Stigmasta-5,24(28)-dien-3β-ol 24E-ethylidenecholesta-5,24(28)-dien-3β-ol ($\Delta^{5,24E}$) CAS # 17605-67-3 $C_{29}H_{48}O$	*Attheya longicornis (5), Attheya septentrionalis (1), Attheya septentrionalis (1), Bacteriastrum hyalinum (4), Chaetoceros calcitrans (2), Chaetoceros muelleri (39), Chaetoceros sp. (10), Corethron criophylum (45), Coscinodiscus sp. (tr), (4.7), Cyclotella cryptica (4), (5), Delphineis brightwellii (15.3), Detonula confervacea (5), (4), Ditylum brightwellii (26), Eucampia antarctica (86), halassiosira punctigera (1), Helicotheca thamensis T (13), Minidiscus trioculatus (1), Skeletonema costatum (1), (2), Skeletonema sp. (4), Skeletonema subsalsum (tr), Thalassiosira aff. Antarctica (3), Thalassiosira gravida (4), Thalassiosira gravida (5), Thalassiosira pseudonana (5), (9), Thalassiosira stellaris (6.5), Thalassiosira weissflogii (0.3-14)*	Protective in diabetes models, anti-oxidant (Lee et al. 2004) anti-inflammatory (Aldini et al. 2014)

Table1.1 ...continued

Sterol common name and other nomenclatures	Diatom sources	Bio-functionality
Isofucosterol 5-Avenasterol 28-isofucosterol 29-isofucosterol [24(24^1)Z]-Stigmasta-5,24(24^1)-dien-3β-ol [24(24')Z]-Stigmasta-5,24(24')-dien-3β-ol 24Z-Ethylidenecholesta-5,24(28)-dien-3β-ol [24(28)Z]-Stigmasta-5,24(28)-dien-3β-ol ($\Delta^{5,24Z}$) CAS # 18472-36-1 C$_{29}$H$_{48}$O	*Attheya septentrionalis (10), Attheya septentrionalis (9), Haslea sp. (10), Amphipora hyaline (0.9), Amphora sp. (0.6), Asterionellopsis glacialis (42), Attheya longicornis (5), Attheya septentrionalis (10), Attheya septentrionalis (tr), Chaetoceros calcitrans (6), Chaetoceros muelleri (3), Chaetoceros socialis (8), Chaetoceros sp. (4), Corethron criophylum (6), Coscinodiscus sp. (tr), Cyclotella cryptica (tr), (2), Cylindrotheca closterium (3), Cylindrotheca fusiformis (7.6), Detonula confervacea (3), (3), Ditylum brightwellii (13), Halassiosira punctigera (1), Haslea ostrearia (4.6), Helicotheca thamensis T (28), Hyalosira sp. (1), Lithodesmium undulatum (29), Minidiscus trioculatus (8), Navicula pelliculosa (1), Nitzschia closterium (2), Papiliocellulus sp. (tr), Phaeodactylum tricornutum (2.5), (1), Skeletonema costatum (1), (4.1), Skeletonema sp. (2.5) Skeletonema subsalsum (11), Thalassiosira pseudonana (4), (4), Thalassiosira stellaris (2.2), Thalassiosira weissflogii (0.2-7), Toxarium sp. (17)*	Protective diabetes models, hypocholesterolemic (Singh et al., 2013)
Sitosterol β-sitosterol 24-Ethylcholesterol Stigmast-5-en-3β-ol 24α-ethylcholest-5-en-3β-ol (Δ^5) (24α=24R) CAS # 83-46-5 C$_{29}$H$_{50}$O	*Amphipora hyaline (59.5), Amphora sp. (3.9), Haslea ostrearia (95.4), Phaeodactylum tricornutum (1-3.5), Ske letonema costatum (3.5), Skeletonema sp. (2.7), Thalassiosira rotula (11.3), Thalassiosira stellaris (0.9)*	Activity in models of immune function (Alappat et al., 2010) Hypocholesterolemic (De Jesus Raposo et al. 2013)
Stigmasterol Stigmasterin (22E)-Stigmasta-5,22-dien-3β-ol 24α-Ethylcho-lesta-5,22E-dien-3β-ol ($\Delta^{5,22E}$)(24α=24S) CAS # 83-48-7. C$_{29}$H$_{48}$O	*Amphipora hyaline (33.9), Amphiprora alata (41), Amphiprora paludosa (89), Amphora coffeaeformis (96), (88), Amphora sp. (95.5), Coscinodiscus sp. (tr), Delphineis brightwellii (0.8), Entomoneis cf. alata (55), Eucampia antarctica (1), Helicotheca thamensis T (4), Hyalosira sp. (tr), Lithodesmium undulatum (1), Paralia sp. (39), Paralia sulcata (26), Phaeodactylum tricornutum (30.5-7), Proboscia alata (2), Stephanopyxis palmeriana (10), Stephanopyxix turris (24), Surirella sp. (50), Synedra fragilaroides (2)*	anti-osteoarthritic properties (Gabay et al., 2010)

Numbers in parentheses represent the concentration of the individual sterol, as a percentage of the total sterols. "tr" (trace) indicates relative abundances of < 0.5% (Barrett et al., 1995; Breteler et al., 2005; Gladu et al., 1991; Giner & Wikfors, 2011; Ponomarenko et al., 2004; Rampen et al., 2010; Véron et al., 1996)

1.2.2 Biosynthesis of sterols in diatoms

Despite extensive chemical profiling of sterol species, models of sterol biosynthesis in diatoms are incomplete. The core reactions involved in isoprenoid and sterol biosynthesis are typically conserved across different taxonomic groups, while downstream or diversifying reactions remain uncharted (Lohr et al., 2012). Biosynthesis of sterols in diatoms is likely to have diversified during environmental specialization and gene transfer (Chan et al., 2012). The number of diatom species is estimated to be at least 30,000 (Mann & Vanormelingen, 2013). Thus far, only the sterol biosynthesis pathway of *P. tricornutum* has been empirically reconstructed (Fabris et al. 2014). This pathway contains features of both plant and fungal pathways, utilizes a multi-functional isopentenyl diphosphate isomerase/squalene synthase (IDISQS) enzyme, and an alternative squalene epoxidase enzyme (AltSQE) found to be widespread in eukaryotes (Pollier et al., 2019).

A model of sterol biosynthesis in diatoms is shown in Fig. 1.2. The first stage in sterol biosynthesis (Fig. 1.2) is the formation of the isomeric compounds isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), the building blocks of all terpenoids. These precursors can be synthesised through either the mevalonate (MVA) pathway or the methylerithriol phosphate (MEP) pathway (Fig. 1.2). Both pathways are simultaneously functional in plants, but the MVA pathway is missing in some green and red algae (Lohr et al., 2012). Massé et al., 2004 found that the diatom *Rhizosolenia setigera* produces sterols by the MVA pathway, while the diatom *Haslea ostrearia* biosynthesizes these compounds via the MEP pathway. In organisms that harbour both the MVA and MEP pathways, the first generally provides the substrates for the sterol biosynthesis. In diatoms The MVA and MEP pathways may both be present and capable of producing isoprenoid precursors for terpenoid biosynthesis (Desmond and Gribaldo 2009; Sasso et al., 2012).

In the MVA pathway, three molecules of acetyl-CoA are converted into an IPP molecule, which is then converted into DMAPP by an IPP isomerase (IDI) (Lichtenthaler, 2000). In plants, the enzymes catalysing the MVA pathway are usually located between cytosol and peroxisomes, and provide precursors for brassinosteroids, polyprenyls and triterpenoid secondary metabolites. In contrast, the MEP pathway produces both IPP and DMAPP from pyruvate and glyceraldehyde-3-phosphate (Fig. 1.2A), and is localised to the plastids. This pathway is known to produce precursors for plastidic isoprenoids, such as carotenoids and phytol.

The isoprenoid building blocks IPP and DMAPP are then converted into geranyl diphosphate (GPP) and farnesyl diphosphate (FPP) in multiple steps via a geranyl pyrophosphate synthase (GPPS). FPP is then further condensed into the thirty-carbon compound squalene, which is the ultimate precursor to sterols (Fig. 1.2B). The way in which these condensation reactions occur in diatoms appears to be different from that of animals, plants or fungi, where FPP is normally converted into squalene by a squalene synthase (SQS). In diatoms this enzymatic function is fused with an IPP isomerase (Athanasakoglou et al., 2019; Fabris et al. 2014). Subsequently, squalene is converted into 2,3 epoxysqualene. In mammals, the epoxidation of squalene is the first committed and rate-limiting step in the pathway (Gill et al., 2011). In diatoms, this reaction is catalysed by an alternative squalene epoxidase (AltSQE), which belongs to the fatty acid hydroxylase superfamily and differs from the conventional flavoprotein SQE used by plants, animals and fungi (Pollier et al., 2019).

2,3 epoxysqualene is then cyclised by the enzyme oxidosqualene cyclase (OSC) producing either lanosterol or cycloartenol (Fig. 1.2C). Lanosterol is commonly produced in animals, fungi and some algae, while cycloartenol is produced in plants and most algae. It remains unclear which diatoms biosynthesize sterols via cycloartenol versus lanosterol (Rampen et al., 2010).
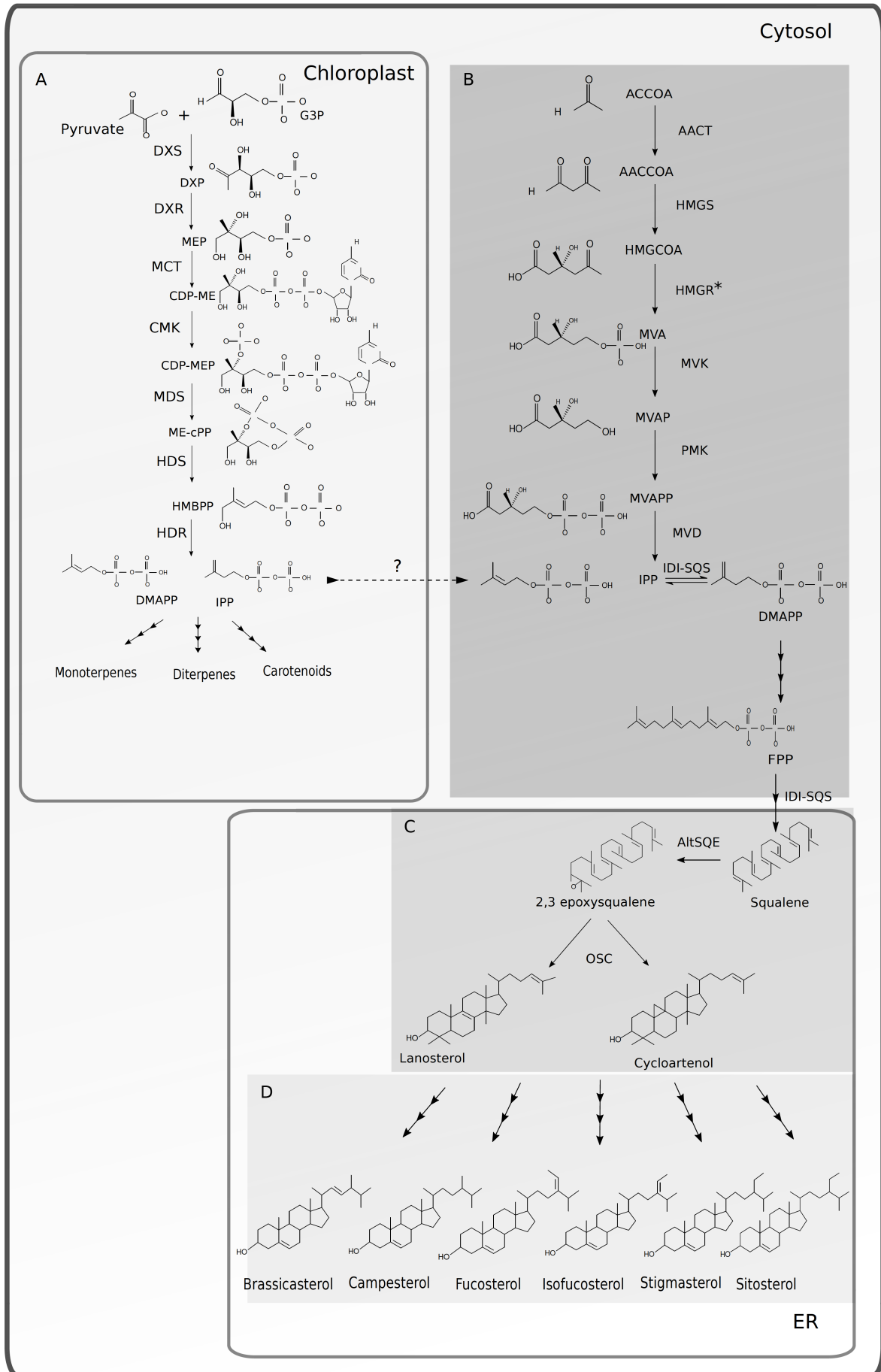
Figure 1.2: Overview of the MVA and MEP pathway and hypothetical early conserved steps of sterol biosynthesis in diatoms. (A) 2-C-methyl-D-erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate (MEP) pathway. DXP, 1-deoxy-D-xylulose-5-phosphate; MEP, 2-C-methyl-D-erythritol-4-phosphate; CDP-ME, 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol; CDP-MEP, 2-phospho-4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol; ME-cPP, 2-C-methyl-D-erythritol-2,4-cyclodiphosphate; HMBPP, 1-hydroxy-2-methyl-2butenyl-4-diphosphate. DXS, 1-deoxy-d-xylulose 5-phosphate synthase; DXR, 1-deoxy-d-xylulose 5-phosphate reductoisomerase; MCT, 2-C-methyl-d-erythritol 4-phosphate cytidylyltransferase; CMK, 4-(cytidine 5′-diphospho)-2-C-methyl-d-erythritol kinase; MDS, 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase; HDS, 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase; HDR, 4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase. Dashed line indicates putative crosstalk between MEP and MVA pathway (B) Mevalonate pathway (MVA) and upstream reactions in the sterol biosynthesis. ACCOA, acetyl-CoA; AACCOA, aceto-acetyl-CoA; HMGCOA, 3-hydroxy-3-methylglutaryl-coenzyme A; MVA, mevalonate; MVAP, mevalonate phosphate; MVAPP, mevalonate diphosphate; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPP, geranyl diphosphate; FPP, farnesyl diphosphate. The MVA pathway starts with two molecules of acetyl-CoA that are condensed to form acetoacetyl-CoA (ACCOA) by an acetoacetyl-CoA thiolase (AACT). Subsequently, an aldol condensation reaction takes place between acetoacetyl-CoA and a third molecule of acetyl-CoA to form 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) by a hydroxymethylglutaryl-CoA synthase (HMGS). The next reaction has been identified as the major rate controlling step in plants (Rodríguez-Concepción, 2006), in which the HMG-CoA is converted to mevalonate by the 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR). Next, two phosphorylation reactions (MVK, Mevalonate kinase; PMK, phosphomevalonate kinase) followed by a decarboxylation (MVD, mevalonate diphosphate decarboxylase) gives rise to IPP which is converted to DMAPP by an isopentenyl diphosphate:dimethylallyl diphosphate isomerase (IDI). After the formation of IPP and DMAPP, the hypothetical core sterol biosynthesis begins. HMGR* is a membrane protein, anchored in the endoplasmic reticulum; all the other six enzymes involved in the MVA pathway are soluble proteins (Lohr et al. 2012). GPPS, Geranyl pyrophosphate synthase; FPPS, farnesyl diphosphate synthase; SQS, squalene synthase; SQE, squalene epoxydase; OSC, oxidosqualene cyclase. (C) core sterol biosynthesis located in the endoplasmic reticulum (ER) (D) specialised portion of sterol biosynthesis.

In *P. tricornutum*, cycloartenol is produced from 2,3-oxidosqualene by an oxidosqualene cyclase (Fabris et al. 2014). Cycloartenol is then converted into obtusifoliol, and subsequently a C-14 demethylation reaction occurs. This reaction is performed by a 14-α demethylase that is conserved among eukaryotes and also present in diatoms. In the organisms that synthesize sterols via lanosterol, the C-14 demethylation reaction occurs immediately after the 2,3-oxidosqualene is cyclised. Additional enzymes such as C24-sterol reductases, C24-methyltransferases, 4-α methyl oxidases, members of the superfamily Cytochromes P450 may catalyse further reactions to generate diverse sterols (Fig. 1.2D), but these steps remain uncharacterised.

1.2.3 Regulation of sterol biosynthesis

As sterols are essential for the proper functioning of cellular membranes, it is vital for the cells to regulate the amounts and identities of the sterol compounds that they produce. The mechanisms of sterol homeostasis have been studied in mammals and yeast (Howe et al., 2016; Wollam & Antebi, 2011), but the mechanisms by which microalgae sense the levels of sterols and transmit signals to control them is not well understood in comparison to other species (Espenshade and Hughes 2007).

Table 1.2 summarises the principal enzymes involved in sterol sensing and regulation in mammals, and their homologues in other groups of organisms. In mammalian systems, a principal regulatory mechanism involved in sterol regulation is the sensing of cholesterol. The Sterol Regulatory Element Binding Proteins (SREBPs) are proteolytically cleaved from the endoplasmic reticulum when cholesterol levels are low; they then migrate to the nucleus and activate target genes containing Sterol Regulatory Element (SRE) sequences in their promoters (Espenshade and Hughes 2007). SREBPs are normally anchored to the ER and bind to the SREBP cleavage activating protein, SCAP. SCAP possesses a sterol-sensing domain (SSD) enabling it to sense cholesterol and prevent SRBP release when cells are replete of sterol molecules (Fig. 1.3) (Espenshade and Hughes 2007). Enzymes involved in cholesterol metabolism, such as 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (HMGR), contain SRE sequences in their gene promoters, which confer regulatory sensitivity according to cholesterol levels in the cell (Espenshade and Hughes 2007). HMGR is a highly conserved enzyme that catalyses a rate-limiting step in the MVA pathway and therefore is a key enzyme in the maintenance of sterols homeostasis (Burg & Espenshade, 2011).

Table 1.2: Principal protein domains known to be involved in the mechanism of sterol homeostasis in mammals, and the presence of putative homologues for these domains in other model organisms.

	Organism	SREBP(Sterol Regulatory Element Binding Protein)	SCAP(SREBP cleavage activating protein)	Insig(insulin induced gene)	S1P(Site-1 protease)	S2P(Site-2 protease)	Reference
Mammal	mice	+	+	+	+	+	(Espenshade and Hughes 2007)
Fungi	*S. pombe*	+	+	+	–	–	(Hughes et al., 2005)
	S. cerevisiae	–	–	–	–	–	(Vik & Rine, 2001)
	C. neoformans	+	+	–	–	+	(Chang et al., 2007; Chun et al. 2007)
Animals	*D. melanogaster*	+	+	–	+	+	(Rosenfeld & Osborne, 1998; Seegmiller et al., 2002)
	C. elegans	+	+	–	–	+	(Espenshade and Hughes 2007)
Plants	*A. thaliana*	–	–	–	+	+	(Vriet et al., 2013)

In fungi, the transcriptional control of sterol synthesis has been reported to be similar to that of mammals. Fission yeasts contain homologs of SREBP and SCAP that are able to sense the fungal sterol ergosterol (Burg and Espenshade 2011; Hughes et al., 2005), and
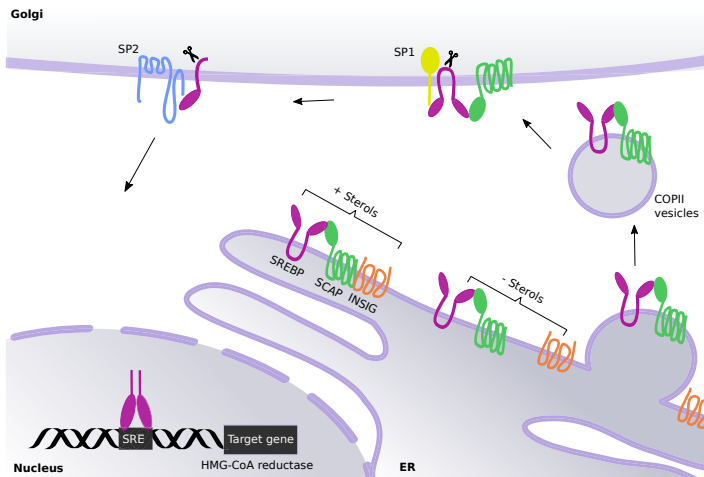
Figure 1.3: The SREBP pathway in mammals from (Espenshade & Hughes, 2007). When the levels of cholesterol are high, Insig binds SCAP-SREBP in the Endoplasmic Reticulum (ER). In sterol depleted cells, INSIG unbinds SEBPR-SCAP allowing them to enter in COPII vesicles. The vesicle is transported to the Golgi where two sequential proteolytic cleavage events take place: SREBP is released by Site-1 (S1P) and Site-2 (S2P) proteases. SREBP travels to the nucleus to activate target genes involved in sterol biosynthesis (HMG-CoA reductase) by binding to sterol regulatory element (SRE) sequences in gene promoters.

the biosynthesis of ergosterol is sensitive to oxygen availability (Todd et al., 2006). In *S. cerevisiae*, the Zn2/Cys6 transcriptional regulators Upc2p and Ecm22p were found to serve as SREBPs, differing from the helix-loop-helix SREBP transcription factors in mammals (Burg and Espenshade 2011; Vik & Rine, 2001).

SREBP homologues are not found in plants (Hannich et al. 2011), but mechanisms for the regulation of HMGR in plants have been proposed. The gene *suppressor of dry2 defects1 (SUD1)* encodes a putative E2 ubiquitin ligase that appears to regulate HMGR activity without apparent changes in protein content (Doblas et al., 2013). Moreover, a phosphatase 2A is reported to control levels of HMGR (Leivar et al., 2011). Additionally, it was found that a type RING-membrane anchor (RMA)-like E3 ubiquitin ligase is expressed to control the activity of HMGR. The expression of the E3 ubiquitin ligase is directly proportional to the accumulation of triterpene saponins, secondary metabolites produced in plants via the MVA pathway (Pollier et al., 2013).

An HMGR enzyme is present in the diatoms *P. tricornutum* and *T. pseudonana*, and it is plausible that its regulation affects secondary metabolism in a similar manner to that of plants. The introduction a non-native oxidosqualene cyclase from *Lotus japonicus* into the diatom *P. tricornutum* re-directed the flux of 2,3 epoxysqualene to the production of other triterpenoids, such as lupeol; as consequence, the MVA pathway and HMGR were upregulated while the total production of sterols was significantly decreased (D'Adamo et al. 2018). Enzymes like C24-sterol reductases, which catalyses the last steps in the synthesis of sterols, were also upregulated. This study suggests that the diatom *P. tricornutum* possesses a regulation system that is sensitive to sterol levels in the cell.

Whether these regulatory mechanisms similarly operate in diatoms and other microalgae is not yet known, since few if any conventional sterol regulatory proteins or SRE *cis-*

regulatory signals have been identified in diatom genomes. It is possible that microalgae regulate their sterols systems differently; for example in the eustigmatophyte *Nannochloropsis oceanica*, It has been proposed that sterol and fatty acid homeostasis relies on sterol-sensitive feedback regulation by 1-deoxy-D-xylulose 5-phosphate synthase (DXS). DXS is the first committed enzyme in the MEP pathway, which is the only mechanism present in *N. oceanica* for the production of isoprenoids and sterols (Lu et al., 2014). Similar investigations remain to be done with regard to the sensing and regulation of sterol biosynthesis in diatoms.

1.3 Diatoms as a source of phytosterols

1.3.1 Diatoms as a microbial platform

Diatoms are adaptable to environmental challenges encountered in dynamic and competitive marine environments. Under fluctuating light conditions, 40% of the photosynthetic electrons were converted into biomass by the diatom *P. tricornutum*, while only 14% were converted into biomass by the green alga *Chlorella vulgaris* (Wagner et al., 2006). Various diatoms can survive under large ranges in pH (pH 3.3–10.6), temperature (4–65°C) (Hildebrand et al., 2012) and salinity (0–50+ g/L). Among thirty-four benthic diatom isolates, four species where able to grow in salinities ranging from 0.5 to 15% (Clavero et al. 2000). Additionally, diatoms can succeed under a highly dynamic system of nitrogen limitation, as well as endure high mixing rates and turbulence. Mechanisms of homeostasis and the storage of nutrients in vacuoles allow diatoms to endure extended phases of limitation in culture (Tozzi et al., 2004).

Diatoms can be more photosynthetically efficient than higher plants in the conversion of light into biomass (Hildebrand et al. 2012), and have evolved concentrating mechanisms that actively assimilate nutrients, CO_2 and bicarbonate from their environment (Mitchell et al., 2013; Shen et al., 2017). These features make them robust and efficient converters of solar energy into biomass and bioproducts under ideal conditions (Wang & Seibert, 2017), and it is possible to improve properties such as photosynthetic efficiency through cellular engineering (Fu et al., 2017).

Diatoms are currently in use for the production of biomass and bioproducts in several applications: the production of supplements for aquaculture (Lebeau & Robert, 2003*a*; López-Elías et al., 2005), eicosapentaenoic (EPA) acid production (Tanaka et al., 2017) and biocrude extraction (Wang & Seibert, 2017). Despite the discovery of many high-value compounds produced by diatoms, such as lipids, oxylipins, sterols, pigments, halogenated compounds isoprenoid and other hydrocarbons (Stonik & Stonik, 2015), their production is still limited to the laboratory scale.

The most commonly used production methods for diatom cultivation include: open

ponds; bags; tubular bioreactors; perfusion cell bleeding; and helical tubular photo-bioreactors (Lebeau & Robert, 2003*b*). Several open pond systems for the production of diatom biomass have also been successfully implemented. In Hawaii and China, outdoor ponds (with CO_2 addition) have high growth rates (Wang & Seibert, 2017). Reported yields vary from 75 to 420 MT $ha^{-1}year^{-1}$ of biomass (Huntley et al., 2015; Wang & Seibert, 2017; Weyer et al., 2010). Recently, Wang & Seibert, 2017, reported a semi-continuous process to grow diatoms with an average yield of 132 MT $ha^{-1}year^{-1}$ of biomass in 5 years.

1.3.2 Biological activities of phytosterols

The effect of phytosterols on human health has been studied for at least sixty years. The Board of Food Standards Australia New Zealand (Federal Register of Legislative Instruments F2015L00440) and The European Atherosclerosis Society Consensus Panel have classified sterols as safe compounds to use in foods, as well as in the administration to patients for the treatment of hypercholesterolemia (Zampelas, 2014).

Table 1.1 lists diatom-produced sterols for which biological activities have been reported. Sitosterol was first used as a supplement to reduce serum cholesterol levels (Pollak, 1953). The hypocholesterolemic activity of phytosterols is based on their ability to inhibit cholesterol absorption in the intestinal lumen. A proposed mechanism of action that the uptake of phytosterols may be competitive or antagonistic in comparison to cholesterol, leading to a reduction of cholesterol absorption (Smet et al., 2012). Clinical trials to evaluate the effects of plant sterols and stanols, a saturated form of sterols, on the lipid profile of patients with hypercholesterolemia resulted in positive effects of these compounds in the reduction of LDL (Athyros et al., 2011; Hallikainen et al., 2008; Párraga-Martínez et al., 2015). The US Food and Drug Administration (FDA) has approved the inclusion of fi − sitosterol, campesterol and stigmasterol in foods to reduce the risk of heart disease (FDA 2005). Edible sterols are present in food products such as spread, cream cheese, yoghurt and milk (Lin et al., 2016).

The reduction of circulating carotenoid levels has been reported as a risk of phytosterols consumption (Kritchevsky & Chen, 2005). A review of the efficacy and safety of phytosterols in the management of blood cholesterol levels reports that levels of α-tocopherol and β-carotene in plasma are reduced in order of 10% to 25% when consuming 1.5 g/d of sterols and stanols (Katan et al., 2003). Clinical trials addressed this problem by adding fruits and vegetables to the diet of subjects consuming plant sterols which effectively maintained proper carotenoid levels (Noakes et al., 2002). Efficacy and secondary effects of phytosterols consumption have been comprehensively reviewed in several studies (Abumweis et al., 2008; Párraga-Martínez et al., 2015).

Additional and secondary effects of phytosterol consumption have been reviewed in multiple studies (Abumweis et al., 2008; Párraga-Martínez et al., 2015), including stud-

ies exploring health-related aspects of sterols with regard to cancer (Shahzad et al., 2017), diabetes (Wang et al., 2017) and inflammation (Aldini et al. 2014). Fucosterol was reported to decrease serum glucose concentration in streptozotocin- and epinephrine-induced diabetic rats (Lee. et al., 2004), and to inhibit enzymes implicated in the pathogenesis of diabetic complications (Jung et al., 2013). Aldini et al. evaluated the effect of phytosterols as anti-inflammatories in a murine experimental model of colitis (Aldini et al. 2014). New biological activities reported for sterols, as well as their bioavailability, metabolism, dose and toxicity must all be established prior to consideration as targets for production in microbial systems (Espín et al., 2007; Sauer & Plauth, 2017).

1.3.3 Production of phytosterols from plants

The current source of phytosterols for commercial use is by-products from the tall oil fractions of plant material (Fernandes & Cabral, 2007). Extraction includes chemical refinement, degumming, neutralisation, bleaching and deodorisation. This process is carried out under vacuum conditions (1–8 mbar) and high temperatures (180–270°C). The final distillate contains 0-5% sterol esters and 2-15% free sterols (Fernandes and Cabral 2007). Recent efforts to achieve higher sterol content include new extraction techniques, such as supercritical fluid extraction from plant matrices. The process is efficient, but higher in cost (Uddin et al., 2015). Similar processes may apply for the extraction of sterols from microalgal biomass. The highest content of sterols in vegetable sources occur in cauliflower (0.41% d.w.) and broccoli (0.34% d.w.) (Table 1.3). The sustainability of producing pure sterols from food sources has been questioned (Tenenbaum, 2008), and the engineering of plants for higher sterol yields has been only mildly successful (Table 1.3).

Table 1.3: Sterol and oil contents cited from different sources on a dry weight (d.w.) basis.

Source	Total sterols (% d.w.)	Reference
Phaeodactylum tricornutum	0.34	(Ahmed et al. 2015)
Chaetoceros muelleri	0.2	(Ahmed et al. 2015)
Fruits	0.05–0.3	(Piironen et al., 2003)
Vegetables	0.025–0.41	(Piironen et al., 2003)
Plant in vitro cultures	0.00014–0.04	(Miras-Moreno et al., 2016)
Engineered plant in vitro cultures	0.00037–0.074	(Miras-Moreno et al., 2016)
Source	**Oil content (% d.w.)**	**Reference**
Corn	4.2	(Thompson et al., 1973)
Soybean	19	(Wolf et al., 1982)
Cylindrotheca sp.	16–37	(Chisti, 2007)
Yeast	58–72	(Whiffin et al., 2016)
Nitzschia sp.	45–47	(Whiffin et al., 2016)
Phaeodactylum tricornutum	20–30	(Chisti, 2007)
Thalassiosira pseudonana	21–31	(Brown et al., 1996)
Source	**Oil in liters per hectare**	**Reference**
Microalgae	58,700	(Chisti, 2007)
Oil palm	5950	
Coconut	2689	
Soybean	446	

1.3.4 Features of diatoms for phytosterol production

Diatoms contain similar phytosterol content to plant sources on a dry weight basis (Table 1.3). Chisti, 2007 indicated that the specific cellular yield of oil obtained from microalgae is approximately 30 times greater than palm oil, the most productive plant crop for oil production. However, high contents of oil were achieved by growing microalgae in limited media conditions which often results in lower biomass production. As sterol content typically correlates with oil content in natural extracts, high oil-producing diatoms may also produce higher levels of sterols. According to Griffiths & Harrison, 2009, the lipid content in diatoms under nutrient-replete conditions is between 21–51% of dry cell mass, whereas in Chlorophyta species the lipid content is between 13–36% of dry cell mass. Table 1.3 lists the amount of sterol produced from different sources, including diatoms.

In comparison to other microbial platform for metabolites production such as bacteria and yeast, diatoms are photosynthetic organisms with reduced carbon source requirements. Bacteria produce hopanoids as analogue to sterols for membrane purposes (Sáenz et al., 2015), with some rare sterols occurring in non-production strains (Wei et al., 2016). Yeast can produce high oil content (Table 1.3) and has been engineered to produce other sterols than only ergosterol (Souza et al., 2011), but is limited in natural sterol diversity.

1.3.5 Challenges for diatom production systems

Although diatoms produce valuable phytosterols, obtaining useful amounts of natural products from these and other microalgae remains challenging. Issues such as controlling growth conditions, contamination, maintenance of productive growth rates through seasonal change, and operational costs remain to be solved (Lebeau & Robert, 2003*b*).

Maintaining optimal conditions of light, temperature, nutrients and inorganic carbon in large scale microalgae cultures is challenging. Both indoor (Satoh et al., 2013) and outdoor (López-Elías et al., 2005; Matsumoto et al., 2017; Sato et al., 2014) mass cultivation of diatoms has been conducted, including outdoor cultivation using natural sunlight (Chen et al., 2011). Light limitation and photoinhibition are issues affecting solar energy conversion into biomass (Melis, 2009), and the use of artificial light can be costly. Temperature control is also a concern, with fluctuations in temperature occurring on daily and seasonal bases (Huang et al., 2017).

Downstream processing of microalgae is currently a challenge due to the free-living planktonic nature of microalgae in large volumes of growth medium. Harvesting methods include centrifugation, flocculation and filtration, each incurring significant costs related to dewatering and drying of biomass (Kim. et al., 2013; Tan et al., 2018). For reasons such as this, an economically competitive production of biofuels from microalgae is not currently feasible (Wang & Seibert, 2017). Nevertheless, higher-value microalgal products can be produced for net gains (Borowitzka, 2013). Technical-economic analysis of diatom-derived

phytosterol blends for consumption may compare favourably to live microalgal biofuels (Malik et al., 2015), which compete against the partly-refined ancient plant and algal "fossil fuels" already found in natural oil deposits.

1.4 Investigation and optimisation of sterol production in diatoms

While diatoms are an attractive source of natural products including sterols at the cellular level, improvements in growth, harvesting, production and metabolic yields will be required to obtain bulk yields of useful material from diatom cultures. Since sterols are involved in cellular processes that interact with environmental conditions, such as cellular signalling, membrane fluidity and permeability, changes in sterol profiles and productivity can vary according to changes in the environment, growth conditions, and metabolic regulation.

1.4.1 Growth conditions

Table 4 describes some of the environmental conditions generally used to grow diatoms for the production of sterol compounds.

1.4.1.1 Light intensity and spectral quality

While some diatoms are able to utilise organic carbon sources to grow mixotrophically (Villanova et al., 2017), the principal research strains perform best in the presence of light ranging from approximately 100 to 300 μmol photons m^{-2}s^{-1} of photosynthetically active radiation (PAR). Véron et al., 1996 found that the sterol content of *P. tricornutum* changed according to temperature and light spectral quality (red, blue, yellow and green-colored light). At 13 °C, maximum sterol content was observed under red light (0.15% of the biomass dry weight), and at 23 °C, maximum sterol content was observed under white light (0.069% of the biomass dry weight). Piepho et al., 2011 reported that sterol content in the freshwater diatom *Cyclotella meneghiniana* increased with light intensity at high phosphorus concentrations, but decreased with light intensity at low phosphorus concentrations. Thus, the effect of light and nutrient conditions may have interacting and variable effects on the sterol metabolism and products of diatoms.

Table 1.4: Culture conditions of diatoms reported for the production of sterol compounds.

Diatom	Culture medium	Temperature (°C)	Light intensity (μmol.m^{-2}s^{-1})	Light color	Light:dark (h)	Reference
Rhizosolenia setigera	f/2 or G2	20–25	70–100	white	12:12	(Barrett et al. 1995)
Skeletonema costatum						
Skeletonema sp.						
Thalassiosira rotula						
Thalassiosira stellaris						
Coscinodiscus sp.						
Haslen ostrearia						
Fragilaria pinnata						
Thalassiothrix heteromorpha						
Thalassionema nitzschioides						
Amphora sp.						
Amphiprora hynlina						
Cylindrotheca fusiformis						
Nitzschia closterium						
Attheya ussurensis sp.	f	20	70	white	12:12	(Ponomarenko et al., 2004)
Delphineis sp.	f/2	18	200	white	12:12	(Giner and Wikfors 2011)
Ditylum brightwellii						
Triceratium dubium						
Phaeodactylum tricornutum	ES–Tris II	13, 23	100	blue, red, yellow, green, white	12:12	(Véron et al., 1996)
Thalassiosira weissflogii	f/2	15	150	–	16:8	(Breteler et al. 2005)

1.4.1.2 Temperature

Sterols mediate the fluidity of cellular membranes under different temperature conditions, and may be actively regulated accordingly (Dufourc, 2008). In plants, phytosterols regulate membrane cohesion to sustain the functional state of membranes during temperature shifts. This is achieved through the formation of small domains in the membrane between the branched ethyl groups in plant sterols (Beck et al., 2007). Similarly, cholesterol in mammalian cells help to maintain membrane microfluidity over a temperature range of 5–65 °C (Dufourc, 2008). Ergosterol in fungi and hopanoids in bacteria serve a similar role in membrane adaptation to temperature changes (Dufourc, 2008).

It is not yet known to what extent temperature affects sterol content and membrane in diatoms. As mentioned above, Véron et al., 1996) showed that the sterol content in *P. tricornutum* at 13 °C increased by more than 50% under all the light spectra evaluated, compared to sterols produced at 23 °C. In contrast, Piepho et al., 2012 showed that the freshwater diatom *Cyclotella meneghiniana* produced higher amounts of sterols at 25 °C than at 10 °C, and that this depended on the supply of phosphorus.

1.4.1.3 Medium composition

The growth of most diatoms relies on the availability of nutrients including silicon (Si), nitrate (N), phosphate (P) and carbon (C), and metals (Round et al., 1990). The limitation of one or more essential nutrients often leads to an increase of lipid production in diatoms (D'Ippolito et al., 2015), as growth arrest results in the diversion of photosynthetic potential

into hydrocarbon production. This creates a tradeoff between growth and the stimulation of lipid-accumulating states with respect to total yield (Chauton et al., 2013; Yang et al., 2013; Eizadora et al., 2009). In obligate silicifying diatoms such as *T. pseudonana* and *C. cryptica*, silicate starvation is sufficient to induce lipid accumulation (Davis, Abbriano, Smith & Hildebrand, 2017; Traller & Hildebrand, 2013).

Interestingly, the production of 24-methylcholesta-5,24(28)-dien-3β-ol in *Thalassiosira weissflogii*, under nitrogen limitation drops from 5.4 to 2.6 μg of sterol per milligram of carbon (Breteler et al., 2005). Piepho et al. (2012) reported that *Cyclotella meneghiniana* produced 5 μg of sterol per milligram of carbon concentration in the presence of a 50 μM phosphate medium, while a concentration of 1.7 μg of sterol per milligram of carbon was obtained in the presence of a 2.5 μM phosphate medium. In the same study, the sterol concentration was reported to be higher at 100 μM silicate than at 50 μM or 200 μM silicate.

Salinity is important parameter that may also affect sterol production, as salt concentration affects membrane fluidity (Russell, 1989). Recently, a study of gene expression of *Halocafeteria seosinensis*, a halophilic protist, revealed that sterol biosynthetic genes were highly up–regulated and/or highly expressed at high salinities (Harding et al., 2017).

1.4.2 Genetic experimentation and engineering in diatoms

In conjunction with an increasing understanding of diatom metabolism, genetic technologies for the manipulation of microalgal strains offer a direct means of investigating novel aspects of their biosynthetic pathways. New genetic engineering tools have been developed for diatoms, including native and inducible promoters and terminators, selectable markers, as well as DNA delivery and transformation techniques (Huang & Daboussi, 2017). These new tools make diatoms accessible to genetics-based biochemistry research and the biotechnology sector.

Diatoms have been genetically probed in order to characterise the function of enzymes in metabolic pathways (Gong et al., 2013; Huang and Daboussi, 2017) and to create lipid-overproducing strains for the sake of biofuel research (Barka et al., 2016; Daboussi et al., 2014; Hamilton et al., 2015; Levitan et al., 2015; Ma et al., 2014; Niu et al., 2016; Trentacoste et al., 2013; Xue et al., 2015; Yao et al., 2014; Zhu et al., 2016). Modulation of native gene expression has also been achieved using antisense RNA and RNAi (Trentacoste et al., 2013), and diatom genes have been successfully edited using TALENs and CRISPR/ Cas9 (Daboussi et al., 2014; Hopes et al., 2016; Nymark et al., 2016; Slattery et al., 2018; Weyman et al., 2015).

Benchmarks for eukaryotic microbial engineering have been set using model yeast strains of *S. cerevisiae*. For example, yeast can be engineered to produce sterols other than ergosterol (Souza et al., 2011). The production of low amounts of lupeol, a triterpenoid produced in plants and not present in diatoms was recently achieved by the introduction of

plant enzymes into *P. tricornutum* (D'Adamo et al., 2019). This altered expression levels of genes involved in the mevalonate pathway and decreased sterol yields, indicating that the secondary metabolism of diatoms is potentially malleable. Additionally, the presence of both MEP and MVA pathways in diatoms might provide both a reliable pool of isoprenoid precursors for the synthesis of sterols and heterologous terpenoids (Vavitsas et al., 2018), as well as a buffer against microbial defects in the modified basal metabolisms of engineered strains.

Altering the sterol metabolic pathway is complex (Liao et al., 2014; Shin et al., 2012; Wang et al., 2012). Due to the abundance of unidentifiable diatom genes and the importance of sterols in maintaining membrane fluidity and other membrane functions, the discovery, silencing, knocking down, adding or deleting of genes affecting sterol synthesis will be challenging. Re-direction of carbon to metabolic pathways of interest without compromising the biomass productivity is one of the major issues (Huang and Daboussi 2017). Risk assessments of the cultivation of genetically modified algae in large scale must be conducted and verified (Henley et al., 2013). The production of useful compounds from genetically modified and transgenic organisms is currently limited to qualified laboratories and fully contained indoor cultivation systems. To protect the environment from potential risks related to the release of transgenic organisms, containment standards do not currently permit outdoor cultivation of genetically modified algae for production purposes (Davis, Crum, Corbeil & Hildebrand, 2017, Hempel & Maier, 2012, Ro et al., 2006). However, trial experiments indicate that organisms can be contained in carefully controlled outdoor systems (Szyjka et al., 2017). In order to efficiently study and alter sterol and secondary metabolism in diatoms and other microalgae, a combination of direct genetic techniques, high-throughput phenotyping, metabolomics and systems-based modeling may be the best approach.

1.5 Conclusions

Diatoms are broadly distributed and evolutionarily diversified microalgae that produce a wide range of potentially valuable phytosterol compounds. Despite broad characterization of natural sterol species produced by diatoms, the details of their differently evolved biosynthetic and regulatory pathways are still being determined. Further research into these aspects of phytochemistry and physiology can help us to understand the evolution of sterol diversity in diatoms, as well as the potential for diatoms and related species to produce useful amounts of terpenoids and secondary metabolites as interesting bioproducts.

The demand for biologically sourced sterol products in the pharmaceutical, nutraceutical and food markets is expanding, encouraging the pursuit of new sources of naturally and synthetically produced sterols. The adaptation and engineering of easily culturable

diatoms into existing research and systems for industrial and sustainable cultivation can provide an alternative and additional source of attractive and interesting phytosterols for natural bioproducts markets.

Acknowledgments

Chapter 2

Levels of diatom minor sterols respond to changes in temperature and salinity

Author Contributions

A.C.J.M. designed and performed all experiments, and wrote the manuscript. J.A. designed, advised and materially supported experiments and assisted in writing. P.J.R. advised and materially supported experiments and assisted in writing.

Abstract

Diatoms are a broadly distributed and evolutionarily diversified group of microalgae that produce a diverse range of sterol compounds. Sterols are triterpenoids that play essential roles in membrane-related processes in eukaryotic cells. Some sterol compounds possess bioactivities that promote human health and are currently used as nutraceuticals. The relationship between sterol diversity in diatoms and their acclimation to different environments is not well understood. In this study, we investigated the occurrence of different sterol types across twelve diatom species, as well as the effect of temperature reduction and changes in salinity on the sterol contents of three model diatom species. In the diatoms Thalassiosira pseudonana, Phaeodactylum tricornutum and Chaetoceros muelleri, we found that changes in the relative contents of minor sterols accompanied shifts in temperature and salinity. This may be indicative of acquired adaptive traits in diatom metabolism.

Key words: Diatoms, sterols, temperature, salinity, phytosterols

2.1 Introduction

Sterols are essential triterpenoids present in all eukaryotes Volkman, 2005. They contribute to the maintenance of important cellular dynamics such as membrane permeability, signaling and modulation of membrane-bound protein functions (Dufourc, 2008). Besides their importance in signaling, sterol metabolism has been reported to function as a "molecular fossil" to track important evolutionary events such as oxygen availability in the ocean–atmosphere system (Gold et al., 2017). Plant sterols, known as phytosterols, have beneficial effects on human health including demonstrated cholesterol-lowering (Ras et al., 2014) properties, anti-inflammatory (Aldini et al., 2014) and anti-diabetic activities (Wang et al., 2017). Phytosterols are currently commercialized as nutraceuticals in many food products such as spread, cream cheese, yoghurt and milk (Lin et al., 2016). The Board of Food Standards Australia New Zealand (Federal Register of Legislative Instruments F2015L00440) and The European Atherosclerosis Society Consensus Panel have classified sterols as safe compounds to use in foods, as well as in the administration to patients for the treatment of hypercholesterolemia (Zampelas, 2014).

In animals and higher plants, sterols are involved in the synthesis of secondary metabolites such as steroid hormones (Valitova et al., 2016). Unlike animals and fungi, which mostly contain a single sterol, plants and algae contain a diverse range of sterols (Desmond & Gribaldo, 2009; Hannich et al., 2011). Diatoms (Bacillariophyceae), in particular, produce a high diversity of unique and rare sterols: at least twenty-five different sterols have been empirically identified in over one hundred diatom species (Rampen et al., 2010). Diatoms are unicellular phytoplankton that serve as primary photosynthetic engines in the global

carbon cycle (Benoiston et al., 2017). They are evolutionarily diversified microalgae that are broadly distributed in different aquatic environments (Armbrust, 2009). Diatoms are an important source of sterols in the marine ecosystem due to their predominance in the ocean (Rampen et al., 2010). Although the sterol profiles of diverse diatom species have been studied, the influence of environmental conditions on the abundance and identities of species-specific sterol repertoires is not fully known (Jaramillo-Madrid et al., 2019). Environmental conditions such as salinity (Ahmed et al., 2015) and the availability of nutrients including nitrogen, phosphate and silicate appear to affect sterol production in some microalgae (Breteler et al., 2005; Piepho et al., 2012).

Since sterols are involved in cellular processes that interact with environmental conditions, such as cellular signaling and membrane fluidity (Dufourc, 2008), the regulation of sterol content by diatom species might be sensitive to changes in the environment and growth conditions. It is not known whether the wide diversity of sterols compounds in diatoms respond to changes in temperature, as in plants, or whether they are involved in complex signaling functions such as cell death (Gallo et al., 2017). The presence of 23-methyl sterols, cyclopropyl sterols and 27-norsterols is thought to provide a defence mechanism against grazers, potentially interfering with the sterol metabolism of predatory organisms (Giner & Wikfors, 2011). Furthermore, conditions like salinity have been reported to affect important cellular processes in diatoms such as biosilicification (Vrieling et al., 2007).

In this study, we surveyed the sterol contents of twelve diatoms representing multiple clades and environments. We then evaluated the effects of a reduction in temperature and changes in salinity on the cell growth and sterol profiles of three commonly cultured mesophilic model diatoms: the centric diatom *Thalassiosira pseudonana*, the pennate diatom *Phaeodactylum tricornutum* and the chaetocerid diatom *Chaetoceros muelleri*.

2.2 Materials and Methods

2.2.1 Diatom species

The species Phaeodactylum tricornutum (CCMP632), Thalassiosira pseudonana (CCMP13-35), Thalassiosira oceanica (CCMP1005) and Chaetoceros muelleri (CCMP1316) were obtained from the National Centre for Marine Algae and Microbiota. Lauderia annulata, Thalassiosira rotula and Chaetoceros curvisetus were isolated from Sydney harbour and further identified by 18S sequencing. The cells were maintained in L1 medium at 18 °C under cool white continuous light (100 μmol photons m^{-2} s^{-1}). Polar diatoms *Eucampia sp.*, *Fragilariopsis sp.*, *Pseudonitzshia sp.* and *Nitzshia sp.* were first isolated from Prydz Bay, Davis Station, Antarctica (66°S, 77°E) during the Austral Spring (November 2014) and maintained in 0.2 μm filtered seawater (salinity 35 ppt) enriched with an adjusted L1 stock medium for

all Antarctic cultures. Cultures were maintained under 50 μmol photons m^{-2} s^{-1} (14:10 hour light: dark cycle) at 3 °C.

2.2.2 L1 medium composition

To prepare 1 L of L1 medium, 1 mL of NaNO$_3$ (75 g L^{-1}), NaH$_2$PO$_4 \cdot$ H$_2$O (5 g L^{-1}), Na$_2$SiO$_3$ $\cdot$ 9 H$_2$O (30 g L^{-1}), trace element solution (see composition below) and 0.5 mL of vitamin solution (see composition below) were added to 1 L of artificial seawater (see composition below), as described by NCMA in the algal media recipes section (http://ncma.bigelow.org/).

   Artificial seawater: NaCl (21 g L^{-1}), Na$_2$SO$_4$ (4.09 g L^{-1}), KCl (0.7 g L^{-1}), NaHCO$_3$ (0.2 g L^{-1}), KBr (0.1 g L^{-1}), H$_3$BO$_3$ (0.03 g L^{-1}), NaF (0.003 g L^{-1}), MgCl$_2 \cdot$ 6H$_2$0 (11.1 g L^{-1}) and CaCl$_2 \cdot$ 2H$_2$O (1.54 g L^{-1}).

   Trace element solution: Na$_2$EDTA $\cdot$ 2H$_2$O (1.17 $\times$ 10^{-5} M), FeCl$_3 \cdot$ 6H$_2$O (1.17 $\times$ 10^{-5} M), MnCl$_2 \cdot$ 4H$_2$O (9.1 $\times$ 10^{-7} M), ZnSO$_4 \cdot$ 7H$_2$O (8 $\times$ 10^{-8} M), CoCl$_2 \cdot$ 6H$_2$O (5 $\times$ 10^{-8} M), CuSO$_4 \cdot$ 5H$_2$O (1 $\times$ 10^{-8} M), Na$_2$MoO$_4 \cdot$ 2H$_2$O (8.22 $\times$ 10^{-8} M), H$_2$SeO$_3$ (1 $\times$ 10^{-8} M), NiSO$_4 \cdot$ 6H$_2$O (1 $\times$ 10^{-8} M), Na$_3$VO$_4$ (1 $\times$ 10^{-8} M) and K$_2$CrO$_4$ (1 $\times$ 10^{-8}-8 M).

   Vitamin solution: Thiamine $\cdot$ HCl (2.96 $\times$ 10^{-7} M), biotin (2.05 $\times$ 10^{-9} M) and cyanocobalamin (3.69 $\times$ 10^{-10} M).

2.2.3 Growth and harvesting of diatoms

All diatom species were grown in flasks containing 500 mL of L1 medium at the specified light and temperature conditions. Cells were harvested during the late exponential growth phase by centrifuging at 4000g for 10 minutes. Diatom pellets were washed with Milli-Q water to eliminate salt excess, freeze-dried to determine dry matter weight and kept at -20 °C until sterol extraction.

2.2.4 Reduced temperature experiments

The diatom species were cultured in flasks containing 1 L of L1 medium at 18 °C under continuous light with an intensity of 100 μmol photons m^{-2} s^{-1}. When the cultures reached exponential growth phase, they were transferred to 4 °C to test the effect of rapid cooling on growth and sterol content. To track changes in the sterol profile, 100 mL of culture was sampled immediately before the transition to 4 °C, and then every 24 h until 96 h elapsed. Control cultures remained at 18 °C and were harvested after 96 h. Samples were used for cell counting and sterols extraction. The experiments were performed in triplicate. Biomass was harvested by centrifuging at 4000g for 10 minutes. Diatom pellets were washed with Milli-Q water to eliminate salt excess, freeze-dried to determine dry matter weight and kept at -20 °C until sterol extraction.

2.2.5 Assay for salinity tolerances of diatom species

To determine the tolerance of different levels of salinity on the growth of diatoms species, screening was carried out in 96-well plates. To adjust salinity of L1 medium different amounts of sodium chloride in artificial seawater were added: 0, 8, 16, 24, 32, 40, 48 and 56 g l^{-1}. Other trace salts components remained the same. Total salt concentrations for amended L1 medium evaluated were 10, 17, 25, 30, 39, 47, 53 and 61 parts per thousand (ppt). The diatoms *P. tricornutum*, *T. pseudonana* and *C. muelleri* were inoculated in triplicate in 96-well plates for each salinity level. Changes in relative growth rates were estimated using a fluorescence plate reader (Infinite M100 pro, Tecan Austria GmbH, Austria), with excitation at 485 nm and emission at 680 nm (Spilling & Seppälä, 2017).

2.2.6 Cultivation at altered salinities

The diatoms *T. pseudonana*, *P. tricornutum* and *C. muelleri* were inoculated in triplicate 1 L-flasks with L1 medium 10 ppt, normal salinity conditions (30 ppt), and high salinity content of 61 ppt. For *T. pseudonana*, the highest salinity was 53 ppt. After five days, 100 mL of culture was harvested by centrifuging at 4000g for 10 minutes. Diatom pellets were washed with Milli-Q water to eliminate salt excess, freeze-dried to determine dry matter weight and stored at -20 °C until sterol extraction.

2.2.7 Extraction and analysis of sterols by GC–MS

For sterol extraction, dry cell matter was heated in 1 mL of 10% KOH ethanolic solution at 90 °C for one hour. Sterols were extracted from cooled material in three volumes of 400 μL of hexane. An internal standard, 5-α-cholestane, was added to each sample. Hexane fractions were dried under a gentle N_2 stream, and derivatized with 50 μL of 99% BSTFA + 1% TMCS at 70 °C for one hour. The resulting extractions were re-suspended in 50 μL of fresh hexane prior to GC–MS injection.

Gas chromatography/mass spectrometry (GC–MS Agilent 7890, Agilent Technologies, Inc.) analysis was performed using an instrument equipped with a HP-5 capillary column (30 m; 0.25 mm inner diameter, film thickness 0.25 μm) coupled to an Agilent quadrupole MS (5975 N) instrument. The following settings were used: oven temperature initially set to 50 °C, with a gradient from 50 to 250 °C (15.0 °C min^{-1}), and then from 250 to 310 °C (8 °C min^{-1}, hold 10 min); injector temperature = 250 °C; carrier gas helium flow = 0.9 ml min^{-1}. A split-less mode of injection was used, with a purge time of 1 min and an injection volume of 2 μL. Mass spectrometer operating conditions were as follows: ion source temperature 230 °C; quadrupole temperature 150 °C; accelerating voltage 200 eV higher than the manual tune and ionization voltage 70 eV. Full scanning mode with a range from 50 to 650 Dalton was used.
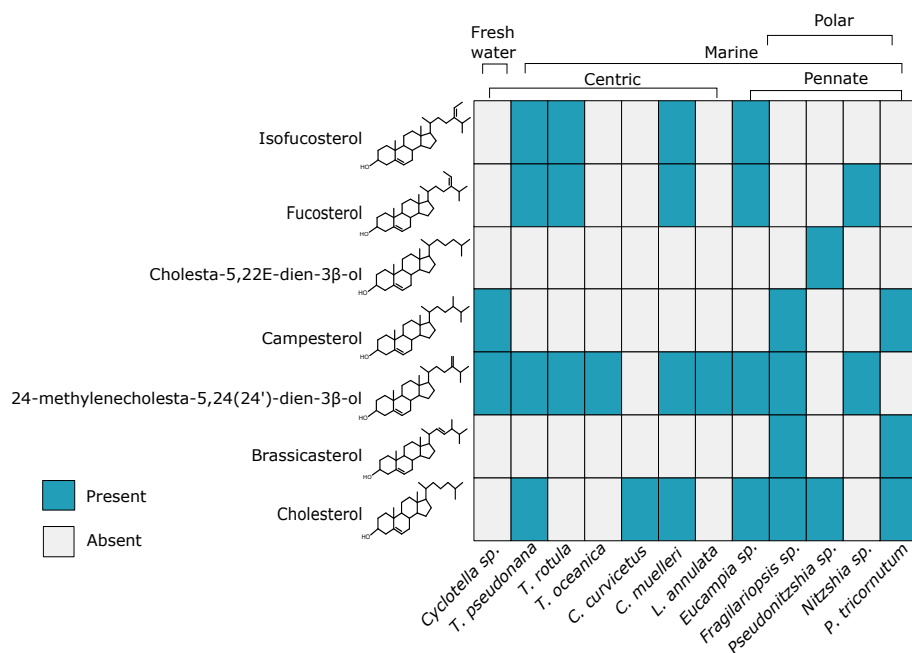
Figure 2.1: Sterol distribution among 12 different diatom species. Diatom species were grown at 18 °C except by polar diatoms cultivated at 3 °C.

Sterol peaks were identified based on the retention time, mass spectra and representative fragment ions compared to the retention times and mass spectra of authentic standards. The NIST (National Institute of Standards and Technology) library was also used as a reference. The area of the peaks and deconvolution analysis was carried out using the default settings of the Automated Mass Spectral Deconvolution and Identification System AMDIS software (v2.6, NIST). Peak area measurements were normalized by both the weight of dry matter prior to extraction, and the within-sample peak area of the internal standard 5a-cholestane. Sterol standards used to calibrate and identify GC–MS results in this study included: cholest-5-en-3-β-ol (cholesterol); (22E)-stigmasta-5,22-dien-3β-ol (stigmasterol); stigmast-5-en-3-β-ol (sitosterol); 24-methylcholest-5-en-3β-ol (campesterol); (22E)-ergosta-5,22-dien-3-β-ol (brassicasterol); (24E)-stigmasta-5,24-dien-3β-ol (fucosterol); 9,19-cyclo-24-lanosten-3β-ol (cycloartenol); 5-α-cholestane and the derivatization reagent bis(trimethyl-silyl) trifluoroacetamide and trimethylchlorosilane (99% BSTFA + 1% TMCS) and were obtained from Sigma-Aldrich, Australia. Sterol levels in terms of μg of sterol per mg of dry weight were calculated using calibration curves. Due to the lack of standards, 24-methylcholesta-5,24(24′)-dien-3β-ol and (24Z)-stigmasta-5,24-dien-3β-ol (isofucosterol) were quantified using the calibration curve of cholesterol and fucosterol, respectively, supposing that they have similar responses.

2.2.8 Statistical analysis

All plots were generated using R: A language and environment for statistical computing, version 3.3.2 (Team et al., 2016). All experiments were conducted in triplicate. The analyses performed were a Student's t-test and Spearman's rank correlation (ρ). p values were adjusted using Bonferroni correction for multiple comparisons. Differences were considered significant at $p < 0.05$.

2.3 Results

The sterol contents of twelve diatom species evaluated here agreed with earlier studies (Rampen et al., 2010). Each diatom has shown a unique sterol profile that differs significantly between closely related species (Fig. 2.1). The most commonly found sterol was 24-methylcholesta-5,24(24')-dien-3β-ol, which was present in all but three species: *C. curvisetus, P. tricornutum* and the polar diatom *Pseudonitzshia sp.* The occurrence of various sterol types across different diatoms did not appear to clearly segregate by clade or environment (Fig. 2.1) (Medlin & Kaczmarska, 2004; Rampen et al., 2010), suggesting a more complex route to evolutionary diversification of sterol production in diatoms.

To investigate dynamic responses of diatom sterol production to changes in their environment, the diatom species *C. muelleri, T. pseudonana* and *P. tricornutum* were separately subjected to i) a reduction in temperature and ii) changes in salinity.

2.3.1 A shift to cold temperature (4 °C) caused species-specific growth effects and changes in minor sterols

Under normal conditions, *P. tricornutum* produced brassicasterol as its principal sterol type, as well as minor amounts of campesterol and cholesterol (Fig. 2.2). *C. muelleri* principally produced cholesterol and fucosterol, as well as minor amounts of 24-methylcholesta-5,24(24')-dien-3β-ol, isofucosterol and traces of an unknown sterol. *T. pseudonana* produced mainly 24-methylcholesta-5,24(24')-dien-3β-ol, as well as smaller amounts of campesterol, fucosterol, isofucosterol and cholesterol. Cholesterol was the only common sterol detected in all the three diatom species (Fig. 2.2).

In response to a shift to cold temperature (4 °C), the diatoms *C. muelleri, P. tricornutum* and *T. pseudonana* exhibited different growth effects (Fig. 2.3). The cell density of *C. muelleri*, which was isolated in 24 °C seas near Hawaii, decreased after 24 h at 4 °C. *T. pseudonana*, which was isolated from a temperate seasonal environment (10–20 °C), maintained cell density at 4 °C but did not multiply further. In contrast, *P. tricornutum* continued to grow at 4 °C, despite a reported optimal temperature range of 11–21 °C (Fig. 2.3) (Andersen et al., 1997).
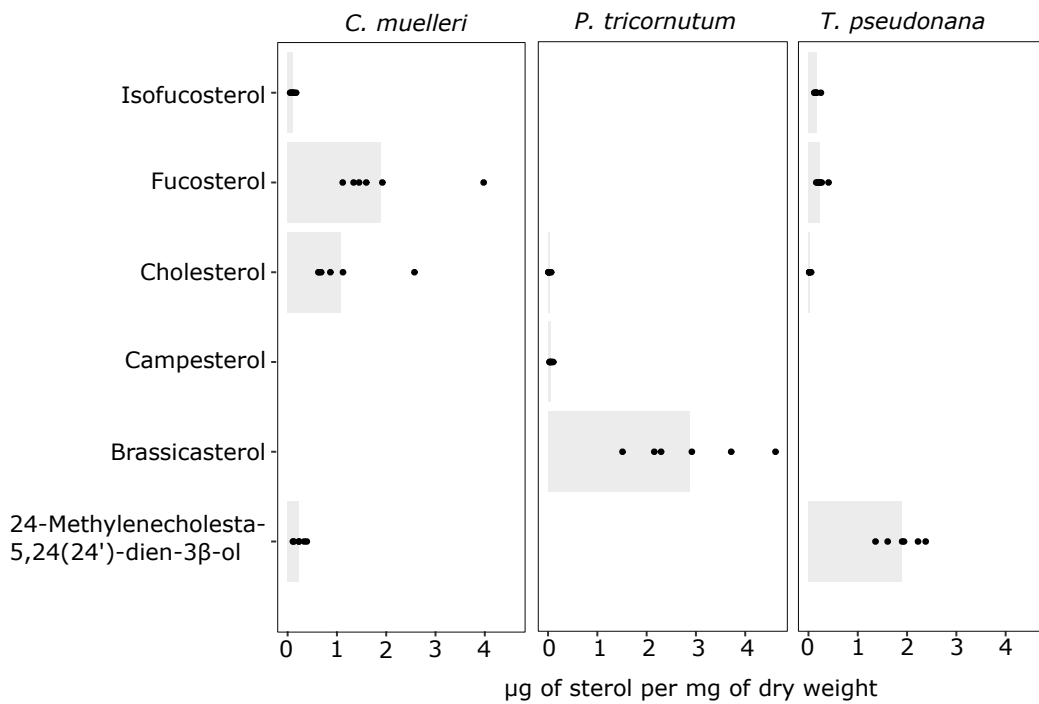
Figure 2.2: Sterol abundance and distribution in the diatoms *Chaetoceros muelleri, Phaeodactylum tricornutum* and *Thalassiosira pseudonana* growing under standard conditions: L1 medium at 18 °C under cool white continuous light (100 μmol photons m^{-2}s^{-1}; n = 6).

Changes in sterol content during the shift to lower temperature appeared to correlate with survival. In *C. muelleri*, sterol content did not change significantly, with the exception of a marginal decrease in fucosterol (Spearman's ρ: -0.75, p < 0.0012; Fig. 2.4, t-test, p = 0.015; Fig. 2.5). In *T. pseudonana*, multiple sterols appeared to shift in relative abundance, but with insufficient replicates to provide confident conclusions. In *P. tricornutum*, campesterol increased (Spearman's ρ: 0.91, p < 0.0000016, t-test, p = 0.03), while cholesterol decreased (Spearman's *rho*: -0.58, p < 0.0046, t-test, p = 0.037; Figs. 2.4 and 2.5).

2.3.2 Species-specific tolerance to different salinities

To determine ranges of salinity in which each diatom species was able to grow, we performed a screening of eight different salt concentrations in amended L1 medium, including the standard salinity level of 30 ppt that was used for the cultivation of marine diatoms. The diatoms *C. muelleri* and *T. pseudonana* were tolerant to salinities between 17 and 39 ppt (Fig. 2.6). At 47 ppt and above, impairment in growth was observed, especially for *T. pseudonana*. At low salinity levels (10 and 17 ppt), the growth of *C. muelleri* and *T. pseudonana* was affected. *P. tricornutum* continued to grow well at low salinity, which is consistent with brackish marine environments from which it was isolated (Andersen et al., 1997) (Fig. 2.6).
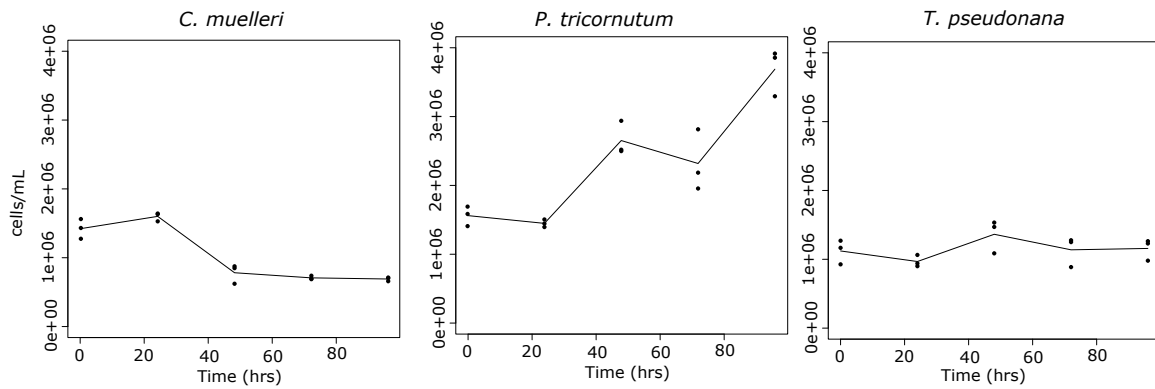
Figure 2.3: Growth curves of diatoms under 4 °C. Time 0 represents the beginning of the temperature shock (n = 3).

2.3.3 Salinity affects the relative contents of non-principal sterols in diatoms

To measure the effect of salinity on the sterol content of diatoms, we grew *C. muelleri* and *P. tricornutum* in medium containing salt at concentrations of 10 ppt, 30 ppt and 61 ppt, and *T. pseudonana* in medium containing salt at concentrations of 10 ppt, 30 ppt and 53 ppt. The amount of the principal sterol found in *C. muelleri*, cholesterol, was not affected by changes in salinity. 24-Methylcholesta-5,24(24')-dien-3β-ol decreased at 61 ppt (Spearman's ρ: -0.9, p < 0.001; Fig. 2.7). In *T. pseudonana*, levels of its principal sterol 24-methylcholesta-5,24(24')-dien-3β-ol were not affected by changes in salinity, but fucosterol and isofucosterol increased in proportion to salinity concentration (Fig. 2.7). At low salinity (10 ppt), fucosterol was reduced by 50% compared to standard salinity (30 ppt), and increased by 20% at high salinity (53 ppt; Spearman's ρ: 0.9, p < 0.001). In *P. tricornutum*, brassicasterol and campesterol levels were negatively correlated with salinity concentration: at low salinity (10 ppt) both sterols increased relative to standard salinity (30 ppt) and decreased at high salinity (61 ppt; Spearman's ρ: -0.74, p < 0.02). Traces of cholesterol were detected in *T. pseudonana* and *P. tricornutum* only under standard salinity conditions (30 ppt; Fig. 2.7).

2.4 Discussion

2.4.1 The composition of sterol types in different diatoms is not simply explained by clade or environment

In this study, profiling of sterol compounds in 12 broadly distributed, evolutionarily divergent diatom species resulted in a wide distribution of sterol compositions that could not be explained simply by diatom clade or environmental niche (Fig. 2.1). This lack of a
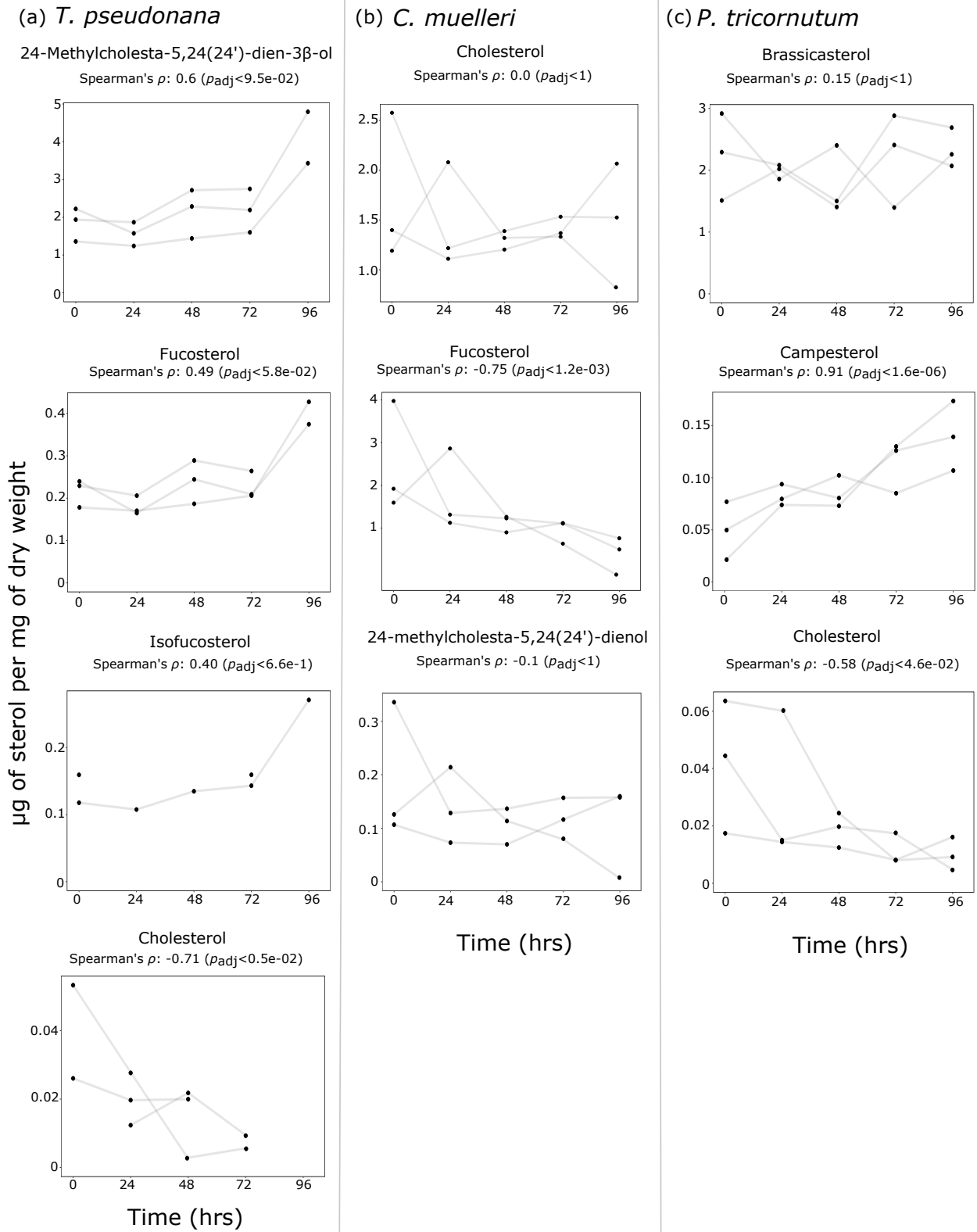
Figure 2.4: Sterol levels in diatoms growing at 4 °C. Abundance is given in terms of μg of sterol per mg of biomass dry weight (n = 3).

general relationship between diatom clades and sterol patterns has been noted previously, and precludes the use of sterol profiling as a means of unambiguous phylogenetic assignment (Rampen et al., 2010). Species-specific differences in sterol composition are likely related to evolution and diversification of sterol biosynthesis. Diatoms share a common core of early sterol biosynthetic enzymes that is unique to diatoms (Jaramillo-Madrid et al., 2019; Pollier et al., 2019), and lineage-specific divergence likely occurs in downstream reactions (Jaramillo-Madrid et al., unpublished data). Typical C-22 desaturated sterols such as brassicasterol have been found in *P. tricornutum*, consistent with the presence of sterol 22-desaturase enzymes, whereas the same desaturases have not been detected in the centric diatoms *T. pseudonana* and *C. muelleri* (Jaramillo-Madrid et al., unpublished data).

The relative proportions of sterol types produced by each species also appear to be a specifically evolved property, and the abundance of sterol compounds is different in each diatom (Fig. 2.2). Similarly, in plants, the balance between 24-methylsterols and 24-ethylsterols varies by individual species (Valitova et al., 2016). Moreover, different plant sterols play particular functions within the cell, such as precursors of plant hormones that regulate growth and development in different tissues (Valitova et al., 2016). Besides sterols occurring in free forms, conjugated sterols such as esterified and glycosylated sterols have been found in diatoms, but their biological role is still unclear (Véron et al., 1996, 1998). Sterol sulphates are involved in signaling to induce cell death in the marine diatom *Skeletonema marinoi* often present in blooms (Gallo et al., 2017). Additional specific functional roles of sterol types in diatoms beyond are not yet known.

2.4.2 *P. tricornutum* thrives and shifts its sterol content at a reduced temperature

In laboratory experiments, we found a correlation between changes in the sterol profile of the diatom *P. tricornutum* and a significative reduction of culture temperature. In plants, phytosterols regulate membrane cohesion to sustain the functional state of membranes during temperature shifts (Valitova et al., 2016). This is achieved by the presence of phytosterols with additional ethyl groups branched on C-24, unlike cholesterol, which is the main sterol found in mammals, which reinforces attractive van der Waals interactions resulting in increased membrane cohesion (Beck et al., 2007). We found that in the diatom *P. tricornutum* the amount of campesterol, a C-24 methyl branched sterol, increased gradually during the temperature shock unlike cholesterol that was significantly reduced after 96 h of low temperature growth (Figs. 2.4 and 2.5). However, brassicasterol, the most abundant sterol found in *P. tricornutum*, levels remain unchanged (Figs. 2.4 and 2.5). These results suggest that a similar strategy used by plants for fine tuning of non-principal sterols (Beck et al., 2007) may also operate in diatoms.

Further research is required to describe to what extent different temperatures affect
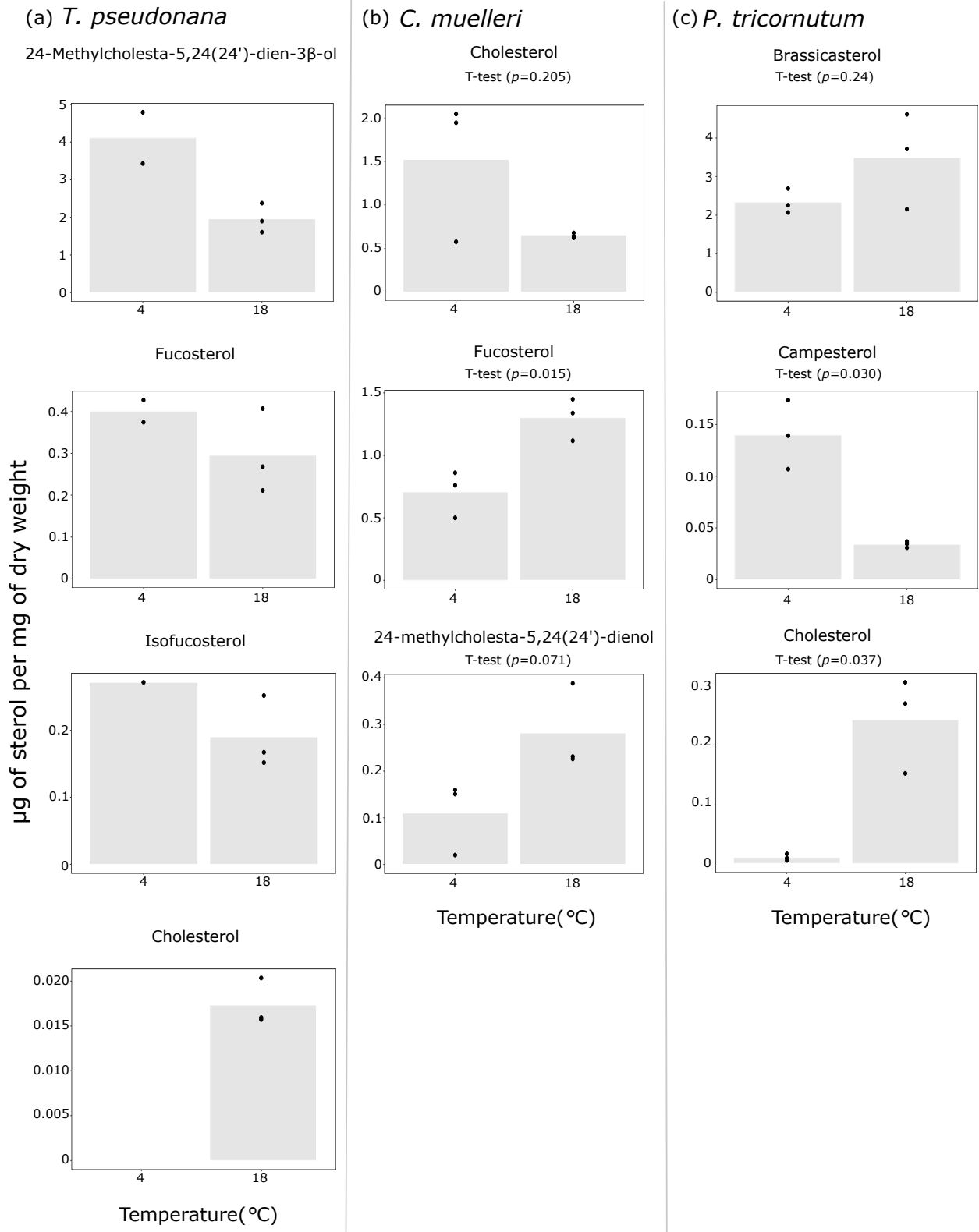
Figure 2.5: Sterol levels in diatoms growing at 4 °C and 18 °C after 96 h of temperature shock. Abundance is given in terms of μg of sterol per mg of biomass dry weight (n = 3). A t-test is provided for samples with n = 3.
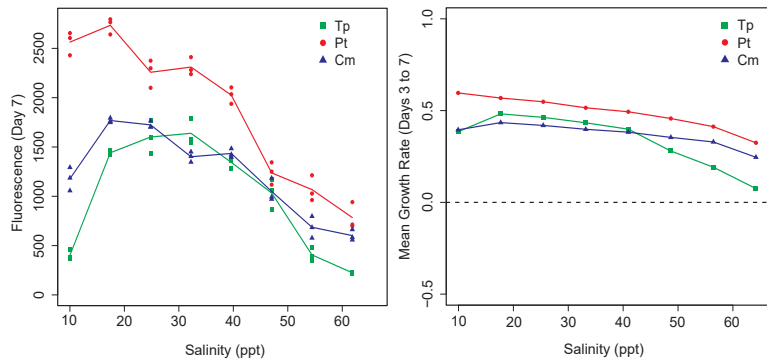
Figure 2.6: Screening of growth yield and growth rate versus the salinity level in the diatoms *C. muelleri* (Cm), *P. tricornutum* (Pt) and *T. pseudonana* (Tp; n = 3). Screening was performed in 96-well plates. Growth was estimated using fluorescence with excitation at 485 nm and emission at 680 nm.

sterols composition in diatoms, how this is regulated, and whether general trends can be reliably detected. It was reported that the sterol content in *P. tricornutum* at 13 °C increased by more than 50% compared to sterols produced at 23 °C (Véron et al., 1996), but an increase of this magnitude was not observed in our study. In contrast, Piepho et al. (2012) showed that the freshwater diatom *Cyclotella meneghiniana* produced lower amounts of sterols at 10 °C than at 25 °C, and that this also depended on the supply of phosphorus (Piepho et al., 2012).

The effect of a shift to cold temperature on the relative contents of sterols in *T. pseudonana* and *C. muelleri* was less conclusive than for *P. tricornutum* (Fig. 2.3). It is possible that the temperatures chosen were too low for these species to effectively re-acclimate, or that sterol regulation is not involved in the tolerance or sensitivity of these diatoms to low temperature.

2.4.3 Minor sterols respond to changes in salinity and temperature

L1 medium and seawater salinity is about 30 ppt and 32 ppt respectively (Guillard & Hargraves, 1993). Thus, the best growth for the three marine diatom species tested was expected in this salinity range. We found that the diatoms *C. muelleri* and *T. pseudonana* grew well within in a range of sodium chloride between 25 and 39 ppt (Fig. 2.6). Sensitivity to high salinity observed in *T. pseudonana* may agree with the theory that this species is an ancestrally freshwater diatom (Alverson et al., 2011). Surprisingly, *P. tricornutum* appeared to have a better growth rate under low salinity levels, rather than at 30 ppt or higher. This species is commonly reported to grow well in a range of salinities between 20 and 30 ppt (Liang et al., 2014), but its tolerance to lower salinities is consistent with the near-shore, brackish and urban environments from this species is commonly isolated (Andersen et al.,

1997). Other diatoms such as *Chaetoceros gracilis* and six different species of Skeletonema were reported to be tolerant to a wide range of salinities (Balzano et al., 2010; Tokushima et al., 2016).

The principal sterols of *C. muelleri*, *P. tricornutum* and *T. pseudonana* were not significantly affected by salt concentration. However, significant variations were observed for sterols constituting a minor portion of the total sterols (Fig. 2.7). This suggests that less abundant sterols might be involved in mechanisms of response to osmotic shock, as it was previously reported for *Pavlova lutheri* (Ahmed et al., 2015). The diatom sterols that were responsive to changes in salinity, such as 24-methylcholesta-5,24(24′) -dien-3β-ol, campesterol and fucosterol are also intermediaries in the sterol biosynthetic pathways of the diatoms *C. muelleri*, *P. tricornutum* and *T. pseudonana* respectively (Jaramillo-Madrid et al., unpublished data). The concentrations of accessory sterols could be controlled through regulation of the activities of C-24 alkyl transferases that catalyse terminal reactions in sterol biosynthesis. The detection of cholesterol in *T. pseudonana* and *P. tricornutum* only under normal salinity conditions (30 ppt) suggests that sterols with methylene, methyl and ethylidene groups on C-24, such as 24-methylcholesta-5,24(24′)-dien-3β-ol, campesterol and fucosterol, respectively, may play a more potent role in adaptive changes to osmotic and temperature stresses. An alternative theory for observed changes in these non-principal sterols is a simple loss in metabolic turnover of sterol intermediates; however, the apparent ease with which *P. tricornutum* thrived in low temperature and low salinity conditions disagrees with that interpretation. Some of the sterol compounds accumulated under temperature and salinity shock, campesterol and fucosterol have been confirmed to have cholesterol-lowering and anti-cancer activities (Luo et al., 2015). This capacity of regulation might represent an advantage when selecting a microalga species for large scale production of sterol compounds.

Other important membrane components such as fatty acids have been reported to fluctuate under different conditions of temperature and salinity. Total fatty acids content decreases in the diatom *P. tricornutum* after a temperature shift from 25 to 10 °C, whereas eicosapentaenoic acid and polyunsaturated fatty acids considerably increased (Jiang & Gao, 2004). Moreover, the diatoms *P. tricornutum* and *Haslea ostrearia* seem to have a similar mechanism of adaptation to high temperatures by adjusting length and number of unsaturation in their fatty acids compounds (Dodson et al., 2014). In the same way, salinity plays an important role in lipid accumulation. A transcriptomics analysis of the oleaginous diatom *Nitzschia sp.* revealed that genes involved in lipid biosynthesis were up-regulated under salinity stress (Cheng et al., 2014). Similarly, triacylglycerol content decreased under high salinity levels in the diatom *C. gracilis* (Tokushima et al., 2016). The levels of pigments, which are catalysed from the same precursors as sterols (Delgado-Vargas et al., 2000), were reported to remain steady under different salinity conditions in the diatom *Pseudo-nitzschia australis.* However, secondary metabolites like toxic domoic acid was reported to increase under low salinity levels (Ayache et al., 2019). This indicates that diatoms possess mecha-
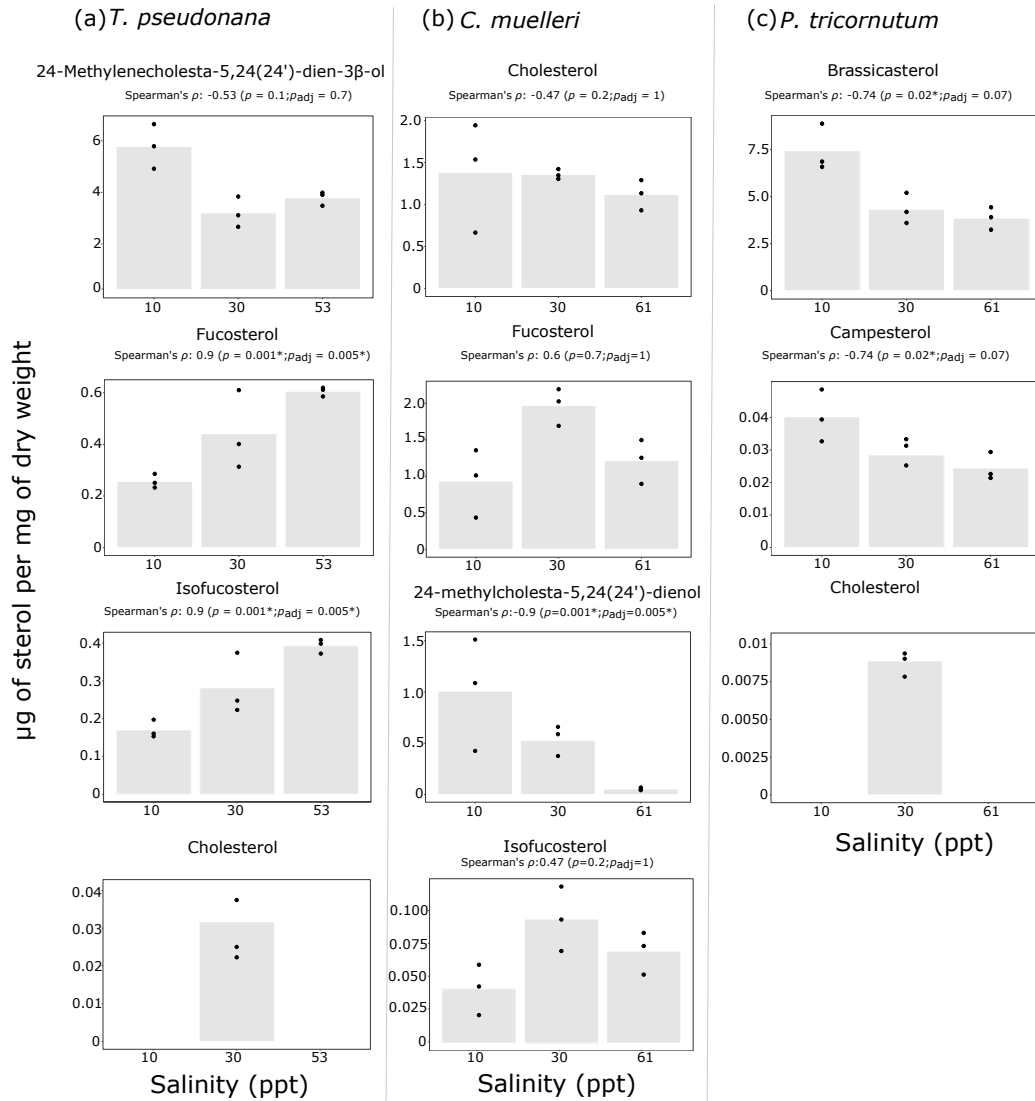
Figure 2.7: Sterol levels in diatoms growing at different salinity levels. Abundance is given in terms of μg of sterol per mg of biomass dry weight (n = 3).

nisms to adapt to changes in environmental conditions by modifying contents of primary metabolites such as sterols.

2.5 Conclusions

In this study, we found that while the sterol types produced by different diatom species were not simply explained by the environment or clade, changes in the relative contents of minor sterols accompanied shifts in temperature and salinity in three commonly cultivated diatoms. This study provided insight into the role that sterol and sterol diversity might play in the capabilities of diatoms to adapt and survive under changing environmental conditions. The augmentation of less abundant sterols with different physical effects on

membrane cohesion might provide a simple means of tuning membrane dynamics, in contrast to bulk shifts in the levels of principal sterols. Thus diatoms might join halotolerant microalgae *Dunaliella salina* (Zelazny et al., 1995) and *Halocafeteria seosinensis* (Harding et al., 2017) as microalgal species whose regulation of sterols is included in adaptation and acclimations to new and dynamic environments. Further metabolic characterization of these new and species-specific reactions may help to clarify the relationship between metabolic evolution and the environment.

Funding

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

Chapter 3

The unique sterol biosynthesis pathway of three model diatoms consists of a conserved core and diversified endpoints

Author Contributions

A.C.J.M. designed and performed all experiments, and wrote the manuscript. J.A. designed, advised and materially supported experiments, sequenced and analyzed RNA-seq data, performed bioinformatics, and assisted in writing. M.F. designed and advised experiments, reviewed and assisted in writing. P.J.R. advised and materially supported experiments.

Published in Algal Research on June 2020:

Abstract

Diatoms produce a wide diversity of sterols amongst different species, the biosynthesis and conservation of which is not yet fully understood. To investigate the conservation and divergence of sterol biosynthesis pathways amongst diatoms, we performed comparative metabolic profiling and transcriptomics for a centric diatom (*Thalassiosira pseudonana*), a pennate diatom (*Phaeodactylum tricornutum*) and a chaetocerid (*Chaetoceros muelleri*) in response to inhibitors of enzymes involved in sterol biosynthesis.These three model diatoms, which are representative of distinct clades, share a unique core phytosterol biosynthesis pathway, that relies on a terbinafine-insensitive alternative squalene epoxidase, and the cyclization of 2,3 epoxysqualene into cycloartenol by a conserved oxidosqualene cyclase. Lineage-specific divergence in the synthesis of sterol precursors was found in the species analysed. Cholesterol synthesis in diatoms seems to occur via cycloartenol rather than lanosterol. The diversification of natural products produced by each species appears to occur downstream of all experimentally targeted enzymes, suggesting adaptive specialization in terminal synthesis pathways.

Key words: diatoms, phytosterols, sterol biosynthesis, metabolism, transcriptomics, triterpenoids

3.1 Introduction

Diatoms are a large group of eukaryotic microalgae that arose through secondary endosymbiosis, and are renowned for their ecological distribution (Armbrust, 2009). Diatoms have genetically diversified their physiology, metabolism and natural products while adapting to dynamic environments (Moustafa et al., 2009). Among these metabolic products are an expanded repertoire of phytosterols (Rampen et al., 2010), a class of essential terpenoids that are involved in the regulation of membrane dynamics, signaling, and membrane-bound protein functions in higher plants, algae, fungi, and vertebrates (Dufourc, 2008). Moreover, sterol sulfates have been implicated in programmed cell death in certain species (Gallo et al., 2017). Sterol diversity has also been used as a "molecular fossil" to track important evolutionary events, such as molecular oxygen availability in the ocean–atmosphere system (Gold et al., 2017). Phytosterols are also considered useful natural products due to their wide range of biological applications (Luo et al., 2015). The principal therapeutic and nutraceutical properties of phytosterols include cholesterol-lowering (Ras et al., 2014), anti-inflammatory (Aldini et al., 2014) and anti-diabetic activities (Wang et al., 2017).

Diatoms derived from a secondary endosymbiosis between a red alga, a cyanobacterium and a phagotrophic eukaryote (Armbrust, 2009). They possess genes from two eukaryotic genomes, as well as genes of putative bacterial origin (Armbrust et al., 2004; Bowler et al.,

2008). It is likely that novel enzymes and regulatory features of sterol biosynthesis are responsible for the natural diversity of diatom sterol repertoires. While some organisms rely on the production of a single sterol type, such as cholesterol (animals) or ergosterol (fungi), plant and algal species produce and employ a greater variety of phytosterols. In diatoms, over 25 different sterols have been empirically identified in over one hundred diatom species (Rampen et al., 2010). The detailed biosynthesis of triterpenoids and sterols in diatoms is only partially understood (Jaramillo-Madrid et al., 2019).

Sterol precursors can be produced either by the mevalonate (MVA) pathway, or by the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (Lohr et al., 2012). In organisms that harbor both MVA and MEP pathways, such as plants, the MVA pathway generally provides the substrates for sterol biosynthesis. In the MVA pathway, three molecules of acetyl-CoA are transformed into isopentenyl diphosphate (IPP), which is then converted into dimethylallyl diphosphate (DMAPP) by an isopentenyl diphosphate isomerase (IDI, E.C. 5.3.3.2). IPP and DMAPP are condensed in geranyl diphosphate (GPP) which, together with another molecule of IPP, are converted to farnesyl diphosphate (FPP) by a farnesyl-diphosphate synthase (FPPS, 2.5.1.10). In a series of subsequent reactions, FPP is converted into squalene.

Squalene is the universal precursor of triterpenoids including sterols, sapogenins, and steroid hormones. The monooxygenation reaction that converts squalene into 2,3 - epoxy-squalene is considered a rate-limiting step in this pathway (Gill et al., 2011), and is carried out by either the squalene epoxidase (SQE, E.C. 1.14.14.17) in plants, animal and fungi, or an alternative SQE (AltSQE, E.C. 1.14.19.-) enzyme that is found in diatoms and several other organisms (Pollier et al., 2019). The subsequent committed step in sterol biosynthesis involves the cyclization of 2,3- epoxysqualene into either lanosterol or cycloartenol by an oxidosqualene cyclase (OSC, E.C. 5.4.99.7 or 5.4.99.8). In animals, 2,3-epoxysqualene is cyclized into lanosterol to yield cholesterol in a series of reactions (Cerqueira et al., 2016). Similarly, in fungi lanosterol is further converted into ergosterol. In contrast, the synthesis of sterols in plants occurs via cycloartenol (Fig 3.1) (Lohr et al., 2012). The sterol pathway of the pennate diatom *P. tricornutum* includes features of both plant and fungal metabolism (Fabris et al., 2014).

It is not clear to what extent the sterol biosynthesis pathways are conserved among different diatoms, and where in these pathways different species have diverged to produce an array of different sterols. Although sterol synthesis in the pennate diatom *P. tricornutum* has been characterized (Fabris et al., 2014), its genome differs significantly from other diatom species, such as the centric diatom *T. pseudonana*, which shares only about 57% of its genes as a result of ninety million years of divergent evolution (Bowler et al., 2008). Moreover, other large diatom clades such as the chaetocerids (e.g. *C. muelleri*) are metabolically unexplored.

In this study, three different diatom species were treated with chemical inhibitors of en-
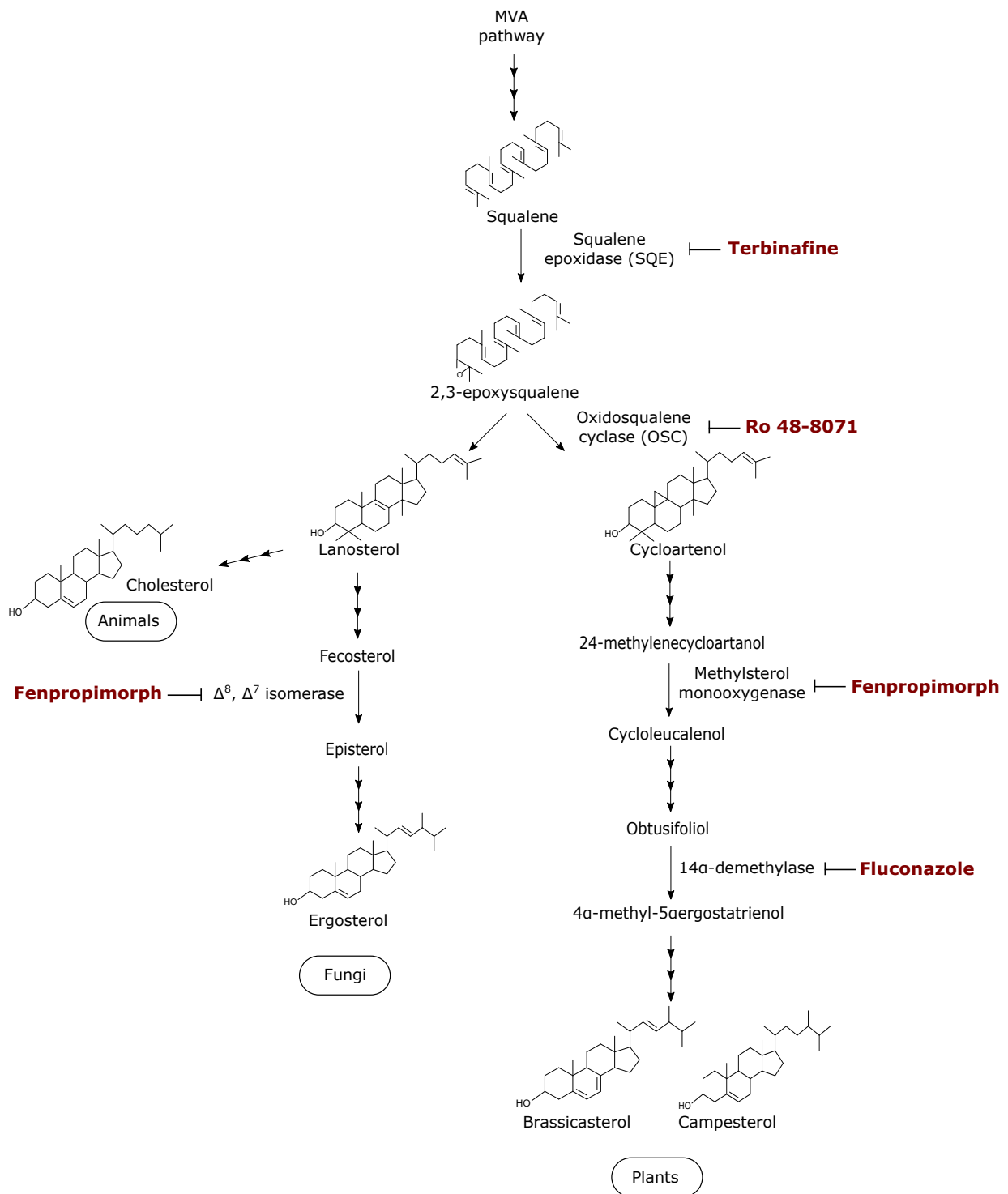
Figure 3.1: General sterol biosynthesis pathway in model eukaryotic organisms (plants, animals and fungi). Squalene epoxidase (E.C. 1.14.14.17); Oxidosqualene cyclase (E.C. 5.4.99.8); Methylsterol monooxygenase (E.C. 1.14.18.9); 14α-demethylase (E.C 1.14.14.154); Δ^8, Δ^7 isomerase (E.C. 5.3.3.5).

zymes participating in the biosynthesis of sterols. These consisted of *Phaeodactylum tricornutum*, a pennate diatom that has served as the model species for metabolic studies (Fabris et al., 2014), *Chaetoceros muelleri*, a representative of one of the largest genus of diatoms (*chaetocerids*) and a favored aquaculture species, and *Thalassiosira pseudonana*, a model centric species. Intermediate compounds arising from targeted inhibition were identified by GC–MS, and combined with genomic and transcriptomic information to empirically reconstruct the sterol biosynthesis pathways for each species.

3.2 Materials and Methods

3.2.1 Diatoms

The species *P. tricornutum* (CCMP632), *T. pseudonana* (CCMP1335) and *C. muelleri* (CCMP13-16) were obtained from the National Centre for Marine Algae and Microbiota (https://ncma.bigelow.org/cms/index/index/). The cells were maintained in L1 medium (Guillard & Hargraves, 1993) at $18\,^{\circ}C$ under continuous cool white light (100 μmol photons $m^{-2}\,s^{-1}$).

3.2.2 Dose/effect assay for sub-lethal effective concentrations of inhibitors.

To determine the concentration of the inhibitors (terbinafine, fenpropimorph, fluconazole and Ro 48-8071) needed to produce an effect on the growth of the three diatom species, a preliminary screening using different concentration of inhibitor was carried out in 96-well plates. The concentrations evaluated were 0.1, 1, 3, 10, 30, 100, 300 μM for terbinafine, fenpropimorph and fluconazole. For Ro 48-8071, the tested concentrations were 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 10 μM. The species *P. tricornutum*, *T. pseudonana* and *C. muelleri* were inoculated in triplicate in 96-well plates. When the culture reached 1×10^6 cell ml^{-1} the inhibitor was added (n=3). Changes in relative growth rates were estimated using a fluorescence plate reader (Infinite M100 pro, Austria), with excitation at 485 nm and emission at 680 nm. Chemicals terbinafine, fenpropimorph, fluconazole were obtained from Sigma-Aldrich, Ro 48-8071 was obtained from Cayman Chemical.

3.2.3 Growth and collection of diatoms for analysis under inhibited conditions.

Diatoms were grown in tubular photobioreactors containing 1.2 l of L1 medium at 21 $^{\circ}$C in continuous light (100 μmol photons $m^{-2}\,s^{-1}$) (Fig. 3.S1). When the cultures reached 1×10^6 cell ml^{-1} density, samples were treated in triplicate with Ro 48–8071 (2 μM), fluconazole (10 μM), fenpropimorph (10 μM) or terbinafine (10 μM) (n = 3). Samples were harvested

after 24, 48, 72 and 96 h for sterols and RNA extraction. Negative controls were set up by adding equivalent volumes of the respective solvent of each inhibitor: methyl acetate for Ro 48-8071; and dimethyl sulfoxide for terbinafine and fenpropimorph.

3.2.4 RNA extraction and transcriptome sequencing.

50 ml of algae culture were harvested by vacuum filtration. Filters were flash frozen and kept at -80 °C until extracted. Total RNA was extracted using Tri Reagent (Sigma Aldrich) and chloroform followed by isopropanol precipitation. Cells were re-suspended in 1 ml Tri Reagent and then allowed to sit at room temperature for 5 min. Then 0.25 ml of chloroform was added, mixed and incubated for 5 min. Samples were centrifuged at 4 °C, 16000 g for 15 min; the upper aqueous layer was removed to a new tube. 1.5 μl of GlycoBlue^{TM}(15 mg ml^{-1} , Invitrogen^{TM}) and 0.5 ml of isopropanol were added and mixed. After 5 min at room temperature, samples were centrifuged for 4 °C, 16000 g for 10 min. The pellet was washed once with 75% ethanol. Finally, the pellet was re-suspended in 50 μl RNase-free water. Traces of possible DNA carryover throughout the extraction were eliminated using the Invitrogen TURBO DNA-free^{TM}Kit. RNA samples were enriched for poly-adenylated mRNAs, barcoded sequencing libraries were generated, and paired end 2x100bp sequencing was performed on a single flow cell lane of an Illumina NovaSeq 6000 by the Ramaciotti Centre for Genomics and the University of New South Wales, Australia.

3.2.5 Transcriptome analysis and enzyme annotation.

De-multiplexed paired-end fastq mRNA-seq reads (NCBI SRA project number: PRJNA524-702) (n = 16, Table 3.S1) were quality-filtered using Trimmomatic (Bolger et al., 2014) as part of the Trinity de novo RNA-seq assembly software (version 2.8.4) (Grabherr et al., 2011). To rule out cross-contamination between species that might conflate ortholog analyses, reads for *C. muelleri* were pre-filtered against the published genomes of *P. tricornutum* (Bowler et al., 2008) and *T. pseudonana* (Armbrust et al., 2004) using HISAT2 (version 2.1.0) (Kim et al., 2015), resulting in the elimination of between 0.08% to 0.25% of reads by sample prior to assembly. Similarly, de novo assemblies of *P. tricornutum* and *T. pseudonana* were generated following read filtering against each other's genomes and the transcriptome of *C. muelleri*. Relative transcript abundances (normalized read counts per gene model, CPM) by read mapping and quantification using HISAT2 and Stringtie (version 1.3.3b) (Pertea et al., 2016). Differential expression was analyzed using EdgeR (version 3.24) (Robinson et al., 2010). Putative homologs between species were predicted using reciprocal best BLASTp (Altschul et al., 1997). Predictions of enzyme activities for putative protein sequences were performed using HMMsearch and HMMscan (Eddy, 1998) based on Hidden Markov Models (HMMs) built from a library of E.C.-based protein alignments (Desai et al., 2011; Kane-

hisa & Goto, 2000; Srivastava et al., 2007). Gene Ontology (GO) term enrichments based on existing annotations of *P. tricornutum* and *T. pseudonana* proteins were computed using a hypergeometric statistic in Python and corrected for multiple hypotheses using the Bonferroni method.

3.2.6 Extraction and analysis of sterols by GC-MS.

100 ml of algae culture was harvested at each time point by centrifuging at 4000 g for 10 min. The diatom pellet was washed with Milli-Q water to eliminate salt excess. The cell pellet was freeze-dried to determine dry matter weight. For sterol extraction, dry cell matter was transferred into a test tube, followed by the addition of 1 ml of 10% KOH ethanolic solution and then heated at 90 °C for 1 h. Once the sample was cold, the free sterols were extracted three times with 400 μl of hexane. An internal standard, 5a-cholestane, was added to each sample prior to hexane extraction. The hexane fraction was dried under a gentle N_2 stream and derivatized with 50 μl of 99% BSTFA + 1% TMCS at 70 °C for 1 h. Finally, 50 μl of hexane was added to each sample to increase sample volume for GC–MS injection.

Gas chromatography/mass spectrometry (GC–MS) analysis was performed using an Agilent 7890 series equipped with a HP-5 capillary column (30 m; 0.25 mm inner diameter, film thickness 0.25 μm) coupled to an Agilent quadrupole MS (5975 N) instrument. The following settings were used: oven temperature initially set at 50 °C with a gradient from 50 °C to 250 °C (15.0 °C min^{-1}), and then from 250 °C to 310 °C (8 °C min^{-1} , hold 10 min); injector temperature 250 °C; and carrier gas helium flow 0.9 ml min^{-1} . A split-less mode of injection was used, with a purge time of 1 min and an injection volume of 2 μl. Mass spectrometer operating conditions were as follows: ion source temperature 230 °C; quadrupole temperature 150 °C; accelerating voltage 200 eV higher than the manual tune; and ionization voltage 70 eV. Full scanning mode with a range from 50 to 650 Da was used.

Sterol peaks were identified based on retention time, mass spectra, and representative fragment ions compared to the retention times and mass spectra of authentic standards. The NIST (National Institute of Standards and Technology) library was also be used as reference. The area of the peaks and deconvolution analysis was carried out using the default settings of the Automated Mass Spectral Deconvolution and Identification System AMDIS software (v2.6, NIST). Peak area measurements were normalized by both the weight of dry matter prior to extraction, and the within-sample peak area of the internal standard 5α-cholestane. Sterol standards used to calibrate and identify GC–MS results in this study included: cholest-5-en-3-β-ol (cholesterol); (22E)-stigmasta-5,22-dien-3β-ol (stigmasterol); stigmast-5-en-3-β-ol (sitos-terol); campest-5-en-3-β-ol (campesterol); (22E)-ergosta-5,22-dien-3-β-ol (brassicasterol); (24E)-stigmasta-5,24-dien-3-ol (fucosterol); 9,19-Cyclo-24-lanosten-3β-ol (cycloartenol); 5-α-cholestane; and the derivatization reagent bis(trimethyl-silyl) trifluoroacetamide and tri-methylchlorosilane (99% BSTFA + 1% TMCS)

and were obtained from Sigma-Aldrich, Australia.

3.3 Results

3.3.1 Sub-lethal enzyme inhibition reveals shared sterol intermediates among diatoms

In the absence of an inhibitor, the diatoms *C. muelleri*, *T. pseudonana* and *P. tricornutum* naturally produced a unique combination of phytosterols: *P. tricornutum* produced brassicasterol and campesterol, *C. muelleri* produced 24-methylcholesta-5,24(24')-dien-3β-ol, fucosterol, isofucosterol and an unidentified sterol, and *T. pseudonana* produced 24-methylcholesta-5,24(24')-dien-3β-ol, campesterol, fucosterol and isofucosterol. Cholesterol was found in all three diatoms (Fig. 3.S2).

To identify intermediates in the sterol pathway, cultures of the diatoms *P. tricornutum*, *T. pseudonana* and *C. muelleri* were treated separately with four different chemical inhibitors of enzymes involved in sterol biosynthesis: terbinafine, Ro 48-8071, fenpropimorph and fluconazole. Each compound is known to inhibit different specific enzymes in the sterol metabolic pathways of other model organisms (Fig. 3.1). Cultures of yeast, plants and animals treated with terbinafine usually accumulate squalene (Nowosielski et al., 2011)]. Ro 48-8071-treated cells are expected to accumulate 2,3-epoxysqualene (Morand et al., 1997). Similarly, fenpropimorph and fluconazole inhibit enzymes at different steps in the sterol biosynthetic pathway (Fig. 3.1) (Campagnac et al., 2009).

In response to treatments with chemical inhibitors (terbinafine: 10 μM; Ro 48-8071: 2 μM; fenpropimorph: 10 μM; fluconazole: 10 μM) (Fig. 3.1; Fig. 3.S3), cultures of the diatoms *P. tricornutum*, *T. pseudonana* and *C. muelleri* exhibited various effects on their growth responses, sterols, and sterol intermediates (Figs. 3.3, 3.4, 3.5). Characteristic end-point sterols were still observed in cells treated with inhibitors, indicating a capacity for each diatom to overcome deliberate sub-lethal inhibition of enzymes in their sterol biosynthesis pathways (Fig. 3.4). In the presence of terbinafine, the end-point sterol brassicasterol was found in *P. tricornutum*, 24-methylcholesta-5,24(24')-dien-3β-ol in *T. pseudonana* and cholesterol in *C. muelleri*. The intermediate 2,3-epoxysqualene accumulated in Ro 48–8071-treated cultures, while squalene was only observed in *P. tricornutum* and *T. pseudonana*. Treatment with fenpropimorph lead to accumulation of several intermediates, including 31-norcycloartanol in *C. muelleri* and *P. tricornutum*, cycloartanol, 24-methylenecycloartanol and fecosterol in *P. tricornutum*, and cycloartenol in the three diatoms studied. Obtusifoliol accumulated in all fluconazole-treated cultures, as well as 4,14 dimethylcholesta-8,24-dienol and 14-methylcholesta-8,24-dienol in *C. muelleri*. The accumulation of fecosterol was observed in *P. tricornutum* during fenpropimorph treatment, but not in *T. pseudonana* or *C. muelleri* (Figs. 3.4, 3.S6, 3.S7). Fecosterol is converted into episterol by a Δ^8, Δ^7 isomerase

(E.C. 5.3.3.5), which is known to be inhibited by fenpropimorph (Campagnac et al., 2009). Weak putative homologs (HMM E-value > 1e-50) for a putative Δ^8, Δ^7 isomerase found in fungi (ERG2) were found in all three diatoms (Table 3.1).

Table 3.1: Proteins containing enzyme domains involved in upstream reactions and sterol biosynthesis. "Orthogroup" refers to separately homologous groups of gene models between species, as predicted by sequence similarity. Numbers in parentheses are E-values for profile-based similarity to known enzyme functions.

E.C. number	Enzyme function	Orthogroup	*C. muelleri*	*P. tricornutum*	*T. pseudonana*
MEP pathway					
2.2.1.7	1-deoxy-D-xylulose-5-phosphate synthase	1	TRINITY_DN1029_c3_g1_i2.p1 (9e-247)	Phatr3_draftJ1689 (3e-244)	574 (3e-248)
		2	TRINITY_DN1029_c3_g1_i3.p1 (1e-220)		
		3	TRINITY_DN17006_c0_g1_i1.p1 (1e-210)		
1.1.1.267	1-deoxy-D-xylulose-5-phosphate reductoisomerase	1	TRINITY_DN2155_c1_g2_i1.p1 (3e-150)		
		2	TRINITY_DN2155_c1_g2_i2.p1 (3e-150)		
		3	TRINITY_DN2155_c1_g2_i3.p1 (3e-150)	Phatr3_J9258 (2e-149)	10943 (3e-151)
2.7.7.60	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase	1	TRINITY_DN801_c9_g1_i1.p1 (3e-70)	Phatr3_J21829 (2e-71)	3622 (2e-73)
2.7.1.148	4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol kinase	1	TRINITY_DN1072_c1_g1_i1.p1 (8e-83)		31907 (2e-84)
		2		Phatr3_EG02383 (5e-82)	
4.6.1.12	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	1	TRINITY_DN1462_c1_g1_i2.p1 (7e-64)	Phatr3_J12330 (2e-61)	37585 (4e-64)
1.17.7.1	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate-synthase (ferredoxin)	1	TRINITY_DN800_c1_g1_i1.p1 (8e-139)	Phatr3_J44955 (4e-136)	29228 (6e-137)
		2	TRINITY_DN800_c1_g1_i2.p1 (3e-128)		
1.17.7.4	4-hydroxy-3-methylbut-2-en-1-yl–diphosphate reductase	1	TRINITY_DN1242_c0_g1_i1.p1 (3e-96)	Phatr3_J41845 (1e-97)	28326 (7e-93)
		2	TRINITY_DN18902_c0_g1_i1.p1 (3e-83)		
		3	TRINITY_DN1242_c0_g1_i2.p1 (4e-71)		
MVA pathway					
2.3.1.9	acetyl-CoA C-acetyltransferase (AACT)	1	TRINITY_DN7658_c0_g1_i2.p1 (9e-177)	Phatr3_J23913 (7e-162)	28651 (2e-156)
		2	TRINITY_DN1971_c0_g2_i2.p1 (4e-124)	Phatr3_J45947 (2e-125)	34809 (2e-128)
		3	TRINITY_DN1971_c0_g2_i1.p1 (2e-94)		
		4	TRINITY_DN7658_c0_g1_i1.p1 (7e-72)		
2.3.3.10	hydroxymethylglutaryl-CoA synthase (HMGS)	1	TRINITY_DN1137_c0_g1_i9.p1 (2e-35)*		
		2		Phatr3_J16649 (1e-148)	
		3			260768 (2e-145)
1.1.1.34	hydroxymethylglutaryl-CoA reductase (HMGR)	1	TRINITY_DN9859_c0_g1_i1.p1 (4e-32)*		
		2		Phatr3_J16870 (6e-213)	33680 (3e-203)
2.7.1.36	mevalonate kinase (MVK)	1	TRINITY_DN1691_c0_g2_i2.p1 (1e-22)*		
		2		Phatr3_J53929 (8e-70)	
		3			2376 (6e-44)*
2.7.4.2	phosphomevalonate kinase (PMK)	1		Phatr3_EG01590 (1e-62)	21866 (5e-43)*
4.1.1.33	diphosphomevalonate decarboxylase (MVD)	1	TRINITY_DN18672_c0_g1_i1.p1 (4e-33)*		
		2		Phatr3_EG02359 (6e-136)	15226 (1e-144)
5.3.3.2	isopentenyl-diphosphate Δ-isomerase (IDI)	1	TRINITY_DN24_c3_g1_i1.p1 (1e-58)	Phatr3_EG02290 (2e-61)	
		2		Phatr3_J12533 (3e-62)	
		3			31266 (1e-59)
2.7.4.26	isopentenyl phosphate kinase	1	TRINITY_DN463_c0_g1_i17.p1 (1e-24)*	Phatr3_EG01825 (7e-39)*	11228 (6e-26)*
Central steps of isoprenoids biosynthesis					
2.5.1.1	dimethylallyltranstransferase	1	TRINITY_DN4739_c0_g1_i1.p1 (1e-108)		
		2	TRINITY_DN4739_c0_g1_i2.p1 (1e-107)	Phatr3_J49325 (7e-107)	30105 (3e-88)
2.5.1.10	farnesyl-diphosphate synthase	1	TRINITY_DN8889_c0_g1_i1.p1 (2e-86)	Phatr3_J19000 (1e-84)	268480 (1e-85)
Sterol biosynthesis					
2.5.1.21	squalene synthase	1	TRINITY_DN24_c3_g1_i1.p1 (4e-124)	Phatr3_EG02290 (2e-119)	268488 (7e-127)
		2	TRINITY_DN24_c3_g1_i2.p1 (4e-103)		
1.14.19.-	AltSQE	1	TRINITY_DN501_c18_g1_i1.p1 ()	Phatr3_J45494 ()	10850 ()

Table 3.1 ...continued

E.C. number	Enzyme function	Orthogroup	C. muelleri	P. tricornutum	T. pseudonana
5.4.99.8	cycloartenol synthase	1	TRINITY_DN929_c1_g2_i1.p1 (2e-244)	Phatr3_EG02293 (5e-228)	277 (1e-233)
2.1.1.41	sterol 24-C-methyltransferase	1	TRINITY_DN649_c0_g1_i4.p1 (6e-18)*		
		2		Phatr3_J10824 (2e-106)	36203 (3e-107)
		3			32029 (2e-106)
1.14.18.9	methylsterol monooxygenase	1	TRINITY_DN964_c3_g1_i2.p1 (4e-23)*	Phatr3_J37371 (2e-21)*	
		2	TRINITY_DN47_c0_g3_i4.p1 (3e-14)*		
		3		Phatr3_J10852 (1e-16)*	
		4			260933 (2e-22)*
1.3.1.72	Δ^{24}-sterol reductase; lanosterol Δ^{24}-reductase	1	TRINITY_DN3357_c0_g1_i1.p1 ()		
		2		Phatr3_EG02562 ()	
		3			268544 (1e-13)*
5.5.1.9	cycloeucalenol cycloisomerase	1		Phatr3_J49447 (1e-96)	7708 (5e-88)
1.14.14.154	sterol 14α-demethylase	1	TRINITY_DN985_c1_g2_i1.p1 (6e-136)	Phatr3_J31339 (7e-134)	33926 (1e-132)
5.3.3.5	Δ^{8}, Δ^{7} isomerase	1	TRINITY_DN1738_c2_g1_i1.p1 ()		10157 ()
		2		Phatr3_J36801 ()	
1.14.19.20	Δ^{7}-sterol 5(6)-desaturase	1	TRINITY_DN501_c18_g1_i1.p1 (2e-27)*		10850 (1e-23)*
		2		Phatr3_J14208 (1e-66)	
1.14.19.41	sterol 22-desaturase	1		Phatr3_J51757 (4e-102)	
1.3.1.70	Δ^{14}-sterol reductase	1	TRINITY_DN1204_c0_g1_i1.p1 (3e-140)	Phatr3_J48260 (3e-145)	22186 (9e-140)
		2			26769 (4e-63)
1.3.1.21	7-dehydrocholesterol reductase	1		Phatr3_J30461 (2e-118)	33794 (1e-93)
2.7.4.26	isopentenyl phosphate kinase	1	TRINITY_DN463_c0_g1_i17.p1 (1e-24)*	Phatr3_EG01825 (7e-39)*	11228 (6e-26)*
1.1.1.170	3β-hydroxysteroid-4alpha-carboxylate 3-dehydrogenase	1	TRINITY_DN3323_c1_g1_i1.p1 (7e-56)	Phatr3_J48864 (3e-53)	4552 (9e-54)
1.3.1.71		1			26769 (1e-181)

*E-value higher than 1e-50

3.3.2 A terbinafine-insensitive alternative squalene epoxidase is a common feature in diatoms

In this study, squalene did not accumulate in three different diatom species treated with terbinafine (10 μM) (Figs. 3.3, 3.S4). This is consistent with previous results for *P. tricornutum*, in which no accumulation of squalene was observed after the addition of up to 600 μM of terbinafine (Fabris et al., 2014). Similarly in *C. muelleri*, a putative AltSQE with high sequence similarity to those found of *P. tricornutum* and *T. pseudonana* was identified, but not a conventional SQE (Table 3.1). Although terbinafine treatments did not cause characteristic signs of SQE inhibition, all three species showed minor alterations in their sterol compositions, suggesting additional potential non-SQE targets of terbinafine (Figs. 3.4, 3.S 3, 3.S4, 3.S5a).

To confirm the occurrence of the squalene epoxidation reaction in the sterol pathways of *C. muelleri* and *T. pseudonana*, the inhibitor Ro 48-8071 was used to selectively block the enzyme OSC, which catalyzes the conversion of 2,3-epoxysqualene either into lanosterol or cycloartenol (Morand et al., 1997). Treatment with Ro 48-8071 resulted in the accumulation of 2,3-epoxysqualene in all three diatom species, and squalene accumulation was observed in *P. tricornutum* and *C. muelleri* (Figs. 3.3, 3.S5c). These results confirm that 2,3-

epoxysqualene is a precursor in the biosynthesis of sterols in *C. muelleri* and *T. pseudonana*, similarly to *P. tricornutum* (Fabris et al., 2014), and that this intermediate is cyclized by a conventional OSC that is inhibited by Ro 48-8071.

3.3.3 Sterols are produced via cycloartenol in three model diatom species

Sterols can be produced via either lanosterol, as in animals, fungi and yeast, or via cycloartenol, as in plants (Nes, 2011). Here, treatment with fenpropimorph resulted in the accumulation of cycloartenol in all the three diatoms (Figs. 3.3, 3.S6), confirming that the sterol biosynthesis pathways proceed via cycloartenol intermediates. This is consistent with previous findings for *P. tricornutum* (Fabris et al., 2014), but it is not known whether this generalizes to additional diatoms. Diatoms treated with fluconazole also accumulated obtusifoliol (Figs. 3.3, 3.S5b, 3.S7), which is an intermediate downstream of cycloartenol in diatom and plant sterol biosynthesis (Jaramillo-Madrid et al., 2019).

An oxidosqualene cyclase (OSC, E.C. 5.4.99.-) that converts 2,3-epoxysqualene into either cycloartenol (E.C. 5.4.99.8) or lanosterol (E.C. 5.4.99.7) is present in all sterol-producing eukaryotic species (Summons et al., 2006). Predicted orthologs of OSC in all three species are listed in Table 3.1. The product specificities of oxidosqualene cyclases (OSC) are also determined by amino acid identities in the active site of the enzyme. Lanosterol synthases are characterized by the presence of the amino acids T381, C/Q449, and V453, while cycloartenol synthases possess Y381, H449, and I453 (Summons et al., 2006) (Fig. 3.S9c). Position 381 can be variable; only position 453 is indicative of specificity for lanosterol (V453) or cycloartenol (I453) (Desmond & Gribaldo, 2009). In accordance with these patterns *P. tricornutum*, *T. pseudonana* and *C. muelleri* appear to utilise a cycloartenol synthase (Fig. 3.S9c).

3.3.4 Cholesterol is produced via a cycloartenol-dependent pathway in diatoms

The only common sterol detected in un-treated cultures of the three diatoms evaluated in this study was cholesterol, which has been found in more than fifty diatom species (Rampen et al., 2010). *C. muelleri* accumulated cholesterol as its principal sterol (Fig. 3.5). It is not known how diatoms produce cholesterol, as it is normally produced via lanosterol (Summons et al., 2006). In addition to cycloartenol, the accumulation of putative cycloartanol and 31-norcycloartanol intermediates was detected during fenpropimorph treatment (Figs. 3.5, 3.S6e, 3.S7). Fluconazole treatment resulted in the presence of putatively identified 4,14 dimethylcholesta 8,24-dienol and 14-methylcholesta 8,24-dienol (Figs. 3.S6d,e; 3.S7). Cholesterol accumulation decreased in the presence of Ro 48-8071 treated cells, suggesting that 2,3-epoxysqualene is a shared precursor in sterols synthesis and cholesterologenesis

(Fig. 3.5). Similarly, cholesterol content diminished in the presence of terbinafine, which presumably has off-target effects over other enzymes participating in sterol biosynthesis and cholesterologenesis. These results suggest potential branching points for the production of cholesterol by some diatoms similar to the one occurring in plants for which cholesterol is synthesized via cycloartenol. A hypothetical model of cholesterol metabolism in diatoms is presented in Fig. 3.5.

3.3.5 Transcriptional dynamics of genes involved in sterol biosynthesis

The blockage of the essential sterol biosynthesis pathway by chemical inhibition caused identifiable metabolic responses in each diatom. To examine gene regulatory responses accompanying this metabolic shift, we analyzed the changes in transcript abundance that resulted from these treatments. Changes occurred for many genes putatively involved in the sterol biosynthesis pathways of all three diatom species upon treatment. The changes in abundance of transcripts encoding known functions in response to the inhibition of sterol biosynthesis are shown in Fig. 3.S8. Putatively orthologous genes shared between all three species that encode functions involved in mevalonate-independent isopentenyl diphosphate synthesis (MEP), carotenoid biosynthesis, and steroid biosynthesis were among those apparently down-regulated in response to blockades of downstream metabolic flux. Potential stress-related functions including proteolysis and the production of polyamines were up-regulated (Fig. 3.S8), indicating an active and adaptive response by these diatoms to metabolic stress.

Within biosynthesis pathways specific to the production of sterols and their precursors, transcripts encoding enzymes that participate in the MEP pathway decreased in relative abundance during the inhibition of sterol biosynthesis in all three diatoms (Fig. 3.2a), while patterns of transcription in the MVA pathway were mosaic and species-specific. Transcripts encoding the enzyme acetyl-CoA C-acetyltransferase (E.C.2.3.1.9), which catalyzes the first step of the MVA pathway, increased in relative abundance in all three diatoms, as did transcripts for the rate-limiting enzyme 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR, E.C. 1.1.1.34). Transcripts encoding enzymes that catalyze the four reactions responsible for converting IPP into squalene all decreased in relative abundance, which is consistent with down-regulation of these genes in response to observed accumulations of squalene and cycloartenol intermediates during inhibition of downstream reactions. In *P. tricornutum* and *T. pseudonana*, transcripts encoding the alternative squalene epoxidase AltSQE increased in relative abundance in response to the inhibition of sterol biosynthesis, whereas the opposite effect was observed in *C. muelleri* (Fig. 3.3).
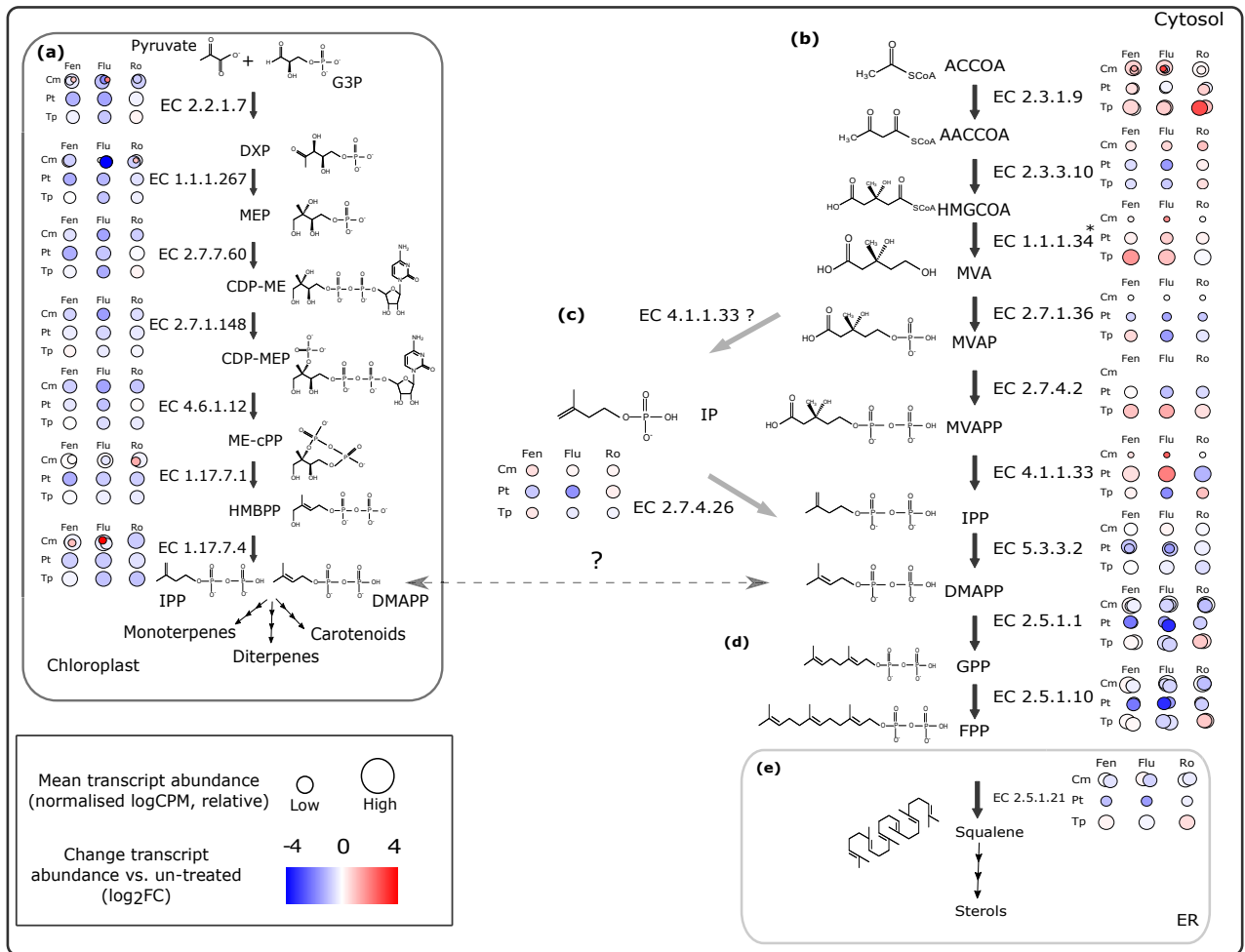
Figure 3.2: Upstream reactions in the sterol biosynthesis pathway of diatoms. (a) MEP Pathway: G3P glyceraldehyde 3-phosphate; DXP: 1-deoxy-D-xylulose-5-phosphate; MEP: 2-C-methyl-D-erythritol-4-phosphate; CDP-ME: 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol; CDP-MEP: 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol; ME-cPP: 2-C-methyl-D-erythritol-2,4-cyclodiphosphate; HMBPP: 1-hydroxy-2-methyl-2butenyl-4-diphosphate. (b) Mevalonate (MVA) Pathway: ACCOA: acetyl-CoA; AACCOA: aceto-acetyl-CoA; HMGCOA: 3-hydroxy-3-methylglutaryl-coenzyme A; MVA: mevalonate; MVAP: mevalonate phosphate; MVAPP: mevalonate diphosphate; IPP: isopentenyl diphosphate; DMAPP: dimethylallyl diphosphate. HMGR (1.1.1.34*) is a membrane protein, anchored in the endoplasmic reticulum; all the other six enzymes involved in the MVA pathway are soluble proteins (Lohr et al., 2012). (c) Alternate MVA Pathway, found in archaea: IP: isopentenyl phosphate (Dellas et al., 2013). (d) Isoprenoids Biosynthesis: GPP: geranyl diphosphate; FPP: farnesyl diphosphate. Dashed line indicates putative crosstalk between MEP and MVA pathway. (e) Sterol biosynthesis: putatively located in the endoplasmic reticulum (ER). Relative transcript abundances (CPM) are shown as shaded circles, with log_2 fold-changes (log_2 FC) relative to untreated control cultures shown in blue (decreased) and red (increased) (n = 16). The sizes of these circles are proportional to average logCPM over all samples. Multiple circles indicate multiple putative genes predicted to encode each respective enzyme function. Cm: *Chaetoceros muelleri*; Pt: *Phaeodactylum tricornutum*; Tp: *Thalassiosira pseudonana*. C: control cultures (no inhibitor added); Fen: Fenpropimorph; Flu: Fluconazole, Ro: Ro 48-8071. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.3.6 Reconstruction of sterol precursor biosynthesis pathways

Functional MVA and MEP pathways have been identified in the diatom *P. tricornutum*, with MVA producing precursors for sterol biosynthesis (Cvejić & Rohmer, 2000). While we did not profile isoprenoid sterol precursors in our experiments, a model of the MVA and MEP pathways based on comparative genomics and transcriptomics is presented in Fig. 3.2. We identified homologs for the seven genes involved in the MEP pathway in each diatom (Table 3.1, Fig. 3.2b).

Orthologs of an IDI-SQS fusion enzyme (IDI, E.C. 5.3.3.2/SQS, E.C.2.5.1.21) that catalyzes the isomerization of IPP to DMAPP in *P. tricornutum* and *Haslea ostrearia* (Athanasakoglou et al., 2019; Fabris et al., 2014) were found in the transcriptomes for *T. pseudonana* and *C. muelleri*, in agreement with the discovery of this novel fusion gene in numerous microalgae (Ferriols et al., 2017). Additional transcripts that contain putative single IDI domains were observed in the diatoms *P. tricornutum* and *T. pseudonana* (Table 3.1, Fig. 3.2d: E.C. 5.3.3.2, Fig. 3.S9a,b). The conversion of FPP into squalene is also catalyzed by this IDI-SQS fusion enzyme (Athanasakoglou et al., 2019; Fabris et al., 2014; Ferriols et al., 2017).

We detected three genes encoding enzymes consistent with prenyltransferase activity in *C. muelleri*, and two each in *P. tricornutum* and *T. pseudonana* (Table 3.1, Fig. 3.2d). Plants employ geranyl diphosphate synthases (GPPS, E.C. 2.5.1.1) to produce GPP from DMAPP and farnesyl-diphosphate synthases (FPPS, E.C. 2.5.1.10) to produce FPP from GPP (Tholl, 2015). In yeast, a single enzyme (ERG20) exhibits both catalytic activities (Klug & Daum, 2014). In our experiments, the relative abundances of transcripts encoding diatom prenyltransferases generally decreased during the inhibition of sterol biosynthesis, with the exception *T. pseudonana* cells treated with Ro 48-8071 (Fig. 3.2d).

3.4 Discussion

3.4.1 A conserved metabolic core for the biosynthesis of sterols in diatoms

In this study, we propose that sterol biosynthesis in diatoms consists of a commonly conserved core, that is characterized by the presence of a terbinafine-insensitive alternative squalene epoxidase (AltSQE) and the conversion of 2,3-epoxysqualene into cycloartenol (Fig. 3.3). This agrees with the prior discoveries that a cycloartenol producing OSC and an AltSQE exists in *P. tricornutum* and numerous other microeukaryotes (Fabris et al., 2014; Pollier et al., 2019). The diatom AltSQE belongs to the fatty acid hydroxylase superfamily, and differs from the conventional flavoprotein SQE used by plants, animals and fungi (Pollier et al., 2019). We confirmed its presence and conservation in other diatom species such as *C. muelleri* (Table 3.1). SQE has been reported to be a rate-limiting step of sterols synthe-
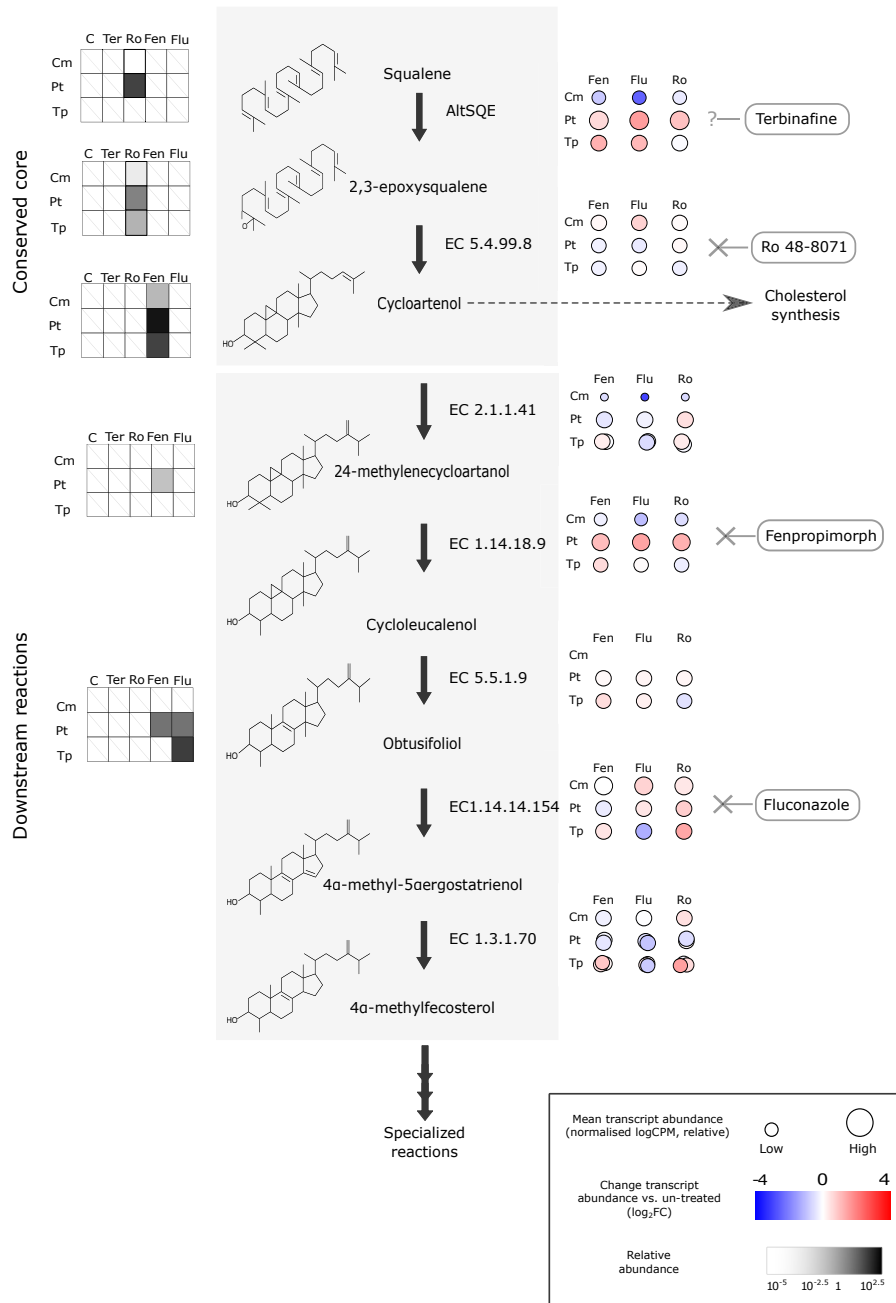
Figure 3.3: Conserved core and shared downstream reactions of the sterol biosynthesis pathway in diatoms. Grey scale heatmaps indicate relative abundances of sterol and intermediate compounds identified by GC–MS, expressed as peak area normalized on a per-sample basis to biomass and an internal standard. Presented sterol level value is the average between the three replicates, n = 3. Relative transcript abundances (CPM) are shown as shaded circles, with log_2 fold-changes (log_2 FC) relative to untreated control cultures shown in blue (decreased) and red (increased) (n = 16). The sizes of these circles are proportional to average logCPM over all samples. Multiple circles indicate multiple putative genes predicted to encode each respective enzyme function. Cm: *Chaetoceros muelleri*; Pt: *Phaeodactylum tricornutum*; Tp: *Thalassiosira pseudonana*. C: control cultures (solvent without inhibitor added); Fen: Fenpropimorph; Flu: Fluconazole, Ro: Ro 48-8071. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sis (Gill et al., 2011). In this study, changes in transcription were observed for AltSQE as a result of pathway inhibition. This may indicate that diatoms respond to the depletion or absence of final sterols through the up or down-regulation of AltSQE.

Diatoms produce a wide diversity of sterols, including cholesterol, which animals and fungi normally produce via a lanosterol precursor. Photosynthetic organisms generally produce sterols via cycloartenol (Nes, 2011); however, the detection of lanosterol in the diatom Haslea sp. (Rampen et al., 2010) makes this generalization unclear. In accordance with conserved residues found in the OSC enzymes of various species (Fig. 3.S9c) and accumulation of cycloartenol during fenpropimorph treatment, *T. pseudonana* and *C. muelleri* are predicted to utilise a cycloartenol synthase, as previously demonstrated for *P. tricornutum* (Fabris et al., 2014). These results support the theory that cycloartenol-based sterol biosynthesis is a predominant feature in diatoms, which may be further explored in additional species.

3.4.2 Species-specific sterol repertoires diversify downstream of a common pathway

Due to common intermediate metabolites observed during the inhibition of sterol biosynthesis, the diversification of species-specific sterol products appears to occur downstream of a common set of reactions. We propose that in all three of the clade-specific diatoms tested, cycloartenol is converted into 4α-methylfecosterol via obtusifoliol in five common enzymatic reactions, as in higher plants, and that species-specific bifurcation of end-point products proceeds downstream of 4α-methylfecosterol.

The existing model for sterol biosynthesis in *P. tricornutum* suggests that 4α-methylfecosterol is processed into sterols via a route similar to that by which fungi produce ergosterol from fecosterol (Fabris et al., 2014). Conversion of fecosterol into episterol is catalyzed by the enzyme Δ^8, Δ^7 isomerase, which is known to be inhibited by the chemical fenpropimorph (Campagnac et al., 2009). In the present study, neither fecosterol nor ergosterol accumulation was detected for the diatoms *T. pseudonana* and *C. muelleri*, suggesting that end-point sterol synthesis occurs differently in these two diatom species as compared to *P. tricornutum*. This may be consistent with ancient metabolic divergence between pennate and centric diatoms (Bowler et al., 2008). As in higher plants, it is possible that these diatoms convert 4α-methylfecosterol into 4α-methylene-lophenol, leading to episterol for the production of 24-methylcholesta-5,24(24')-dien-3β-ol, or to 14-ethylidene-lophenol for the synthesis of isofucosterol. However, putative enzymes found in the diatoms *T. pseudonana* and *C. muelleri* for the catalysis of these steps were only weakly predicted on the basis of profile-based HMMs and sequence similarities (Fig. 3.4, Table 3.1). In contrast to the detection of fungal-like features in *P. tricornutum* (Fabris et al., 2014), the diatoms *T. pseudonana* and *C. muelleri* seem to possess features more consistent with plant sterol biosynthesis.

In plants, isofucosterol is converted into sitosterol, but this sterol was not detected in any of the diatoms evaluated. Thus, we propose that *T. pseudonana* and *C. muelleri* convert isofucosterol into fucosterol, though an enzyme that catalyzes this isomeric conversion has not been identified. Campesterol was also detected in *T. pseudonana*, which may be explained by the activity of a $\Delta24(24')$-sterol reductase (E.C.1.3.1.71) that converts 24-methylcholesta-5,24($24'$)-dien-3β-ol into campesterol, as in plants. *P. tricornutum* in contrast converts ergosterol into campesterol using a Δ^7 sterol reductase (E.C. 1.3.1.21) (Fig. 3.4). The lack of C-22 desaturated sterols in *T. pseudonana* and *C. muelleri* is consistent with the absence of sterol 22-desaturases (E.C. 1.14.19.41) in the predicted proteins of these species. This may reflect a shared feature among centric diatoms, and/or a unique metabolic adaptation in the pennate or Phaeodactylum lineages.

3.4.3 Cholesterol biosynthesis in diatoms shares features with both plants and animals

Cholesterol is the main sterol found in mammals. The presence of cholesterol in at least fifty diatom species is intriguing (Rampen et al., 2010), since lanosterol, the conventional precursor for the formation of cholesterol, has not been widely detected in diatoms. Presence of cholesterol was also observed in the three diatoms analyzed in this study. Mammals rely solely on lanosterol for the production of cholesterol (Cerqueira et al., 2016). Plants have been reported to produce traces of cholesterol by a divergent pathway from phytosterol biosynthesis (Sonawane et al., 2017). In this study, we present a hypothetical pathway for the synthesis of cholesterol in diatoms that derives from cycloartenol, and shares characteristics with cholesterogenesis in plants and synthesis of ergosterol in fungi.

Conversion of cycloartenol into cycloartanol is the first committed step involved in cholesterol metabolism in tomato and potato plants, carried out by a sterol side chain reductase 2 (SSR2) (Sawai et al., 2014). Due to significant levels of cycloartanol observed in fenpropimoph-treated cultures (Fig. 3.5), we hypothesize that a similar reaction is occurring in diatoms by a Δ-24 sterol reductase (E.C. 1.3.1.72).

In this study, *C. muelleri* produced cholesterol abundantly. Homologs of the Δ-24 sterol reductase (E.C. 1.3.1.72) were found in *C. muelleri* and *P. tricornutum* (Table 3.1), as well as a less confident candidate in *T. pseudonana* (Thaps 268544). Cycloartanol is proposed to be converted into 31-norcycloartanol by a methylsterol monooxygenase (E.C. 1.14.18.9) that is sensitive to fenpropimorph (Campagnac et al., 2009). It is likely that 31-norcycloartenol levels were detected due to the effect of fenpropimorph on the enzyme Δ^8, Δ^7 isomerase (E.C. 5.3.3.5), which is further down in the metabolic pathway.

The initial three steps of cholesterogenesis in diatoms are similar to the ones reported for plants (Sonawane et al., 2017). In the model proposed in this study, 31-nor-24(25)-dihydrolanosterol is transformed into 4,14-di-methylcholesta-8,24-dienol and 14-methylcho-

lesta-8,24-dienol, both of which were observed in fluconazole-treated cells. Detailed hypothetical model is presented in Fig. 3.5.

3.4.4 Divergence and dynamics in diatom sterol precursor pathways

Transcriptomics performed in this study confirms that both diatoms *P. tricornutum* and *T. pseudonana* possess a complete and functional MVA and MEP pathway, as previously reported for *P. tricornutum* (Cvejić & Rohmer, 2000). The *C. muelleri* transcriptome contains a complete set of the enzymes participating in the MEP pathway, but not all enzymes to comprise a complete MVA pathway were detected. While false negatives for gene identification are possible, the depth and completeness of the *C. muelleri* transcriptome was comparable to that of *T. pseudonana*, *P. tricornutum*, and other marine eukaryotes (Tables 3.S1, 3.S2).

It is possible that *C. muelleri* contains additional non-conventional enzymes that catalyze conserved reactions in the MVA pathway, or that this species relies on the MEP pathway alone to provide isoprenoid precursors. The diatom H. ostrearia has also been reported to synthesize its main sterol (24-ethylcholest-5-en-3-ol) using precursors supplied by the MEP pathway (Cvejić & Rohmer, 2000; Massé et al., 2004). No homolog for the enzyme phosphomevalonate kinase (PMK, E.C. 2.7.4.2) was detected in *C. muelleri*, suggesting that this reaction might be catalyzed by a divergent enzyme. PMK catalyzes an important and conserved step for the conversion of acetyl coenzyme A (ACCOA) to IPP and DMAPP, via MVA. A confident PMK ortholog was also not detected in other microalgal species including Micromonas pusilla and Nannochloropsis oceanica, which were reported to lack a functional MVA pathway (Lu et al., 2014; Sasso et al., 2012). An alternative MVA pathway for the formation of isoprenoids precursors in archaea has been described (Dellas et al., 2013). In the latter case, the enzyme phosphomevalonate decarboxylase (MPD) converts mevalonate 5-phosphate (MVAP) into isopentenyl phosphate (IP), which is subsequently converted into IPP by the enzyme isopentenyl phosphate kinase (IPK, E.C. 2.7.4.26). Homologs of the enzyme IPK were detected in the three diatoms tested. However, no homologs for the enzyme MPD were found (Table 3.1).

Although gene expression data alone cannot indicate or explain metabolic fluxes, the reduction of transcript abundances in the MEP pathway for all species during the inhibition of sterol biosynthesis (Fig. 3.2a) suggests that the production of isoprenoid precursors is sensitive to changing metabolic conditions. Were these changes in transcript abundance to result in altered levels of metabolic enzymes, it is possible that cross-talk between pathways or cellular compartments may be affected. Studies in plants have shown that cross-talk between MVA and MEP pathways is feasible (Bick & Lange, 2003; Hemmerlin et al., 2003), and that IDD and DMAPP produced via the MVA pathway are mostly used to produce sterols in the cytosol (Kirby & Keasling, 2009). We presume that isoprenoid products of the MVA pathway accumulated or were imported into the plastids as a consequence of blocked

Conserved core reactions

Downstream reactions

4α-methylfecosterol

C. muelleri
T. pseudonana

EC 5.3.3.5

EC 2.1.1.143

24-methylene-lophenol

14-ethylidene-lophenol

EC 1.14.18.9

7-avenasterol

EC 1.14.19.20

5-dehydroavenasterol

EC 1.3.1.21

Isofucosterol

?

Fucosterol

EC 1.14.18.9

Episterol

EC 1.14.19.20

5-dehydroepisterol

EC 1.3.1.21

24-methylcholesta-5,24(24')-dien-3β-ol

EC 1.1.1.170

Fecosterol

EC 5.3.3.5

Episterol

EC 1.14.19.20

5,7,24(24')-ergostatrienol

EC 1.14.19.41

5,7,22,24(24')-ergostatetraenol

EC 1.3.1.70

Ergosterol

EC 1.3.1.21

Campesterol

EC 1.14.19.41

Brassicasterol

EC 1.3.1.71

P. tricornutum

Mean transcript abundance
(normalised logCPM, relative)

Low   High

Change transcript abundance vs. un-treated
(log$_2$FC)

-4   0   4

Relative abundance

10^{-5} $10^{-2.5}$ 1 $10^{2.5}$

Figure 3.4: Specialized reactions in sterol biosynthesis for the pennate diatom *P. tricornutum* (Fabris et al., 2014) and the centric diatoms *C. muelleri* and *T. pseudonana* proposed in this study. Grey scale heatmaps indicate relative abundances of sterol and intermediate compounds identified by GC–MS, expressed as peak area normalized on a persample basis to biomass and an internal standard. Presented sterol level value is the average between the three replicates, n = 3. Relative transcript abundances (CPM) are shown as shaded circles, with log 2 fold-changes (log_2 FC) relative to untreated control cultures shown in blue (decreased) and red (increased) (n = 16). The sizes of these circles are proportional to average logCPM over all samples. Multiple circles indicate multiple putative genes predicted to encode each respective enzyme function. Cm: *Chaetoceros muelleri*; Pt: *Phaeodactylum tricornutum*; Tp: *Thalassiosira pseudonana*. C: control cultures (solvent without inhibitor added); Fen: Fenpropimorph; Flu: Fluconazole, Ro: Ro 48-8071. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sterol synthesis, and that the MEP pathway was down-regulated in response. However, knowledge of exchange between compartmentalized metabolic pathways is incomplete (Zhang et al., 2009), and adequate measurements of intercompartmental flux have not been obtained. Changes in transcriptional levels of MEP pathway might also be a general response to stress caused by background effects of the chemical inhibitors used, as in plants (Xiao et al., 2012).

Transcriptional changes in enzymes such as HMGR (E.C. 1.1.1.34) and AltSQE, which in other organisms participate in sterol homeostasis, suggest that a sterol regulation system is operating in diatoms. Treatment of *P. tricornutum* with ethynylestradiol, which affects enzymes involved in TAG and sterol metabolism, lead to up-regulation of the enzyme hydroxymethylglutaryl-CoA synthase (HMGS, E.C.1.3.3.10) and down-regulation mevalonate kinase (MVK, E.C. 2.7.1.36). The same transcriptional patterns were observed in this study during Ro 48–8071 treatment in the three diatoms tested (Conte et al., 2018). Similarly, the genetic perturbation in the sterol pathway of *P. tricornutum* through the introduction of a heterologous OSC led to a reduction of expression levels of enzymes participating in the MVA and MEP pathways (D'Adamo et al., 2019).

Another key feature of the biosynthesis of terpenoids in diatoms is the presence of a protein fusion between IDI and SQS (IDI, E.C. 5.3.3.2/SQS, E.C. 2.5.1.21), which occurs in at least thirty-three diatom species, including *P. tricornutum* (Fabris et al., 2014) and *T. pseudonana* (Ferriols et al., 2017). Among those species, nineteen also express an additional independent IDI enzyme. Alternative splicing has also been suggested as a mechanism by which independent IDI proteins can be produced (Athanasakoglou et al., 2019). Similarly, in this study we found the IDI-SQS fusion transcript in the three diatom species evaluated, though a distinct independent IDI sequence was not detected in *C. muelleri* (Fig. 3.S9b).

The "central steps" of isoprenoid biosynthesis consist of the conversion of the precursors IPP and DMAPP into prenyl diphosphates such as FPP and GPP. We detected several candidates matching prenyltransferase activities: two candidates in *C. muelleri* for GPPS activity (E.C. 2.5.1.1) and one in both *T. pseudonana* and *P. tricornutum*. One match was found in the three diatoms evaluated for FPPS (E.C.2.5.1.10) (Table 3.1, Fig. 3.2d). It is
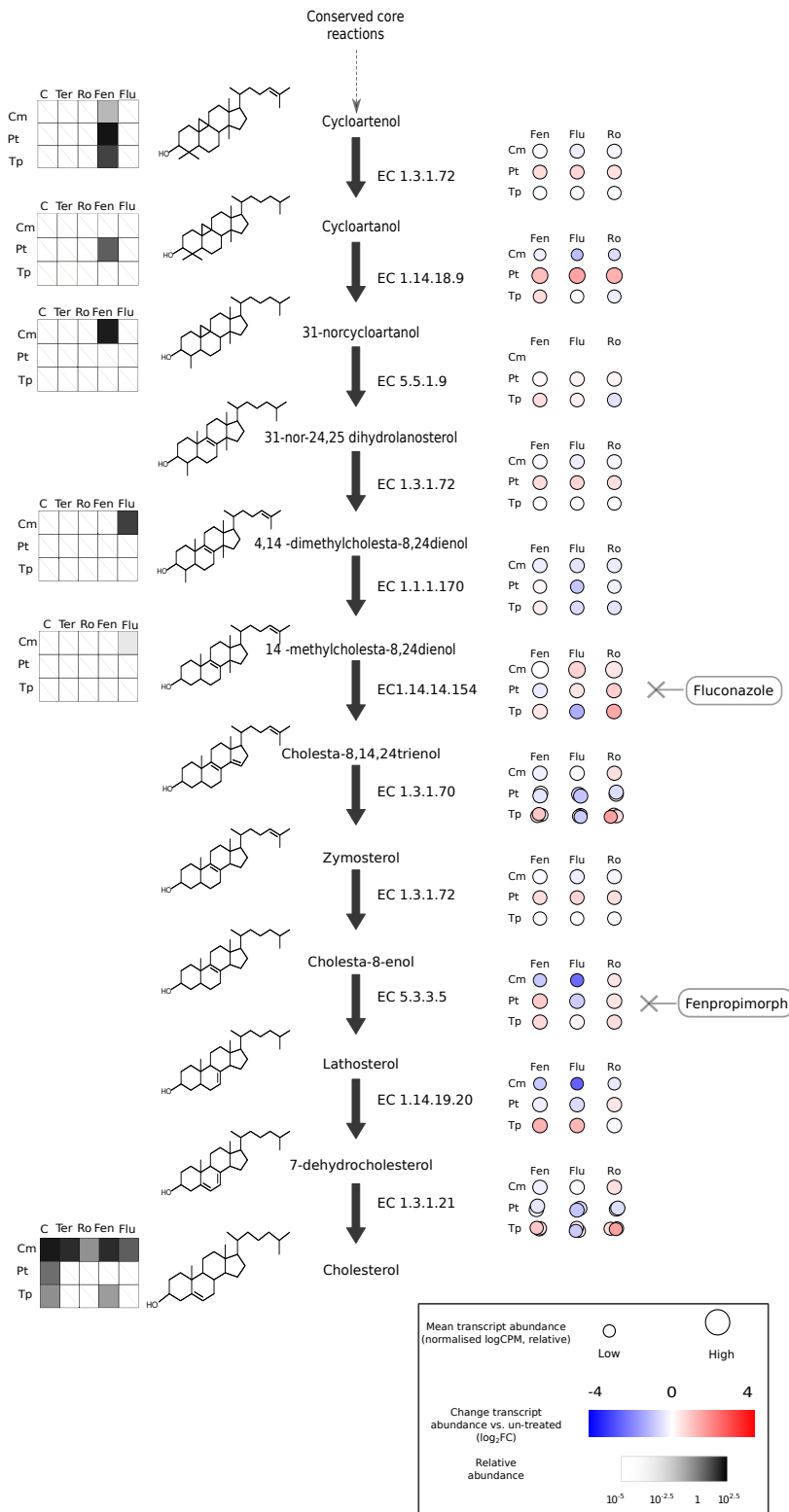
Figure 3.5: A hypothetical model of cholesterol synthesis in diatoms. Grey scale heatmaps indicate relative abundances of sterol and intermediate compounds identified by GC–MS, expressed as peak area normalized on a persample basis to biomass and an internal standard. Presented sterol level value is the average between the three replicates, n = 3. Relative transcript abundances (CPM) are shown as shaded circles, with log 2 fold-changes (Log 2 FC) relative to untreated control cultures shown in blue (decreased) and red (increased) (n = 16). The sizes of these circles are proportional to average logCPM over all samples. Multiple circles indicate multiple putative genes predicted to encode each respective enzyme function. Cm: *Chaetoceros muelleri*; Pt: *Phaeodactylum tricornutum*; Tp: *Thalassiosira pseudonana*. C: control cultures (solvent without inhibitor added); Fen: Fenpropimorph; Flu: Fluconazole, Ro: Ro 48-8071. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

not known which of these prenyltransferases are specifically participating in the synthesis of sterols. Functional characterization of prenyltransferases in the diatom H. ostrearia showed five different candidates, present in different subcellular locations and expressing distinct physiological roles (Athanasakoglou et al., 2019).

3.5 Conclusions

This study deepens our understanding of sterol biosynthesis across representatives of three distinct diatom groups,expanding upon prior findings in *P. tricornutum*, and revealing new metabolic features of two additional model diatoms, *T. pseudonana* and *C. muelleri*. The isoprenoid and sterol biosynthesis pathways of diatoms appear to be responsive to chemical perturbations at a metabolic and transcriptomic level. Changes in the relative abundances of transcripts encoding sterol metabolic functions suggest compensatory mechanisms by which diatoms maintain their metabolic pathways in response to chemical challenges. The genetic plasticity of metabolism in diatoms is evident in the diversity of sterols that different species produce, despite sharing a common core of enzymes and intermediates that is unique to diatoms. Differences in the supply of precursors via MEP pathway in *C. muelleri*, as well as differences in the presence and apparent regulation of enzyme homologs among species point to further aspects of divergence or specialization during the recent and continuing evolution of diatom metabolism.

Statement of informed consent, human/animal rights

No conflicts, informed consent, or human or animal rights are applicable to this study.

CRediT authorship contribution statement

Ana Cristina Jaramillo-Madrid: Conceptualization, Investigation, Funding acquisition, Methodology, Formal analysis, Writing - original draft. **Justin Ashworth**: Conceptualization, Formal analysis, Methodology, Investigation, Funding acquisition, Supervision, Writing - original draft, Writing - review & editing. **Michele Fabris**: Conceptualization, Validation, Writing - review & editing. **Peter J. Ralph**: Supervision, Project administration, Funding acquisition, Writing - review & editing.

Declaration of competing interest

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

Acknowledgements

Supplemental material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.algal.2020.101902.

Table 3.S1: Read count statistics for mRNA sequencing.

Species	Total 100bp pairs ($\times 10^6$)	No. Samples	Median ($\times 10^6$)	Min ($\times 10^6$)	Max ($\times 10^6$)
Chamu	864	16	54	41	66
Phatr	422	16	23	12	50
Thaps	892	16	55	45	86

Table 3.S2: Results of Benchmarks for Universal Single-Copy Orthologs (BUSCO) for de novo transcriptome assemblies.

Species	Complete	Single-Copy	Duplicated	Fragmented	Missing
Thaps	256 (84%)	96 (32%)	160 (53%)	18 (6%)	29 (10%)
Phatr	258 (85%)	10 (34%)	156 (51%)	16 (5%)	29 (10%)
Chamu	263 (87%)	15 (51%)	109 (36%)	12 (4%)	28 (9%)
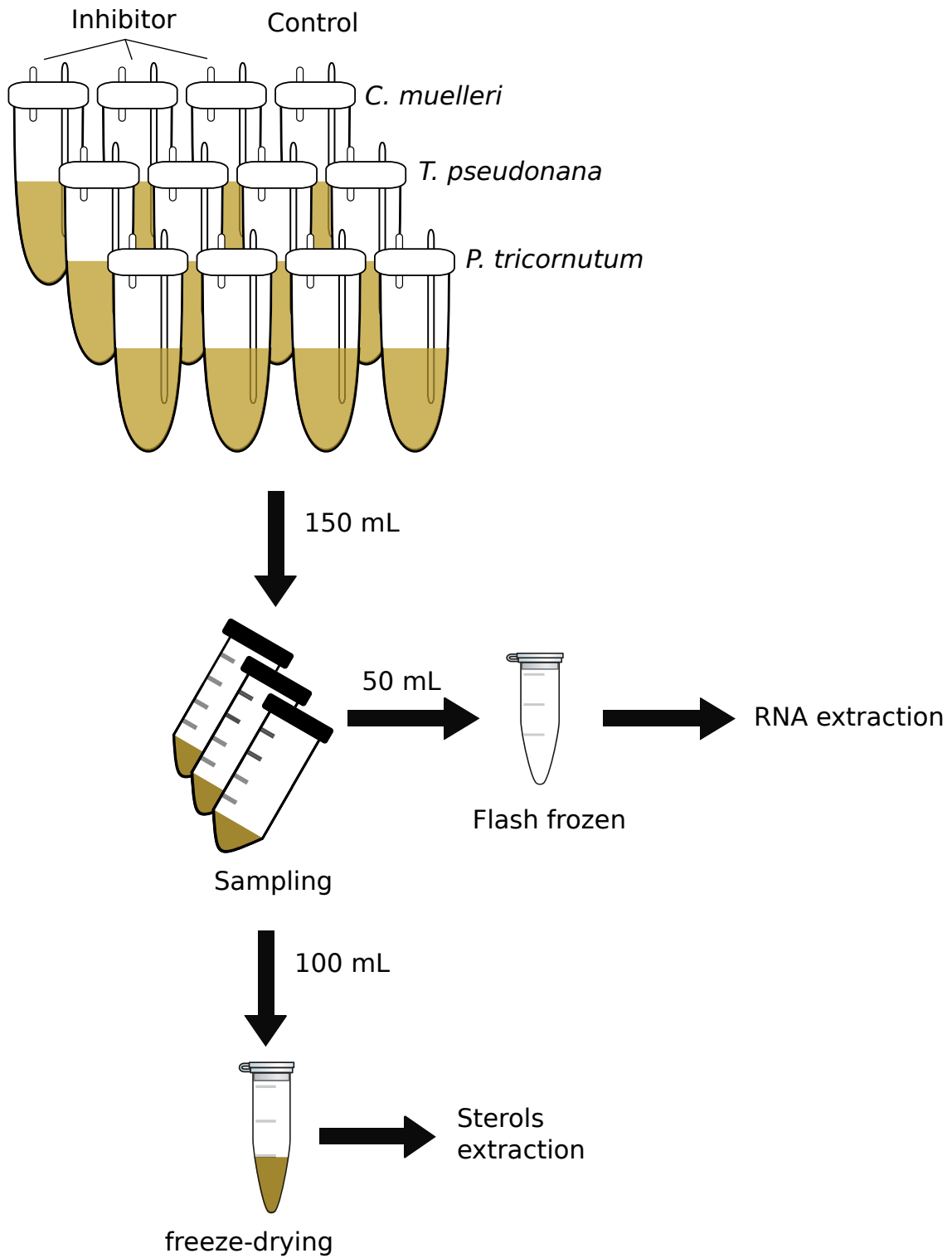
Figure 3.S1: Experimental set up in large scale photobioreactors for inhibitors experiment.
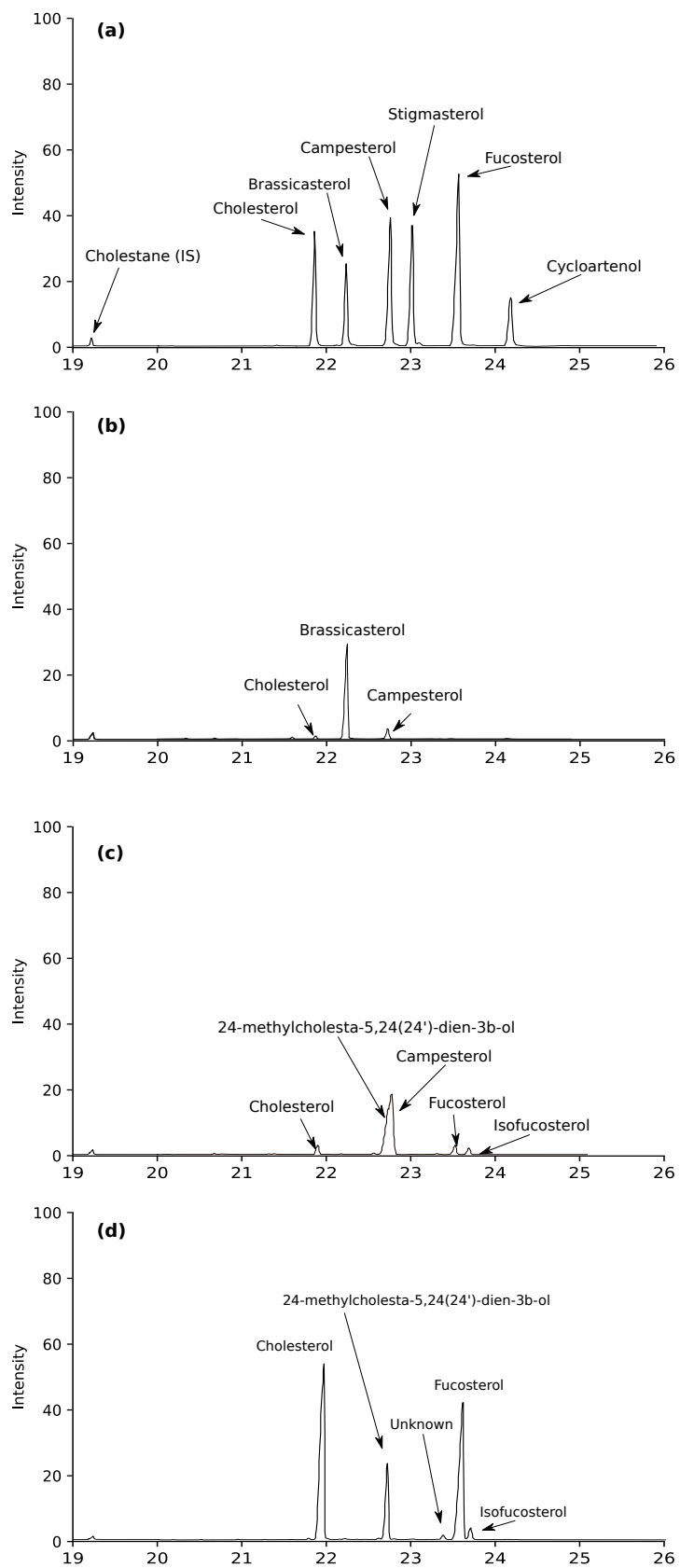
Figure 3.S2: Chromatograms from GC-MS analysis of each diatom specie growing without inhibitor addition. (a) Commercial standards (b) *P. tricornutum* (c) *T. pseudonana* (d) *C. muelleri*.
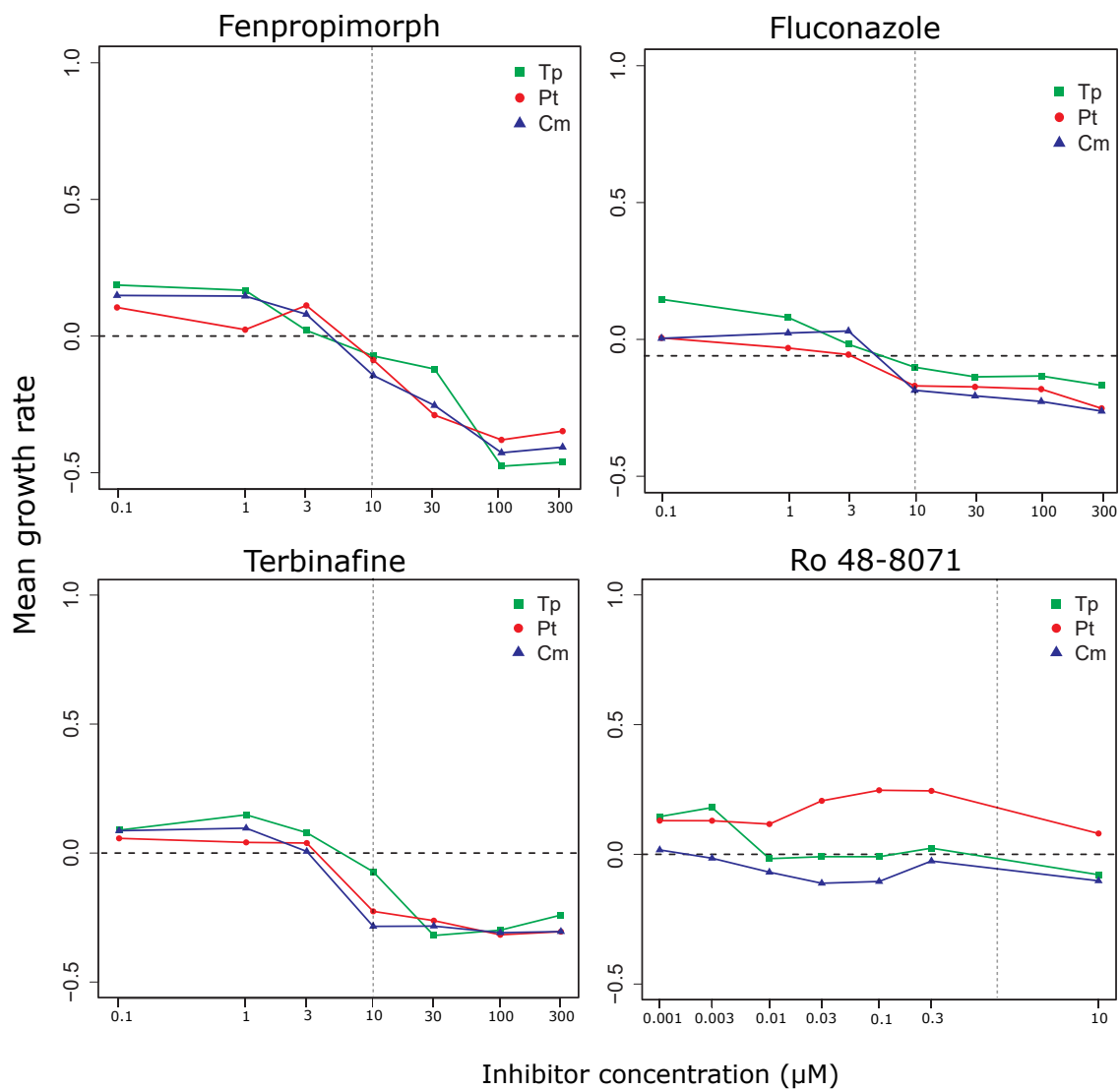
Figure 3.S3: Microscale (200 μL) inhibitor dose-effect screen in 96-well plates.
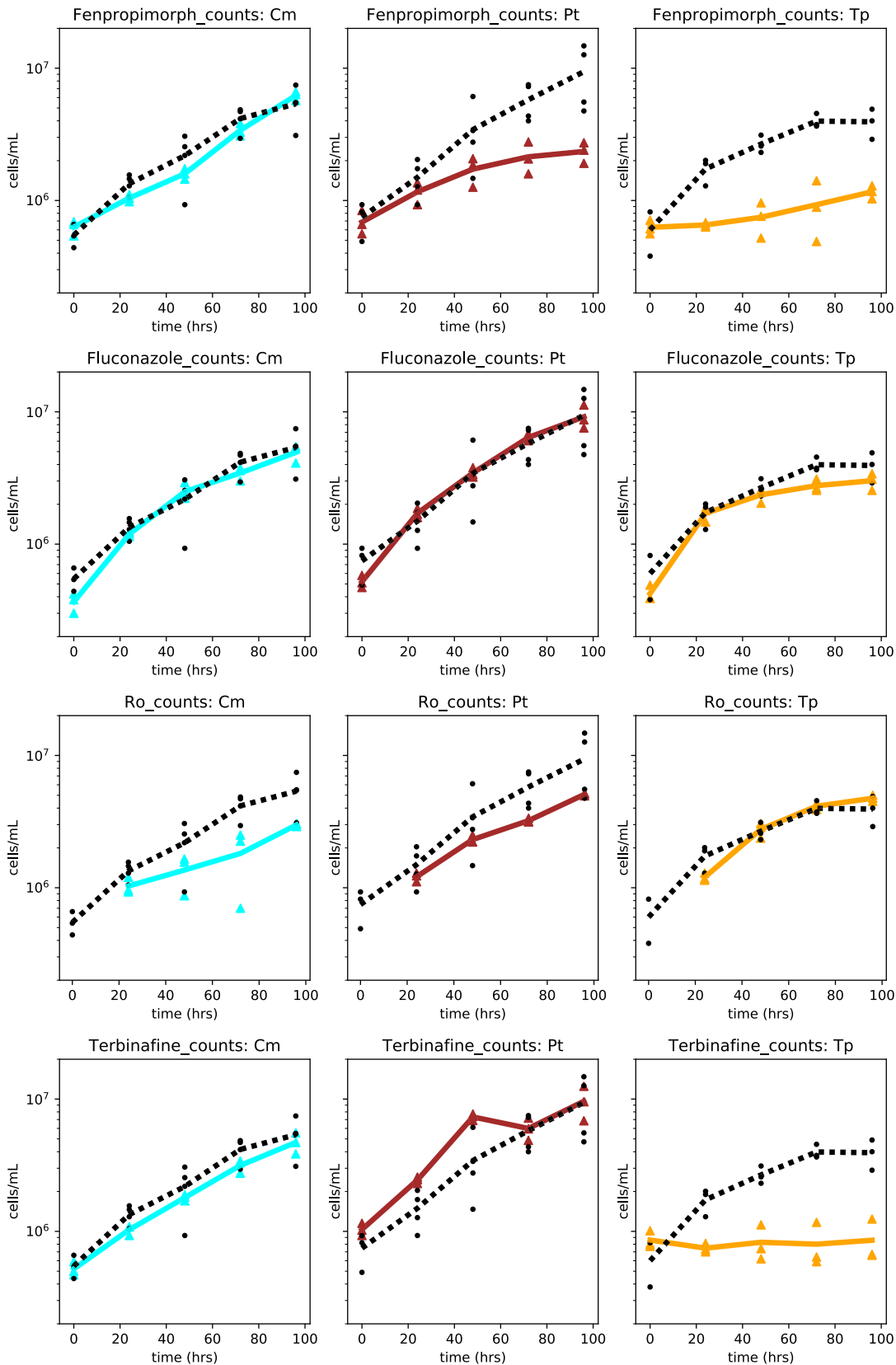
Figure 3.S4: Cell counts during replicate 1.2L aerated batch culture treatment experiments (n=3).
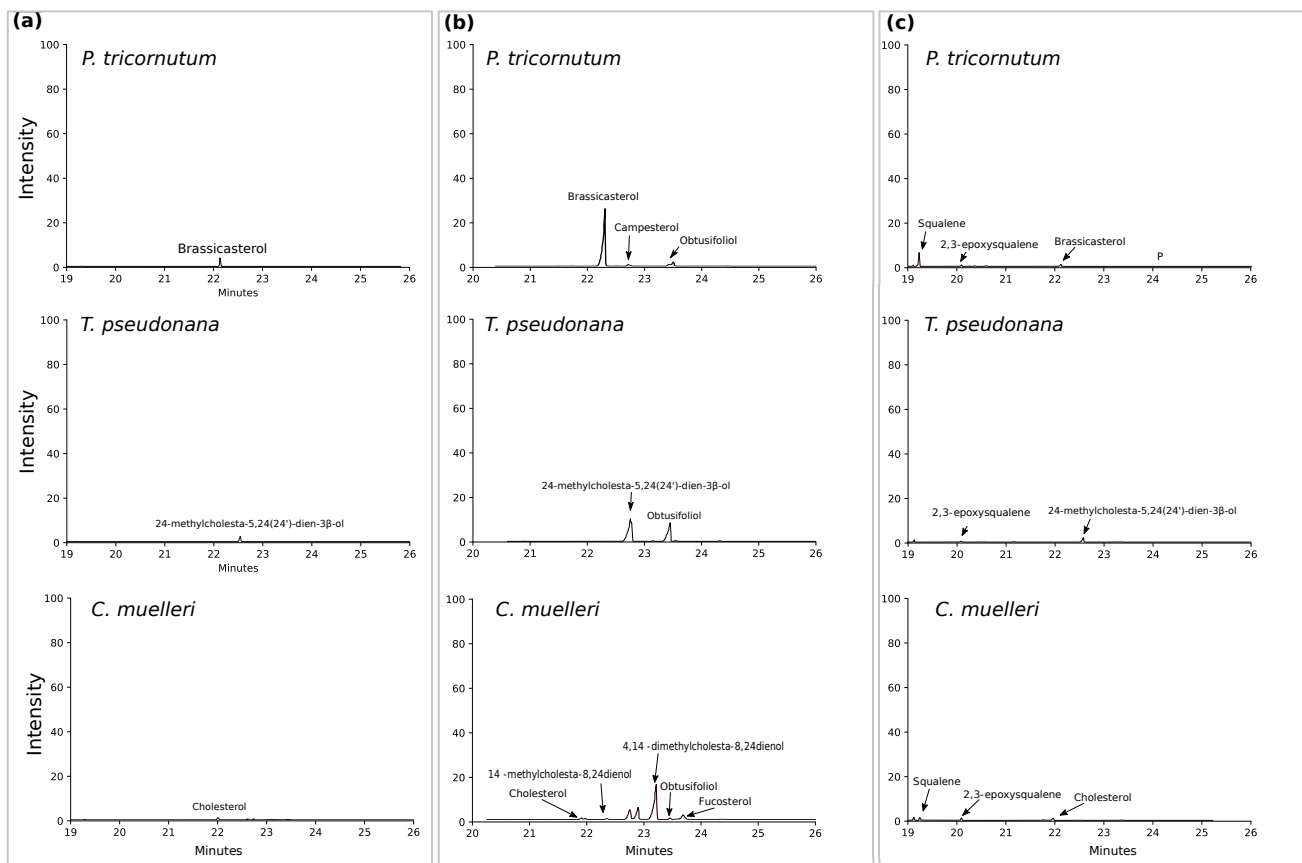
Figure 3.S5: Chromatograms from GC-MS analysis of each diatom specie under inhibitors treatment. (a) Terbinafine (b) Fluconazole (c) Ro 48-8071

Figure 3.S6: Chromatograms from GC-MS analysis of each diatom species under fenpropimorph treatment (a) Cycloartenol commercial standard, (b) Lanosterol commercial standard (c) *P. tricornutum* (d) *T. pseudonana* (e) *C. muelleri.*

Figure 3.S7: Mass spectra of putatively identified sterols.

DE Orthogenes (FDR<=0.001)
Enriched GO Terms (*p*<=0.05)

protein transport
metabolism
intracellular protein transport
protein folding
chromosome organization and biogenesis
small GTPase mediated signal transduction
ubiquitin-dependent protein catabolism
protein catabolism
vesicle-mediated transport
nucleosome assembly
Mo-molybdopterin cofactor biosynthesis
protein biosynthesis
regulation of lyase activity
proline catabolism
polyamine biosynthesis
tricarboxylic acid cycle
glucose metabolism
ATP synthesis coupled proton transport
regulation of transcription, DNA-dependent

nitrogen compound metabolism
carotenoid biosynthesis
pyrimidine base biosynthesis
electron transport
arginine biosynthesis
tricarboxylic acid cycle
mevalonate-independent IPP biosynthesis
beta-1,3 glucan biosynthesis
tRNA 5'-leader removal
nuclear division
leucyl-tRNA aminoacylation
chromate transport
porphyrin biosynthesis
fatty acid desaturation
glycolysis
biotin biosynthesis

glycerol-3-phosphate metabolism
electron transport
cation transport
polyamine biosynthesis
glyoxylate cycle
glycerol-3-phosphate catabolism
glycine catabolism

photosynthesis light harvesting
nitrogen compound metabolism
ATP synthesis coupled proton transport
tricarboxylic acid cycle
malate metabolism
pyrimidine base biosynthesis
ATP biosynthesis
glycolysis
proteolysis and peptidolysis
arginine biosynthesis
energy coupled proton transport
mevalonate-independent IPP biosynthesis
glutamine biosynthesis
main pathways of carbohydrate metabolism
beta-1,3 glucan biosynthesis
regulation of transcription, DNA-dependent
pentose-phosphate shunt, non-oxidative branch
tricarboxylic acid cycle intermediate metabolism
biotin biosynthesis

ubiquitin cycle
protein ubiquitination
rRNA processing
trehalose biosynthesis
protein modification
protein catabolism
vesicle-mediated transport
protein amino acid dephosphorylation
PEP-dependent sugar phosphotransferase system
protein biosynthesis
regulation of transcription, DNA-dependent
'de novo' IMP biosynthesis
electron transport
fatty acid biosynthesis
intracellular signaling cascade
aromatic compound metabolism

regulation of transcription, DNA-dependent
glycolysis
nitrogen compound metabolism
porphyrin biosynthesis
fatty acid desaturation
arginine biosynthesis
protein amino acid phosphorylation
amino acid metabolism
protein ubiquitination
antibiotic biosynthesis
[U/G/C]TP biosynthesis
isoprenoid biosynthesis
protein insertion into membrane
carotenoid biosynthesis
chlorophyll biosynthesis
cell redox homeostasis
spermatogenesis
pantothenate biosynthesis
asparagine biosynthesis
proteolysis and peptidolysis
cellular protein metabolism
RNA processing
tRNA aminoacylation for protein translation
protein biosynthesis
electron transport
signal transduction
photosynthesis light harvesting
aromatic amino acid family biosynthesis
carbohydrate metabolism
mitosis
lipid biosynthesis
DNA replication, synthesis of RNA primer
pyrimidine base biosynthesis

Flu

Fen

Ro

110   71   866
8   14
40   37

Figure 3.S8: GO term enrichments among differentially expressed orthologous transcripts from mRNA-seq transcriptomes.

Figure 3.S9: Sequence alignment for the enzymes IDI, SQS and OSC. (a) Multiple sequence alignment for gene and transcript models of independent IDI (IDI, E.C. 5.3.3.2). (b) Multiple sequence alignment for gene and transcript models of IDI-SQS fusion (IDI, E.C. 5.3.3.2/SQS, E.C. 2.5.1.21) and other non-diatom species with independent SQS gene. (c) Alignment of oxidosqualene cyclase (OSC, E.C. 5.4.99.8) homologs from representative taxa. All diatom species are inside the square. D455 is the catalytic residue. Positions 381, 449, and 453 are differentially conserved between lanosterol synthases (V453) and cycloartenol synthases (I453).

Chapter 4

Over-expression of key sterol pathway enzymes alters intermediates and end-point sterol profiles in the model marine diatom *Phaeodactylum tricornutum*

Abstract

Sterols are a class of triterpenoid molecules with diverse functional roles in eukaryotic cells, including intracellular signaling and regulation of cell membrane fluidity. Diatoms are a dominant eukaryotic phytoplankton group that produce a wide diversity of sterol compounds. The enzymes 3-hydroxy-3-methyl glutaryl CoA reductase (HMGR) and squalene epoxidase (SQE) have been reported to be rate-limiting steps in sterol biosynthesis in other model eukaryotes; however, the extent to which these enzymes regulate triterpenoid production in diatoms is not known. To probe the role of these two metabolic nodes in the regulation of sterol metabolic flux in diatoms, we independently over-expressed two versions of the native HMGR and a conventional, heterologous SQE gene in the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*. Over-expression of these key enzymes resulted in significant differential accumulation of downstream sterol pathway intermediates in *P. tricornutum*. HMGR over-expression resulted in the accumulation of squalene, cycloartenol, and obtusifoliol, while cycloartenol and obtusifoliol accumulated in response to heterologous SQE over-expression. In addition, accumulation of the end-point sterol 24-methylenecholesta-5,24(24')-dien-3β-ol was observed in all *P. tricornutum* over-expression lines, and campesterol increased 3-fold in *P. tricornutum* lines expressing SQE. Minor differences in end-point sterol composition were also found in *T. pseudonana*, but no accumulation of sterol pathway intermediates was observed. Despite the successful manipulation of pathway intermediates and individual sterols in *P. tricornutum*, total sterol levels did not change significantly in transformed lines, suggesting the existence of tight pathway regulation to maintain total sterol content.

4.1 Introduction

Sterols are essential triterpenoids that function as regulators of cell membrane dynamics in all eukaryotic organisms (Dufourc, 2008).In animals and higher plants, sterols participate in the synthesis of secondary metabolites involved in defense mechanisms, and steroid hormones that regulate growth and development (Valitova et al., 2016). Due to their presence in ancient sediments, sterol compounds are used as durable biomarkers to track important evolutionary events (Gold et al., 2017). Sterols of plant origin, known as phytosterols, are used as nutraceuticals for their cholesterol-lowering effects (Ras et al., 2014). Other therapeutic applications such as anti-inflammatory (Aldini et al., 2014) and anti-diabetic activities (Wang et al., 2017) are currently under research. In order to meet increasing demands in the global phytosterols market, about 7–9% per annum (Borowitzka, 2013), diatoms have been proposed as an alternative source of sterols (Jaramillo-Madrid et al., 2019).

Diatoms are primary constituents of phytoplankton communities and principal players in the global carbon cycle. These photosynthetic microorganisms are an important ecolog-

ical group of microalgae present in a great diversity of aquatic environments (Armbrust, 2009). Diatoms exhibit higher photosynthetic efficiencies than plants and are adaptable to environmental challenges encountered in dynamic and competitive marine environments (Hildebrand et al., 2012), which are also characteristics suited to the microbial production of bioproducts. Diatoms are emerging as alternative and sustainable hosts for terpenoids production (D'Adamo et al., 2019; Fabris et al., 2020). As complex organisms with a particular evolution history, diatoms possess a unique metabolism (Fabris et al., 2014, 2012; Jaramillo-Madrid et al., 2020a; Pollier et al., 2019) that can represent an advantage for production of terpenoid such as sterols (Vavitsas et al., 2018).

Diatoms produce high proportions of a large variety of sterol compounds (Rampen et al., 2010). Sterol sulfates appear to be important regulators of diatom bloom dynamics, as they were shown to trigger programmed cell death in the marine diatom *Skeletonema marinoi* (Gallo et al., 2017). Recent studies suggest that sterol biosynthesis is tightly regulated. Levels of intermediate compounds in sterol synthesis change in response to different environmental conditions (Jaramillo-Madrid et al., 2020b) and to the addition of chemical inhibitors (Jaramillo-Madrid et al., 2020a). However, end-point sterol levels remain unchanged under same treatments. Deeper understanding of diatom sterol metabolism will provide ecological insights as well as enable future metabolic engineering efforts for biotechnological applications. In particular, the regulation of the sterol biosynthesis in diatoms is not yet well understood.

Isoprenoid sterol precursors can be synthesized through either the cytosolic mevalonate (MVA) pathway or the plastidial methylerithriol phosphate (MEP) pathway. In most eukaryotic organisms, only one of the two pathways are present (Lohr et al., 2012). However, in plants, both pathways are functional but the MVA provides the substrates for sterol biosynthesis (Vranová et al., 2013). In diatoms, including the model diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, both pathways are functional (Jaramillo-Madrid et al., 2020a). However, there is no evidence for MVA presence in some diatoms such as *Haslea ostrearia* and *Chaetoceros muelleri* (Athanasakoglou et al., 2019; Jaramillo-Madrid et al., 2020a; Massé et al., 2004). In these diatoms, synthesis of isoprenoids may rely solely on the MEP pathway, as is the case for some green and red algae (Lohr et al., 2012). The presence of both the MVA and MEP pathways is an advantage for engineering efforts, as it potentially provides a higher pool of intermediates for isoprenoid production (Jaramillo-Madrid et al., 2020a; Sasso et al., 2012). It has been recently demonstrated that in *P. tricornutum* products from MVA pathway accumulated in the cytoplasm can be used for the production of non-endogenous terpenoids such as geraniol, indicating presence of free GPP pool (Fabris et al., 2020).

In the MVA pathway, three molecules of acetyl-CoA are transformed into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (Fig. 4.1). In plants, fungi, and animals, the 3-hydroxy-3-methylglutaryl-co-enzyme-A reductase (HMGR, E.C. 1.1.1.34)

is one of the key enzymes in the MVA pathway and catalyzes the reduction of HMG-CoA to mevalonate (Friesen & Rodwell, 2004). HMGR is known as main regulator and rate-limiting enzyme in early biosynthesis of sterol and non-sterol isoprenoids in MVA-harboring eukaryotic cells and it is highly regulated at the transcriptional, translational, and post-translational levels (Burg & Espenshade, 2011). In yeast and mammals, HMGR contains a sterol sensing domain (SSD) that is responsible for detecting sterol levels in the cell and maintaining sterol homeostasis (Espenshade & Hughes, 2007). The SSD is located in the N-terminal membrane binding domain of the HMGR enzyme (Burg & Espenshade, 2011). Moreover, it has been reported that genetic manipulations on HMGR, including truncation of N-terminal domain, led to considerable accumulation of terpenes in trans-genic plants and yeast (Bansal et al., 2018; Bröker et al., 2018; Lee et al., 2019). Although HMGR has been extensively characterized in model eukaryotic organisms, little is known about its features in diatoms.

MVA products IPP and DMAPP are subsequently used for the synthesis of squalene, the first committed intermediate in the formation of sterols (Gill et al., 2011) (Fig. 4.1). In plants, fungi, and animals, squalene is converted into 2,3 epoxysqualene. This reaction is conventionally catalyzed by the enzyme squalene epoxidase (SQE, E.C. 1.14.14.17) (Gill et al., 2011). Several studies indicate that SQE is a control point in cholesterol synthesis modulated by sterol levels and post-translationally regulated by cholesterol-dependent proteasomal degradation (Gill et al., 2011; Nagai et al., 2002b). However, diatoms do not possess a conventional SQE, and instead this step is catalyzed by a recently characterized alternative squalene epoxidase (AltSQE) (Pollier et al., 2019). The diatom AltSQE belongs to the fatty acid hydroxylase superfamily, and differs from the conventional flavoprotein SQE. Whether AltSQE has a similar role to SQE in sterol regulation is not known.

While selective inhibitors for AltSQE are not known, the treatment of diatoms with statins, known to inhibit HMGR enzymes, has resulted in perturbation of isoprenoids metabolism that included an overall decrease of total sterols content (Conte et al., 2018; Massé et al., 2004), suggesting that a conventional HMGR enzyme might be involved in the pathway. In this work, we genetically targeted the HMG CoA reduction and squalene epoxidation steps in *P. tricornutum* to provide further insights into the nature of the diatom sterol biosynthesis pathway and its regulatory constraints. Considering the demonstrated challenges in genetically down-regulating the essential genes involved in *P. tricornutum* sterol metabolism (Pollier et al., 2019, Fabris et al., 2014), we chose to investigate these pathway nodes by gene over-expression and sub-cellular localisation.

We generated independent diatom exconjugant lines constitutively expressing (i) either the full-length HMGR, (ii) an N-terminal truncated version of HMGR (tHMGR), or (iii) a heterologous SQE from Nannochloropsis oceanica (NoSQE) over the background of the endogenous AltSQE. By phenotyping these transgenic diatom cell lines, we describe specific changes in several nodes of the sterol biosynthesis pathway and provide evidence for
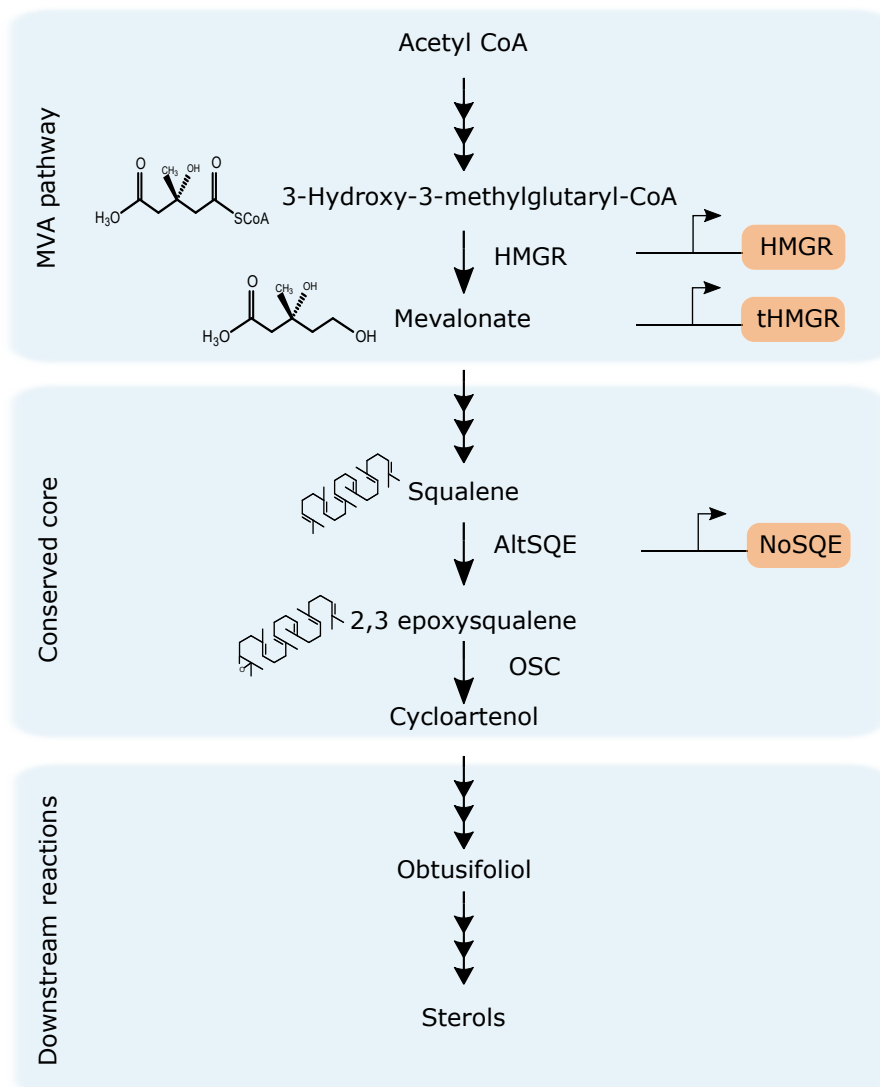
Figure 4.1: Upstream reactions and conserved core of sterol biosynthesis pathway in diatoms and genetic targets overexpressed in this study, highlighted in orange. Mevalonate pathway, MVA; 3-hydroxy-3-methylglutaryl-coenzyme A reductase, HMGR, truncated HMGR, tHMGR, alternative squalene epoxidase, AltSQE, squalene epoxidase from *N. oceanica*, NoSQE, oxidosqualene cyclase, OSC.

regulatory mechanisms unique to diatom sterol metabolism.

4.2 Methodology

4.2.1 Diatom culturing

The species *P. tricornutum* (CCMP632) and *T. pseudonana* (CCMP1335) were obtained from the National Centre for Marine Algae and Microbiota at Bigelow Laboratory (USA). Axenic cultures were maintained in L1 medium (Guillard & Hargraves, 1993) at 18°C under continuous cool white light (100 μmol photons m^{-2} s^{-1}) in a shaking incubator (100 rpm).

4.2.2 Episome construction and transformation

All episomes used in this study were assembled using uLoop assembly method (Pollak et al., 2019). Individual components for episome assembly (L0 parts) were built and domesticated using uLoop assembly syntax. Assembly reactions were performed using the respective uLoop assembly backbones for each level as described by Pollak et al. (2019). After domestication, each L0 part was confirmed by Sanger sequencing. Correct episome assemblies were confirmed by colony PCR and diagnostic restriction digestion. The source of each DNA part and primers used for domestication are listed in Table 4.S1. All L0 parts used to assemble the plasmids used in this work have been deposit in Addgene (Table 4.S 1). Plasmid maps (Figs. 4.S1, 4.S2) and complete plasmid sequence are provided in supplemental material.

Episomes consisted of a pCA-derived backbone (Pollak et al., 2019), CEN/ARS/HIS and OriT sequences, a selection cassette, and an expression cassette (Figs. 4.S1, 4.S2). Sequence OriT required for bacterial conjugation were amplified from pPtPBR11 plasmid (Diner et al., 2016) (Genebank KX523203). Selection markers nourseothricin (NAT) for *T. pseudonana* and blasticidin-S deaminase (BSD) for *P. tricornutum* were driven by elongation factor 2 (EF2) constitutive promoters from corresponding diatom species (*T. pseudonana* v. 3 ID: 269148; *P. tricornutum* v. 3 ID: Phatr3_J49202). Expression cassettes included genes encoding either putative HMGR, tHMGR, or NoSQE each fused at the C-terminus with a mVenus fluorescent protein (Nagai et al., 2002a), and an expression cassette expressing only mVenus was used to assemble an empty control vector (Fig. 4.S1). The open reading frames encoding putative HMGR and tHMGR were amplified from genomic DNA of either *T. pseudonana* (CCMP1335) (Gene ID Thaps_33680) or *P. tricornutum* (CCMP632) (Gene ID Phatr3_J16870) (Table 4.S1). A domesticated, codon-optimized synthetic gene encoding SQE sequence from Nannochloropsis oceanica v.2 CCMP1779 (Gene ID 521007) was obtained from Genewiz® (USA) (complete sequence in Supplemental Material fasta file 1). Expression of target genes were driven by the promoter of elongation factor 2 (EF2) in *T.*

pseudonana and the promoter of predicted protein Phatr3_J49202 in *P. tricornutum*. L0 parts for CEN/ARS/HIS, fluorescent reporter gene mVenus, Phatr3_J49202 promoter and terminator were obtained from Dr. Christopher Dupont (J. Craig Venter Institute, USA). The plasmid pTA-Mob for conjugation (Strand et al., 2014) was obtained from Dr. Ian Monk (University of Melbourne, Australia).

4.2.3 Diatom transformation and screening

Diatoms were transformed by bacterial conjugation (Karas et al., 2015). The transformation protocol for *T. pseudonana* was modified by increasing the starting bacterial density (OD600 to 0.5) and the final incubation and recovery period for transformed diatom culture to 24 hrs prior to selection on plates containing nourseothricin. *T. pseudonana* and *P. tricornutum* colonies resistant to nourseothricin (50 μg ml^{-1}) or blasticidin (10 μg ml^{-1}), respectively, were inoculated in 96-multiwell plates containing 200 μl of L1 medium with 100 μg ml^{-1} of nourseothricin or 10 μg ml^{-1} blasticidin, depending on the diatom species, and subcultured every 5 days. Clonal lines from 96-well plates were screened by detecting mVenus fluorescence using a CytoFLEX S (Beckman Coulter) flow cytometer operated in plate mode. 48 clones of each expression system were screened for *T. pseudonana* and 12 for *P. tricornutum*. A 488 nm laser was used for fluorescence excitation; mVenus fluorescence was detected using a 525/40 nm filter and chlorophyll fluorescence was detected using 690/50 nm filter. 10,000 events were analyzed per sample. Three independent cell lines per construct with the highest median mVenus fluorescence readings were selected for full-scale experiments, including WT and empty vector controls.

4.2.4 Experiments with transgenic diatom cultures

Three replicates of each selected clone were inoculated in 5 ml of L1 medium (100 μg ml^{-1} nourseothricin or 10 μg ml^{-1} blasticidin) and grown for 3 days. Subsequently, cultures were upscaled to 50 ml L1 supplemented with the respective antibiotic for 5 days and these were used to inoculate cultures in L1 medium for sterol analysis experiments. Full-scale experiments were carried out in 200 ml flasks containing 120 ml of L1 medium and antibiotic under continuous light (150 μmol photons m^{-2} s^{-1}) and constant shaking (95 rpm). Cell density and mVenus fluorescence were monitored daily by sampling 200 μl from each culture and transferring it to a 96 well plate for high-throughput flow cytometry analysis. Pulse Amplitude Modulated (PAM) fluorometry was used to estimate photosynthetic activity by comparing fluorescence yield of PSII under ambient irradiance (F) and after application of a saturating pulse (Fm) (Schreiber, 2004). After 48 hours of growth, biomass was harvested by centrifuging at 4000 g for 10 minutes. Diatom pellets were washed with Milli-Q water to eliminate excess salt, freeze-dried to determine dry weight, and kept at

-20°C until sterol extraction.

4.2.5 Extraction and analysis of sterols by GC-MS

For sterol extraction, dry cell matter was heated in 1 ml of 10% KOH ethanolic solution at 90°C for one hour. Sterols were extracted from cooled material in three volumes of 400 μL of hexane. An internal standard, 5α-cholestane, was added to each sample. Hexane fractions were dried under a gentle N2 stream, and derivatized with 50 μL of 99% BSTFA + 1% TMCS (N,O-Bis(trimethylsilyl)trifluoroacetamide, Trimethylsilyl chloride) at 70°C for one hour. The resulting extractions were resuspended in 50 μL of fresh hexane prior to GC-MS injection.

Gas chromatography/mass spectrometry (GC-MS) analysis was performed using an Agilent 7890 instrument equipped with a HP-5 capillary column (30 m; 0.25 mm inner diameter, film thickness 0.25 μm) coupled to an Agilent quadrupole MS (5975 N) instrument. The following settings were used: oven temperature initially set to 50°C, with a gradient from 50°C to 250°C (15.0°C min^{-1}), and then from 250°C to 310°C (8°C min^{-1}, hold 10 min); injector temperature = 250°C; carrier gas helium flow = 0.9 ml min^{-1}. A split-less mode of injection was used, with a purge time of 1 min and an injection volume of 5 μL. Mass spectrometer operating conditions were as follows: ion source temperature 230°C; quadrupole temperature 150°C; accelerating voltage 200 eV higher than the manual tune; and ionization voltage 70 eV. Full scanning mode with a range from 50 to 650 Dalton was used.

Sterol peaks were identified based on retention time, mass spectrum, and representative fragment ions compared to the retention times and mass spectrum of authentic standards. The NIST (National Institute of Standards and Technology) library was also used as reference. The area of the peaks and deconvolution analysis was carried out using the default settings of the Automated Mass Spectral Deconvolution and Identification System AMDIS software (v2.6, NIST). Peak area measurements were normalized by both the weight of dry matter prior to extraction, and the within-sample peak area of the internal standard 5a-cholestane. Sterol standards used to calibrate and identify GC-MS results in this study included: cholest-5-en-3-β-ol (cholesterol); (22E)-stigmasta-5,22-dien-3β-ol (stigmasterol); stigmast-5-en-3-β-ol (sitosterol); campest-5-en-3-β-ol (campesterol); (22E)-ergosta-5,22-dien-3-β-ol (brassicasterol); (24E)-stigmasta-5,24-dien-3β-ol (fucosterol); 9,19-Cyclo-24-lanosten-3β-ol (cycloartenol), 5-α-cholestane; and the derivatization reagent bis(trimethyl-silyl) trifluoroacetamide and trimethylchlorosilane (99% BSTFA + 1% TMCS) were obtained from Sigma-Aldrich, Australia.

4.2.6 Fluorescence imaging

Live diatom transformants expressing mVenus were imaged without fixative with a confocal laser scanning microscope (Nikon A1 Plus, Japan) and photomultiplier tube (PMT) detector. The 488-nm and 637-nm lasers were used for mVenus and chlorophyll autofluorescence, respectively. Gains on the detector were kept constant between samples and controls. Images were acquired with $60\times/1.4$ objective oil immersion objective and processed using imaging software NIS-Elements Viewer 4.0 (Nikon, Japan).

4.2.7 Multiple sequence alignment and phylogenetic reconstruction

Diatom homologue sequences were retrieved either from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) (Johnson et al., 2019; Keeling et al., 2014) database or from GenBank protein database (Table 4.S2) using BLASTp search with *T. pseudonana* HMGR as the query sequence. The species names and corresponding MMETSP ID numbers are listed in Table 4.S2. HMGR from yeast, mammals, and plants were used as outgroups. Sequences from outgroups (Table 4.S2) were obtained from GenBank protein database. Multiple sequence alignments of the full-length protein sequences were performed by MAFFT version 7 program with default parameters and alignments were manually edited by exclusion of ambiguously aligned regions. The maximum likelihood tree phylogenetic tree was constructed using MEGA 6 with partial deletion option. The reliability of obtained phylogenetic tree was tested using bootstrapping with 1000 replicates. Prediction of transmembrane helices in HMGR from diatoms was carried out using the TMHMM Server v. 2.0 with default parameters (Krogh et al., 2001). Conserved motifs in the selected sequences were identified by an InterProScan (Jones et al., 2014) search against all available member databases, including Pfam (protein families) and SUPERFAMILY (structural domains).

4.2.8 Statistical analysis

All plots were generated using R: A language and environment for statistical computing. All experiments were conducted in triplicate. The analyses performed were Shapiro-Wilk to test normality, non-parametric Kruskal–Wallis test and Pairwise Wilcoxon Rank Sum Tests to calculate pairwise comparisons between group levels with corrections for multiple testing. Differences between groups were considered significant at $p < 0.05$.

4.3 Results

4.3.1 Identification of putative HMGR from *T. pseudonana* and *P. tricornutum*

In contrast to the recently discovered AltSQE, little is known about the diatom HMGR enzyme despite being a major regulatory step in sterol biosynthesis. Previous studies demonstrated that specific HMGR inhibitors alter isoprenoids metabolism in the diatoms *P. tricornutum*, *Haslea ostrearia* and *Rhizosolenia setigera* (Conte et al., 2018; Massé et al., 2004). Given the presence of many unusual features in diatoms metabolism (Fabris et al., 2012; Jaramillo-Madrid et al., 2020a; Pollier et al., 2019) we analyzed the conservation of the HMGR sequence among all the diatom species with genomic or transcriptomics sequences available.

The amino acid sequence of *Arabidopsis thaliana* HMGR enzyme (AT1G76490) was used as query to search against the genome sequence of the diatoms *T. pseudonana* and *P. tricornutum* to identify the genes putatively encoding HMGR. Through this analysis, we identified a single copy of a putative HMGR gene located in chromosome 29 in *P. tricornutum* (Gene ID Phatr3_J16870)(Fabris et al., 2014) and chromosome 4 in *T. pseudonana* (Gene ID Thaps_33680). In model eukaryotic organisms, HMGR is characterized by the presence of an N-terminal membrane domain and a C-terminal catalytic region. Sequence alignment analysis revealed differences in membrane domain location among model organisms, while the catalytic region is conserved (Figs. 4.2, 4.3). The C-terminal catalytic domain of HMGR was highly conserved across all the organisms analyzed (Fig. 4.3). The catalytic residues Glu559, Asp767, and His866 (Friesen & Rodwell, 2004) (Li et al., 2014), were also found to be present and conserved in *T. pseudonana* and *P. tricornutum* (Fig. 4.2). HMGR from *Saccharomyces cerevisiae* and *Homo sapiens* possess a sterol sensing domain (SSD) in the transmembrane N-terminal region, which is involved in sterol homeostasis (Burg & Espenshade, 2011), therefore, we analyzed HMGR transmembrane region in diatoms to identify similarities with other model organisms. Most of the analyzed HMGR sequences from diatoms possess three trans-membrane helices in the N-terminal domain, except a few with two or none domains predicted (Table 4.S2). In comparison, plants usually have two domains (Li et al., 2014). We found seven transmembrane domains in *S. cerevisiae* and five in *Homo sapiens* (Table 4.S22). We did not detect similarities with known yeast and mammals SSDs in the N-terminal region in any of the diatoms analyzed (Fig. 4.3).

Figure 4.2: Sequence alignment of HMGR protein (catalytic domain) from model organisms. Asterisks (*) indicate conserved catalytic residues.

4.3.2 Phylogenetic analysis of HMGR and conserved protein domains

Based on the alignments of full-length HMGR protein sequences of twenty-eight diatom species retrieved from whole genome and transcriptome assemblies, a neighbor-joining phylogenetic tree was constructed to study evolutionary relationship of HMGR protein sequence among diatoms (Fig. 4.3). We designated HMGR from yeast, mammals, and plants as outgroups. Pennate and centric diatoms were divided into two different clades (Fig. 4.3). As expected, HMGR from diatoms of the same genus tended to cluster together. Species from the order Thalassiosirales which includes *Thalassiosira* and *Skeletonema* genus are grouped together (Fig. 4.3). Similarly, HMGR from the diatoms *P. tricornutum* and *Fistulifera solaris* that belong to the *Naviculales* order appear to be closely related (Fig. 4.3). Interestingly, we did not find a match for HMGR in the transcriptomic sequences of the diatoms: *Chaetoceros muelleri*, as previously reported (Jaramillo-Madrid et al., 2020a), *Chaetoceros brevis*, *Chaetoceros debilis* and *Chaetoceros curvisetus* (Table 4.S2).

4.3.3 Expression and localization of putative HMGR and tHMGR

While the core reactions in sterol synthesis being conserved in *T. pseudonana* and *P. tricornutum* (Jaramillo-Madrid et al., 2020a), both diatoms produce a distinctive profile of sterol compounds which variates differently upon changing environmental conditions and chemical inhibitors treatment (Jaramillo-Madrid et al., 2020a,b). Additionally, the sterol metabolism of the centric diatom *T. pseudonana* has not been explored to the same depth as the model pennate *P. tricornutum*. To evaluate the effect of over-expression of the rate-limiting enzyme HMGR on sterol accumulation, *T. pseudonana* and *P. tricornutum* were transformed with episomes containing their respective putative HMGR copy driven by a constitutive promoter (Fig. 4.1, Figs. 4.S1, 4.S2). Episomes are maintained extra-chromosomally and therefore enable more consistent expression required for metabolic engineering studies (George et al., 2020). The trans-membrane domains of HMGR enzymes in mammals and yeast have been reported to contain a sterol sensing domain (SSD) that regulates expression and degradation of the enzyme (Kuwabara & Labouesse, 2002). Although our results suggest that diatom HMGR lacks a SSD (Fig. 4.3), we designed an N-terminal truncated version of putative HMGR, tHMGR, to evaluate whether an unknown regulatory sequence is present in N-terminal region affecting activity of HMGR in diatoms. This tHMGR sequence encoded solely the C-terminal catalytically active region of the enzyme.
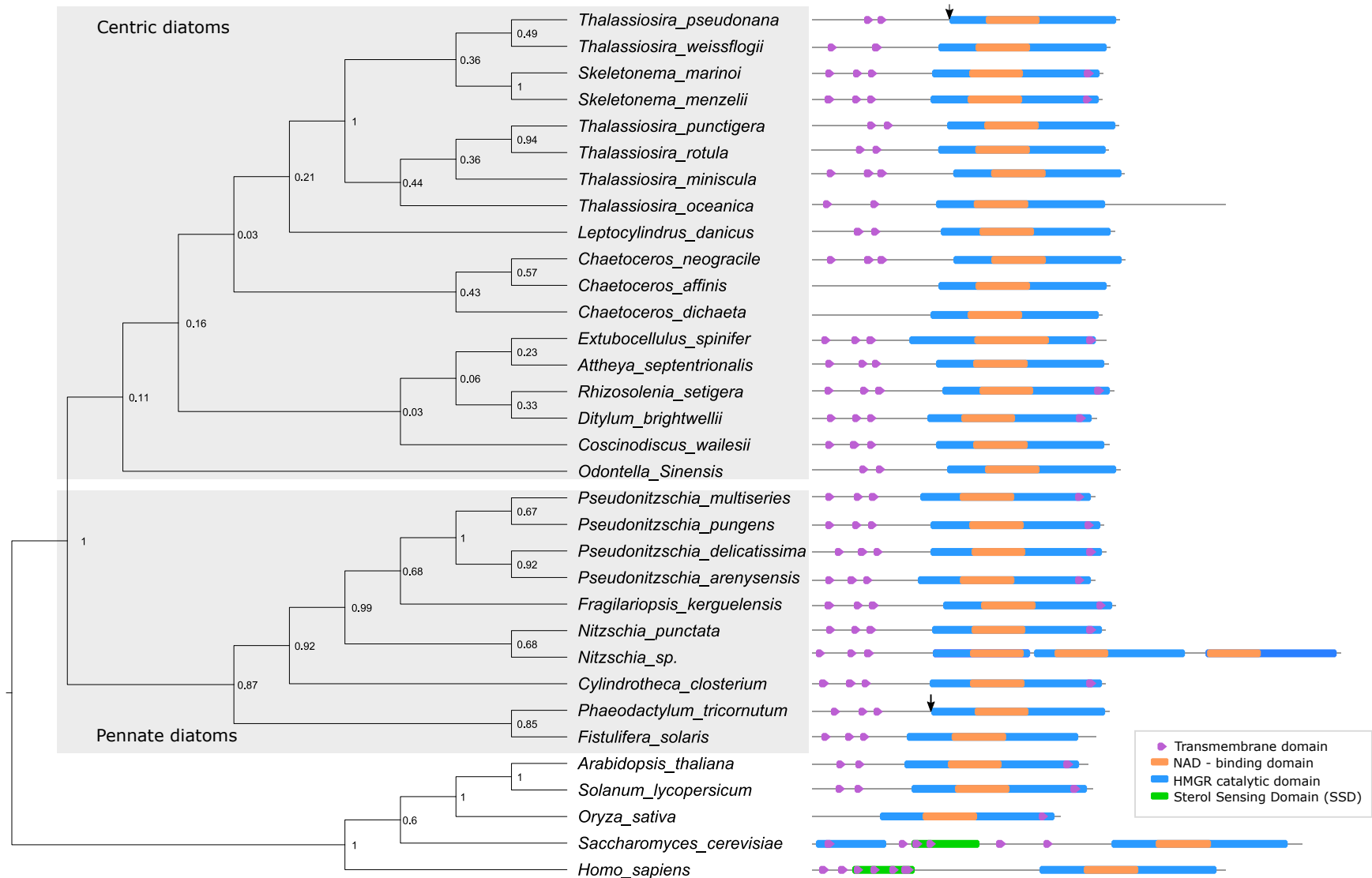
Figure 4.3: Maximum likelihood phylogenetic tree of diatom HMGR proteins and its domains. Numbers at the nodes represent bootstrap support (1000 replicates). HMGR from yeast, mammals and plants were used as outgroups. Arrow represent start of N-terminal truncated version for *P. tricornutum* and *T. pseudonana* used in this study.

Transformation via bacterial conjugation yielded between 50-100 colonies for *T. pseudonana* and around 30-50 colonies for *P. tricornutum* transformed with episomes containing expression cassettes for either HMGR, tHMGR, or the mVenus control. mVenus fluorescence was measured by flow cytometry and used as an indirect proxy of enzyme expression, since each expression system was C-terminal fused with mVenus protein. Three clones per expression system with the highest median mVenus signal were chosen for full scale experiments.

In *P. tricornutum*, time course median mVenus fluorescence in mVenus control clones was 10-fold compared to WT, which confirms the effectiveness of the chosen promoter (Fig. 4.S3). Conversely, median mVenus fluorescence clones expressing HMGR and tHMGR was 1.3-fold compared to WT, indicating an apparent regulation process occurring over the fused proteins (Fig. 4.S3). In *T. pseudonana*, median mVenus fluorescence in HMGR, tHMGR, and mVenus control clones appear similar to WT signal, suggesting that low expression was achieved using the EF2 promoter (Fig. 4.S4). However, confocal microscopy images confirmed expression of mVenus in both diatom species (Figs. 4.4, 4.S3). Different cellular localizations were observed for each genetic construct. Images of control cell lines showed mVenus expression localized in the cytoplasm (Figs. 4.4, 4.S3), while no mVenus fluorescence was detected in WT diatoms. mVenus fluorescence in exconjugants overexpressing HMGR was detected around the chloroplast, suggesting that putative HMGR is localized in the endoplasmic reticulum (ER), which tightly surrounds the chloroplasts in diatoms (Kroth, 2002) (Liu et al., 2016). Conversely, tHMGR localises in the cytoplasm, consistent with truncation of the N-terminal membrane domain (Figs. 4.4, 4.S3).

4.3.4 Influence of HMGR and tHMGR expression on sterol levels in *T. pseudonana* and *P. tricornutum*

Although expression of transgenes appears to be low according to mVenus fluorescence levels, we proceeded to identify changes in sterol profiles in the transgenic lines. Sterols were extracted from exconjugants in the mid-exponential phase, which was the time period with the maximum observed mVenus fluorescence (Figs. 4.S3, 4.S4) and with enough biomass to sample for sterol extraction (determined to be 48 hours growth for *T. pseudonana* and 72 hours for *P. tricornutum*). After 75 hours, cell density of *P. tricornutum* HMGR and tHMGR was 1.4 times lower than WT, while no growth impairment was observed for mVenus exconjugants (Fig. 4.S6). No differences in chlorophyll levels and in effective quantum yield of PSII were observed in *P. tricornutum* exconjugants (Figs. 4.S7, 4.S8). Similarly, no growth impairment, chlorophyll levels or differences in effective quantum yield of PSII compared to WT were observed for *T. pseudonana* exconjugants (Figs. 4.S9, 4.S10, 4.S11).

In *P. tricornutum* over-expressing HMGR, squalene levels were ten times higher in HMGR
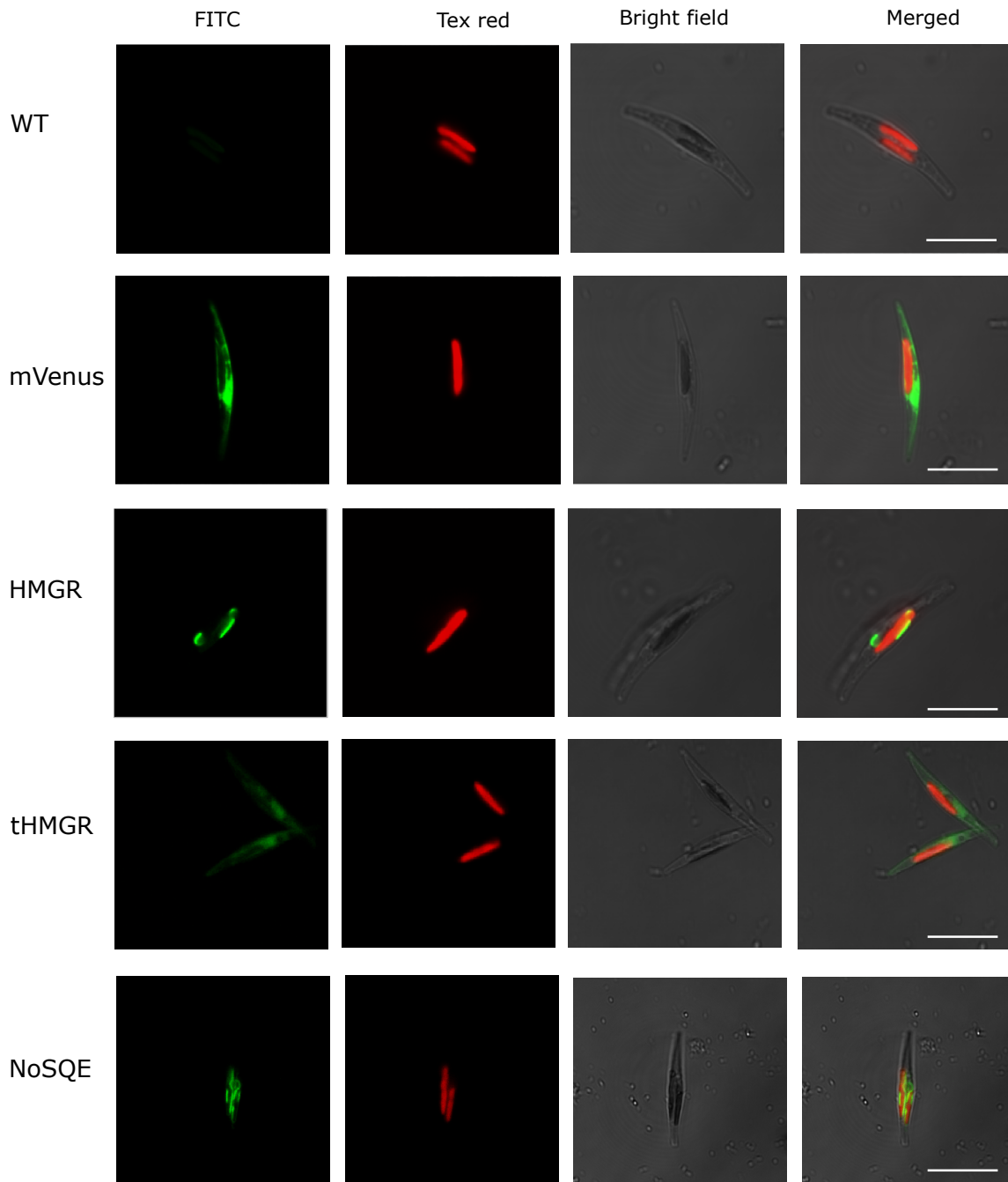
**Figure 4.4:** Confocal microscopy images showing subcellular localization of the mVenus fusion with target proteins in representative transgenic diatoms cells, compared to wild type (WT) as negative control and control cell lines that only expressed mVenus (Venus). 3-hydroxy-3-methylglutaryl-coenzyme A reductase, HMGR, truncated HMGR, tHMGR, squalene epoxidase from N. oceanica, NoSQE. Scale bars correspond to 10 μm.

exconjugants than in WT and mVenus controls. Moreover, a 3-fold increase in cycloartenol and a 2.5-fold obtusifoliol accumulation was detected compared to WT (Fig. 4.5). However, we did not detect the intermediate 2,3 epoxysqualene. Levels of end-point sterol campesterol decreased 2-fold, whereas no significant differences were observed in brassicasterol, the most abundant sterol in *P. tricornutum*. We detected traces levels of end point sterol 24- methylcholesta-5,24(24')-dien-3β-ol in the WT and mVenus controls, which is typically found in centric diatoms and has not been reported in *P. tricornutum* (Rampen et al., 2010). Levels of this end-point sterol were 17 times higher in two independent exconjugant lines expressing HMGR. Similar to *T. pseudonana*, total sterol levels were not affected despite significant changes in individual sterols (Fig. 4.5).

Since expression of mVenus alone, HMGR-mVenus and tHMGR-mVenus in *T. pseudonana* did not appear to be effective based on flow cytometry data (Fig. 4.S5), observed changes on sterol profiles may not be directly related to the over-expression of the targeted enzymes. *T. pseudonana* cell lines transformed with HMGR construct exhibited a decrease in the minor sterols fucosterol and isofucosterol relative to WT control (Fig. 4.S12). However, isofucosterol reduction was also detected in the control expressing only mVenus (Fig. S12). No intermediates were detected. Total sterol levels remained similar in the three independent cell lines studied; no significant differences were observed in comparison to the WT control (Fig. 4.S12).

Expressing the catalytically active region of their putative HMGRs (tHMGR) was expected to reduce regulatory mechanisms that may affect HMGR activity in diatoms. In *P. tricornutum*, sterol changes in cell lines expressing tHMGR were also similar to those expressing HMGR, with campesterol levels 2-fold less abundant than in HMGR transformants (Fig. 4.5). Traces of 24- methylcholesta-5,24(24')-dien-3β-ol were detected as in HMGR expressing lines, 10 times higher compared to WT. We also detected an increase in the intermediates squalene (4-fold), cycloartenol (1.8-fold) and obtusifoliol (2-fold) compared to WT (Fig. 4.5). Squalene levels accumulated in HMGR clones were statistically different to tHMGR clones, being 2.5 times higher in HMGR expressing lines (Fig. 4.5). No changes in total sterol levels and brassicasterol were observed (Fig. 4.5).

No changes in total sterol content were observed in *T. pseudonana* cell lines transformed with tHMGR (Fig. 4.S12). No intermediates were detected. Significant changes in less abundant sterol compounds occurred in transformants including controls expressing only mVenus.
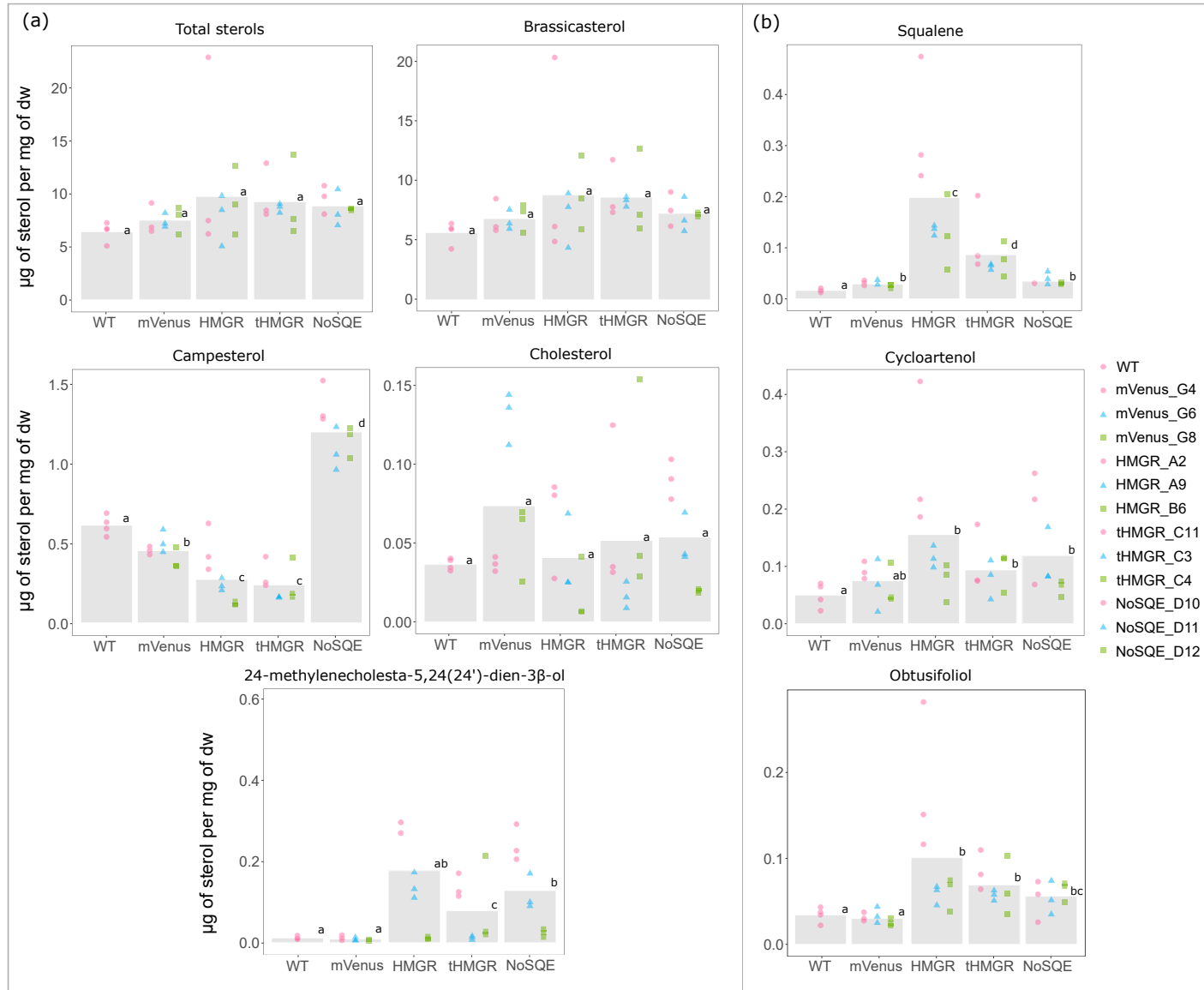
Figure 4.5: Sterol levels in *P. tricornutum* transformants. (a) End-point sterol (b) intermediates accumulation. Identical letters denote no statistically significant differences among groups using the Pairwise Wilcoxon Rank Sum tests (p < 0.05, n = 9).

4.3.5 Heterologous expression of a Stramenopile putative SQE

Diatoms have been reported to employ an alternative squalene epoxidase (AltSQE) that is different from the conventional SQE found in other eukaryotes (Pollier et al., 2019). It has been observed that artificially alter the expression of this enzyme in *P. tricornutum* is particularly challenging (Pollier et al., 2019), suggesting that a strict regulation of endogenous AltSQE may be occurring as is the case for SQE in (Gill et al., 2011; Nagai et al., 2002b). Therefore, we hypothesized that the expression of a heterologous, conventional SQE that could override endogenous regulation, would influence final sterol levels. Consequently, a heterologous putative SQE from the Stramenopile N. oceanica (NoSQE, Nanoce ID 521007) was expressed in the diatoms *T. pseudonana* and *P. tricornutum*.

Similar to HMGR and tHMGR, transformation yield of *T. pseudonana* and *P. tricornutum* with episomes containing expression cassettes for SQE was between 50-100 colonies and 30 – 50 colonies, respectively. Chlorophyll and mVenus fluorescence intensity were comparable to those of diatoms over-expressing putative HMRG and tHMGR (see 3.2 section) (Figs. 4.S4, 4.S5, 4.S7, 4.S10). Confocal microscopy images showed that mVenus fluorescence in NoSQE–mVenus transformants was similarly located to HMGR, indicating ER localization (Figs. 4.4, 4.S3) in both diatom species.

Total sterol content of *T. pseudonana* and *P. tricornutum* exconjugants over-expressing NoSQE remained unchanged compared to WT and mVenus controls (Figs. 4.5, 4.S12). However, in *P. tricornutum* overexpressing NoSQE we observed significant differences in both end-point sterols and sterol intermediates. Downstream intermediates cycloartenol and obtusifoliol exhibited a 1.8-fold increase compared to WT control (Fig. 4.5), while no differences in squalene were observed. The intermediate 2,3 epoxysquelene was not detected in either diatom species. In contrast to HMGR over-expression, campesterol increased by 3-fold (Fig. 4.5). However, the major end-point sterol brassicasterol remained unchanged.

4.4 Discussion

4.4.1 HMGR is largely conserved among diatoms and lacks a conventional sterol sensing domain

To investigate sequence characteristics of the rate-limiting HMGR enzyme in diatoms, we identify the genes putatively encoding the enzyme HMGR from 28 different diatom species. While a putative HMGR homologue was detected in 26 of the diatom species, the failure to detect HMGR transcripts in transcriptomic sequences of some diatoms belonging to *Chaetoceros genus* (Table 4.S2), may suggest that those diatoms may solely rely on the MEP pathway to produce isoprenoids (Jaramillo-Madrid et al., 2020a). The lack of obvious

HMGR transcripts may also have occurred due to low expression or down-regulation of these and other genes related to the MVA pathway under the conditions in which RNA sequencing was performed. Nevertheless, conserved HMGR genes were detected in many diatoms for which genomic or transcriptomic data are available. These HMGR genes diverge between pennate and centric diatoms (Fig. 4.3), which are separated by 90 million years of divergent evolution (Bowler et al., 2008). Presence of putative HMGR in most of the diatom species analysed is an indicator of a functional MVA pathway operating in diatoms, as it was previously reported by transcriptomics analysis in the diatoms *P. tricornutum* and *T. pseudonana* (Jaramillo-Madrid et al., 2020a). The MEP pathway appears functional in the two model diatoms, which indicates that both cytosolic MVA and plastidial MEP are simultaneously operating in *P. tricornutum* and *T. pseudonana*, as it is the case in plants (Jaramillo-Madrid et al., 2020a; Vranová et al., 2013). Presence of both pathways could represent an advantage for terpenoids production, due to a potentially higher metabolic flux and greater pool of available precursors (Vavitsas et al., 2018).

The organization of the N-terminal transmembrane domain of diatom HMGR differs significantly to their mammals and yeast counterparts. While most of the diatoms analyzed in this study presented three membrane spanning sequences (Fig. 4.3, Table 4.S2), mammals possess five and yeast seven. The presence of transmembrane domains is likely related to anchoring the protein within the ER membrane, but the consequences of this structural difference for diatom HMGR in terms of regulation of enzyme expression and activity are unknown. In mammals and yeast, HMGR possess a SSD involved in sensing oxysterol molecules that activate feedback regulation leading to degradation of the protein (Burg & Espenshade, 2011) (Theesfeld et al., 2011). Despite the lack of a conventional sterol-sensing domain in the HMGR enzymes of diatoms (Fig. 4.3), several studies have shown a transcriptional response of MVA enzymes to perturbations in sterol metabolism (D'Adamo et al., 2019)(Jaramillo-Madrid et al., 2020a). The fact that in diatom sequences do not possess a canonical SSD opens several possibilities, including that HMGR may not play the same regulatory role in diatoms as it does in other organisms; i.e. the MVA pathway may be regulated through a different mechanism that does not involve HMGR feedback regulation. Another possible explanation is that HMGR from diatoms and plants possess a non-conventional SSD sequence, a motif with different characteristics than the already described in mammals and yeast.

4.4.2 HMGR and tHMGR over-expression lead to accumulation of sterol pathway intermediates in *P. tricornutum*

In this study, we investigated the response of diatoms to genetic targeting of sterol biosynthesis, through manipulation of the MVA rate-limiting enzyme HMGR. The fluorescent localization of extra-chromosomally expressed HMGR to the membrane surrounding the

plastid is consistent with ER localization of proteins from previous studies in *P. tricornutum* (Kroth, 2002; Pollier et al., 2019)and *T. pseudonana* (Sheppard et al., 2010). Therefore, our study provides evidence that diatom HMGR is localized to the ER, just as in mammals, yeast, and higher plants (Leivar et al., 2005) (Burg & Espenshade, 2011). The truncation of the native HMGR sequence resulted in cytoplasmic localization, demonstrating that the signals for protein targeting are in the N-terminal portion of the protein sequence.

In some cases, over-expression of HMGR has been showed to slightly increase sterol and intermediates production in other organisms. The over-expression of HMGR from the plant *Panax ginseng* in *Arabidopsis thaliana* resulted in a nearly two-fold increase of sitosterol, campesterol, and cycloartenol, while levels of squalene and stigmasterol did not significantly change (Kim et al., 2014). In a more recent study, (Lange et al., 2015) independently over-expressed all genes participating in the MVA pathway, obtaining a significant increase in total sterols when expressing HMGR (3.4-fold) and 3-hydroxy-3-methylglutaryl-co-enzyme-A synthase, HMGS (2-fold). The over-expression of native HMGR in *Arabidopsis* (HMG1) led to high levels of HMGR mRNA, but only a slight increase in HMGR activity, and no changes in leaf sterol levels (Re et al., 1995). Heterologous expression of HMGR from Hevea brasiliensis in tobacco, however, showed an increase in HMGR transcript and total sterol from leaves (Schaller et al., 1995).

In this study, over-expression of endogenous putative HMGR in *P. tricornutum* resulted in an increased accumulation of the intermediates squalene, cycloartenol, and obtusifoliol and decrease in the end-point sterol campesterol. However, we did not detect 2,3 epoxysquelene, indicating differences on the catalytic rates of the enzymes squalene epoxidase and cycloartenol cyclase (Fig. 4.1). While accumulation of squalene is detectable, 2,3 epoxysquelene seems to be rapidly converted into cycloartenol. Accumulation of 2,3 epoxysquelene has been reported before by chemically inhibiting cycloartenol cyclase with Ro 48-8071 (Fabris et al., 2014; Jaramillo-Madrid et al., 2020a).

Accumulation of intermediates suggest that in *P. tricornutum*, over-expression of HMGR boosted production of presumably MVA-derived intermediates, IPP and DMAPP that are subsequently converted into squalene (Fig. 4.1). Even though MVA products also serve as building blocks of other isoprenoids (Lange et al., 2000), perturbation of the rate-limiting step catalyzed by HMGR was sufficient to cause accumulation of downstream intermediates committed to sterol biosynthesis, such squalene, cycloartenol, and obtusifoliol. However, this metabolic bottleneck did not translate into overall increase of sterol compounds; on the contrary, levels of the end-point sterol campesterol were reduced and brassicasterol levels remained unchanged.

Although, fluoresce levels in *P. tricornutum* exconjugants expressing the target enzymes was considerably diminished compared to the expression of mVenus alone (Fig. 4.S5), phenotypic changes in terms of sterol profiles indicates that the level of expression was sufficient to cause disruption in the metabolic pathway (Fig. 4.5). Moreover, since all the

enzymes expressed are membrane proteins, as confirmed by confocal microscopy (Figs. 4.4, 4.S3), fluoresce signal could have been hindered and no direct correlation with expression could be assumed. The absence of detectable sterol pathway intermediates in *T. pseudonana* transformants (Fig. 4.S12) may be related with the promoter chosen for over-expression of the target enzymes, as higher expression may be necessary to instigate the intermediate accumulation observed in *P. tricornutum*. Fluorescence signal for mVenus control in *P. tricornutum* (Phatr3_J49202 promoter) was around ten times higher than in *T. pseudonana* mVenus control (EF2 promoter) through the full-scale experiment (Figs. 4.S4, 4.S5), suggesting that use of stronger promoters for metabolic engineering of *T. pseudonana* should be considered. Levels of fucosterol and isofucosterol decreased in *T. pseudonana* lines overexpressing a putative endogenous HMGR (Fig. 4.S12). However, since similar results were observed for isofucosterol in the control expressing only mVenus, it is uncertain if the observed reduction was a direct consequence of putative HMGR over-expression alone. No intermediates were detected in *T. pseudonana* expressing putative HMGR. End-point sterols 24-methylenecholesta-5,24(24')-dien-3β-ol, cholesterol, campesterol, and total sterol levels were statistically indistinguishable from those obtained with untransformed WT and empty vector control transformants (Fig. 4.S12).

Results suggest that there are several regulation points in sterol biosynthesis in diatoms, including the MVA pathway, conserved core, and specialized downstream reactions. It is also possible that MEP responds to an alteration on the MVA pathway, rebalancing IPP and DMAPP pools, and metabolic cross-talk between this two pathways could potentially occur in diatoms. Despite the core reactions in sterol synthesis being conserved in *T. pseudonana* and *P. tricornutum* (Jaramillo-Madrid et al., 2020a), we observed different responses in lines over-expressing putative HMGR of these two model diatoms (Figs. 4.5, 4.S12). As previously mentioned, these results might be related with the promoters chosen for each species, or could indicate differences on sterol regulation between centric and pennate diatoms that correlates with divergence between putative HMGR from both diatom groups (Fig. 4.3). Differences in sterol profiles from *T. pseudonana* and *P. tricornutum* has been suggested to occur in downstream reactions of sterol synthesis (Jaramillo-Madrid et al., 2020a), and responses to alteration on a key point of MVA pathway suggests possible divergences in regulation mechanisms (this study).

Additional strategies have been developed to avoid regulation of HMGR and increase MVA carbon flux. Truncation of HMGR to remove N-terminal membrane and SSD domain was first reported in plants and yeast with the aim to express only the catalytic domain of HMGR and avoid regulatory effects (Donald et al., 1997; Polakowski et al., 1998). Although HMGR from plants do not contain an SSD sequence (Fig. 4.3), expression of an N-terminal truncated HMGR has been reported to increase sterol levels. Expression of tHMGR from hamster in tobacco resulted in augmented sterol content in leaf tissue (Chappell et al., 1995). Constitutive expression of tHMGR from Hevea brasiliensis in tobacco

resulted in an increasing 11-fold of seed HMGR activities and 2.4-fold increase in total seed sterol content (Harker et al., 2003). However, over-expression of a tHMGR has not always been effective at altering sterol content; the over-expression of tHMGR in yeast resulted in accumulation of squalene, no changes in ergosterol, the final sterol in yeast (Donald et al., 1997; Polakowski et al., 1998). Likewise, in this study we did not observe a statistically significant alteration in total sterols of *T. pseudonana* and *P. tricornutum* after constitutive expression of a tHMGR (Figs. 4.5, 4.S12). Yet, accumulation of the intermediates squalene, cycloartenol, and obtusifoliol was observed in *P. tricornutum*. Interestingly, levels of those intermediates were higher when expressing HMGR, suggesting that truncation may have affected enzyme activity, performance, or access to substrates. Although we observed changes in intermediates and minor sterol levels in *P. tricornutum* expressing HMGR and tHMGR, total sterol levels remained unchanged. These results suggest that diatoms have a tight sterol regulation system that may not be related to the conventional regulation model that involves SSD but rather a complex system with several regulation points not only in the MVA pathway but further down the sterol metabolic pathway.

4.4.3 Levels of end-point campesterol increased after heterologous expression of SQE in *P. tricornutum*

Diatoms possess a distinct AltSQE (Pollier et al., 2019) catalyzing the conversion of squalene into 2,3 epoxysqualene which is then transformed into cycloartenol, the first committed step towards the production of steroids (Fig. 4.1). Whether the presence of an AltSQE confers diatoms with biological advantages is not yet known. Similarly, AltSQE and SQE are mutually exclusive and, to date, no organisms have been found to naturally harbor both (Pollier et al., 2019).

The genetic manipulation of SQE and squalene synthase, SQS, has been extensively used for enhanced production of squalene and triterpenoids (Dong et al., 2018; Lee. et al., 2004) (Gohil et al., 2019). Point mutations in the SQE gene (ERG1) in yeast resulted in accumulation of squalene (Garaiová et al., 2014). Similarly, accumulation of squalene was observed in the green microalgae *Chlamydomonas reinhardtii* after knocking-down the SQE gene, while co-transformation lines with SQE-over-expression and SQE-knockdown yielded similar amounts of squalene (Kajikawa et al., 2015).

This study is the first attempt at elucidating the response of diatoms to the expression of a conventional SQE. We did not observe any significant changes in phenotype of *T. pseudonana* and *P. tricornutum* expressing heterologous NoSQE (Figs. 4.S6-4.S11S11). This suggest that in diatoms there is no apparent toxicity or physiological reason for the mutual exclusivity between AltSQE and conventional SQE. Confocal microscopy images of lines expressing NoSQE-mVenus fusion proteins, revealed that heterologous NoSQE was proximal to the chloroplasts, indicating that diatoms could recognize the ER signal peptide on the het-

erologous NoSQE, localizing it in the ER membrane (Figs. 4.4, 4.S3), just as the endogenous AltSQE (Pollier et al., 2019) and native SQE enzymes are in other species (Laranjeira et al., 2015; Leber et al., 1998).

Significant accumulation of cycloartenol (1.8-fold) and obtusifoliol (1.8-fold) intermediates, but not of 2,3 epoxysqualene was obtained for *P. tricornutum* lines expressing NoSQE (Fig. 4.5b). These intermediates occur after the formation of 2,3 epoxysqualene, which is the product of the reaction catalyzed by SQE (Fig. 4.1). As expected, we did not observed increased accumulation of squalene, which is the substrate for SQE, contrary to accumulation obtained by expressing HMGR which is upstream squalene production (Fig. 4.1). Nevertheless, heterologous expression of NoSQE resulted in a 2-fold increase of campesterol, an end-point sterol. Accumulation of intermediates was higher in *P. tricornutum* cell lines expressing HMGR compared to those expressing NoSQE (Fig. 4.5a). This indicates that intermediates accumulated by MVA pathway manipulation (i.e. HMGR) do not necessarily increase the flux to brassicasterol, suggesting that sterol regulation is occurring at the conserved core point and at other points further down the metabolic pathway. In particular, this might suggest that in *P. tricornutum* the epoxidation of squalene might be involved in pathway flux modulation, as observed in mammals (Gill et al., 2011; Nagai et al., 2002b), and that, to complete the scenario suggested by the results obtained by expressing NoSQE, it is plausible that an additional pathway checkpoint exists at the level of the C_{22}-desaturation (E.C 1.14.19.41, Phatr3_J51757) (Fabris et al., 2014). When treated with fluconazole and fenpropimorph, inhibitors targeting upstream of campesterol, the transcription of Phatr3_J51757 significantly increases (Jaramillo-Madrid et al., 2020a). This further supports that this last reaction in sterol synthesis could be a highly regulated point to maintain stable sterol levels in the cell.

From a biotechnological perspective, our results suggest that to observe changes in end-point compounds, increasing of precursors pool is not enough; genetic manipulation should target other points further down in the metabolic pathway, such as committed steps in sterol synthesis. Availability of intermediates in transformed lines could be used by heterologous pathways plugged to the endogenous (tri)terpenoid synthesis, allowing production of other high-value terpenoids as geraniol (Fabris et al., 2020).

In this study, we provide important insights into the metabolic bottleneck and pathway-level regulation of sterol synthesis. We obtained accumulation of sterol pathway intermediates by over-expression of HMGR, indicating possible metabolic bottleneck(s) downstream of the MVA pathway that may limit flux into end-point sterols. A future co-expression approach to increase end-point sterol compounds in *P. tricornutum* could involve simultaneous expression of enzymes in the conserved core (i.e. SQE, AltSQE, cycloartenol synthase) and enzymes further down such as sterol C_{22} desaturase. Similar co-expression approaches for manipulation of sterol levels in diatoms has not been reported but has proven to be a strategy to increase triterpenoid production in other organisms (Bröker et al., 2018; Dong

et al., 2018; Lange et al., 2015; Zhang et al., 2017).

4.5 Conclusions

The results obtained in this study prove effectiveness of extra-chromosomal expression of key enzymes involved in sterol synthesis to influence levels of specific sterol compounds. We confirmed reported advantages of the use of extra-chromosomal episomes transformed via conjugation such as expression consistency among clones (Figs. 4.S4-4.S11) and no random genome integration (George et al., 2020). Additionally, we demonstrated the convenience of a modular assembling systems as uLoop to build versatile genetic constructs for a functional genetics study with multiple species.

Furthermore, we applied reproducible genetic transformation methods for extra-chromosomal and heterologous expression to probe the biosynthesis of sterol compounds in diatoms. We obtained transgenic lines of diatoms that expressed fluorescent fusion proteins consisting of target enzymes and the mVenus reporter gene. Whilst significant accumulation of intermediates participating in sterol synthesis was observed in *P. tricornutum* transformants. *T. pseudonana* and *P. tricornutum*, transformants did not appear to produce different levels of total sterols. It is presumed that several levels of regulation could be affecting the expression, localization, lifetime, and activities of these introduced genes. The regulatory processes operating in diatoms with regard to sterol homeostasis remains elusive. Further research into the regulatory responses of diatoms to the over-expression and heterologous genes might provide further insights into these processes and improve strategies for more detailed metabolic engineering.

Acknowledgements

Supplemental material

Table 4.S1: L0 parts for construction of episomes using uLoop assembly method (Pollak et al., 2018). Primers sequences used for domestication are presented for L0 each part.

LO part	Gene ID	Description	Template	Primers	Plasmid number
AC_pTpEF2	269148	Elongation factor promoter	*P. tricornutum* genomic DNA	F:AAGCTCTTCATCCggagAGTGTGCAA TGCAGTCAATTCAATAGATATG R:TTGCTCTTCTTCGcattCTTGACG TTCTTTTCTCTTTAATTAATCGCG	
AC_pPtEF2	Phatr3_J35766	Elongation factor promoter	*P. tricornutum* genomic DNA	F:AAGCTCTTCATCCggagTCCATTTTG ACATGTTTCCTAGCTAGAAG R:TTGCTCTTCTTCGcattTGTGTGGAG AGAACGAGCAGCAGCG	
AC_p49202	Phatr3_J49202	Predicted protein JCVI	NA	NA	
AC_CENARSHIS	NA	CEN6-ARSH4-HIS3 episome maintenance sequence from PtPBR11 (Genebank KX523203) (Diner et al., 2016)	NA	NA	
CF_OriT	NA	Origin of transfer, originally called basis of mobilization (bom) (Diner et al., 2016)	PtPBR11 (Genebank KX523203)	F:CAGAAGCTCTTCATCCaatgGATCGT CTTGCCTTGCTCGTCGG R:TTGCTCTTCTTCGagcgATCTTCCGC TGCATAACCCTGCTTCGG	139948
CD_NoSQE	521007	0.0	NA	NA	
CD_TpHMGR	33680	3-hydroxy-3-methylglutaryl-coenzyme A reductase (EC: 1.1.1.34)	*P. tricornutum* genomic DNA	F:AAGCTCTTCATCCaatgAGCCCCAAC GACCCACCCGTCAAAG R:TTGCTCTTCTTCGacctgaCTTTGAA GCAGGCTTGCGATTATGTTG	
CD_PtHMGR	Phatr3_J16870	Truncated 3-hydroxy-3-methylglutaryl-coenzyme A reductase	*P. tricornutum* genomic DNA	F:AAGCTCTTCATCCaatgACGGTGACT ATCAGCAGTAGTATTAG R:TTGCTCTTCTTCGacctgaCTTGGCG GCGGGTTTGCGGTTG	
CD_PttHMGR	Phatr3_J16870	Truncated 3-hydroxy-3-methylglutaryl-coenzyme A reductase	*P. tricornutum* genomic DNA	F:AAGCTCTTCATCCaatgGACTCTATT TCCACCAAGACCAGCGCG R:TTGCTCTTCTTCGacctgaCTTGGCG GCGGGTTTGCGGTTG	
CD_bsd	NA	blasticidin-S deaminase	pST1374-N-NLS-flag-linker-Cas9-D10A (Addgene #51130)	F:AAGCTCTTCATCCaatgGCCAAGCCT TTGTCTCAAGAAGAATCCAC R:TTGCTCTTCTTCGacctgaGCCCTCC CACACATAACCAGAGGG	139945
CD_nat	NA	nourseothricin acetyltransferase	pAGM4723:TpCC_ Urease (Addgene #85982)	F:AAGCTCTTCATCCaatgACCACTCTT GACGACACGGCTTACCGG R:TTGCTCTTCTTCGacctgaGGGGCAG GGCATGCTCATGTAGAGCG	
CD_mVenus		Yellow fluorescent protein	NA	NA	

(continued on next page...)

Table 4.S1 ...continued

LO part	Gene ID	Description	Template	Primers	Plasmid number
DE_Thr6XHis FLAG	NA	General C-terminal protein tag. Thormbine, 6xHIS and FLAG (DYKD-DDDK) tag	DE_Venus-ThrHISFLAG L0 part	F:AAGCTCTTCATCCaggtGCGCTGGTC CCTCGCGGTAG R:TTGCTCTTCTTCGaagcTCACTTATC GTCATCATCCTTGTAGTCG	139949
DE_Venus-ThrHISFLAG	NA	Yellow fluorescent protein with a Thormbine, 6xHIS and FLAG (DYKD-DDDK) tag JCVI	NA	NA	
DE-3xStop	NA	Three stop codon JCVI	NA	NA	
EF_tPtEF2	269148	Elongation factor terminator	*P. tricornutum* genomic DNA	F:AAGCTCTTCATCCGCTTACAGAAAAA CAGACTCATAGGGTAC R:TTGCTCTTCTTCGAGCGGAGACATTA CTCCACACGATGG	
EF_tTpEF2	Phatr3_J35766	Elongation factor terminator	*P. tricornutum* genomic DNA	F:AAGCTCTTCATCCgcttATATCTTCT TGCAACAATGGTAGCG R:TTGCTCTTCTTCGagcgAGCAGGGTT GGTTAGAGATAACTAATG	
EF_t49202	Phatr3_J49202	Predicted protein JCVI	NA	NA	

Table 4.S2: Source and number of transmembrane domains of HMGR sequences used for phylogenetic and domain analysis.

Organism	MMETSP ID	Present	Source	Strain	Gene ID	TM*
Attheya septentrionalis	MMETSP1449	Y	Transcriptomics	CCMP2084	Transcript_1390	3
Chaetoceros affinis	MMETSP0091	Y	Transcriptomics	CCMP159	Transcript_17836	0
Chaetoceros brevis	MMETSP1435	N	Transcriptomics	CCMP164	NA	NA
Chaetoceros curvisetus	MMETSP0716	N	Transcriptomics	Unknown	NA	NA
Chaetoceros debilis	MMETSP0149	N	Transcriptomics	MM31A-1	NA	NA
Chaetoceros dichaeta	MMETSP1447	Y	Transcriptomics	CCMP1751	Transcript_4115	0
Chaetoceros muelleri	NA	N	Transcriptomics	CCMP1316	NA	NA
Chaetoceros neogracile	MMETSP0752	Y	Transcriptomics	CCMP1317	Transcript_689	3
Chaetoceros sp.	MMETSP1429	Y	Transcriptomics	UNC1202	Transcript_18209	0
Coscinodiscus wailesii	MMETSP1066	Y	Transcriptomics	CCMP2513	Transcript_2372	3
Cylindrotheca closterium	MMETSP0017	Y	Transcriptomics	KMMCC:B-181	Transcript_33356	3
Ditylum brightwelli	MMETSP1001	Y	Transcriptomics	GSO105	Transcript_14074	3
Extubocellulus spinifer	MMETSP0696	Y	Transcriptomics	CCMP396	Transcript_715	3
Fistulifera solaris	NA	Y	Genome	JPCC DA0580	GAX29119.1	3
Fragilariopsis kerguelensis	MMETSP0909	Y	Transcriptomics	L2-C3	CAMPEP_ 0196091342	2
Homo sapiens	NA	Y	Genome	NA	3156	5
Leptocylindrus danicus	MMETSP0321	Y	Transcriptomics	B650	Transcript_1399	2
Nitzschia punctate	MMETSP0747	Y	Transcriptomics	CCMP561	Transcript_27671	3
Nitzschia sp	MMETSP0014	Y	Transcriptomics	RCC80	Transcript_11673	3
Odontella Sinensis	MMETSP0160	Y	Transcriptomics	Grunow 1884	Transcript_15934	3
Oryza sativa	NA	Y	Genome	NA	LOC_Os02g48330	0
Phaeodactylum tricornutum	NA	Y	Genome	CCMP632	Phatr3_J49202	3
Pseudonitzschia arenysensis	MMETSP0329	Y	Transcriptomics	B593	CAMPEP_ 0116128146	3
Pseudonitzschia delicatissima	MMETSP0327	Y	Transcriptomics	B596	CAMPEP_ 0116102328	3
Pseudonitzschia multiseries	NA	Y	Genome	CLN-47	288249	3
Pseudonitzschia pungens	MMETSP1061	Y	Transcriptomics	cf. pungens	Transcript_13267	3
Rhizosolenia setigera	MMETSP0789	Y	Transcriptomics	CCMP1694	CAMPEP_ 0178949462	3
Saccharomyces cerevisiae	NA	Y	Genome	ATCC 204508	854900	7
Skeletonema marinoi	MMETSP1428	Y	Transcriptomics	UNC1201	Transcript_20411	3
Skeletonema menzelii	MMETSP0603	Y	Transcriptomics	CCMP793	Transcript_7316	3
Solanum lycopersicum	NA	Y	Genome	NA	Solyc02g082260.3	2
Thalassiosira miniscula	MMETSP0737	Y	Transcriptomics	CCMP1093	Transcript_23358	2
Thalassiosira oceanica	NA	Y	Genome	CCMP1005	91521	2
Thalassiosira pseudonana	NA	Y	Genome	CCMP1335	269148	3
Thalassiosira puntigera	MMETSP1067	Y	Transcriptomics	Tpunct2005C2	Transcript_37597	3
Thalassiosira rotula	MMETSP0403	Y	Transcriptomics	CCMP3096	Transcript_15672	2
Thalassiosira weissflogii	MMETSP1414	Y	Transcriptomics	CCMP1010	Transcript_6235	3
Thalassiothrix antarctica	MMETSP0152	Y	Transcriptomics	L6-D1	Transcript_7954	3

*TM indicates number of transmembrane domains predicted by TMHMM Server v. 2.0 (http://www.cbs.dtu. dk/services/TMHMM-2.0/). One of the three transmembrane domains in *Arabidopsis* is located in C-terminus domain.
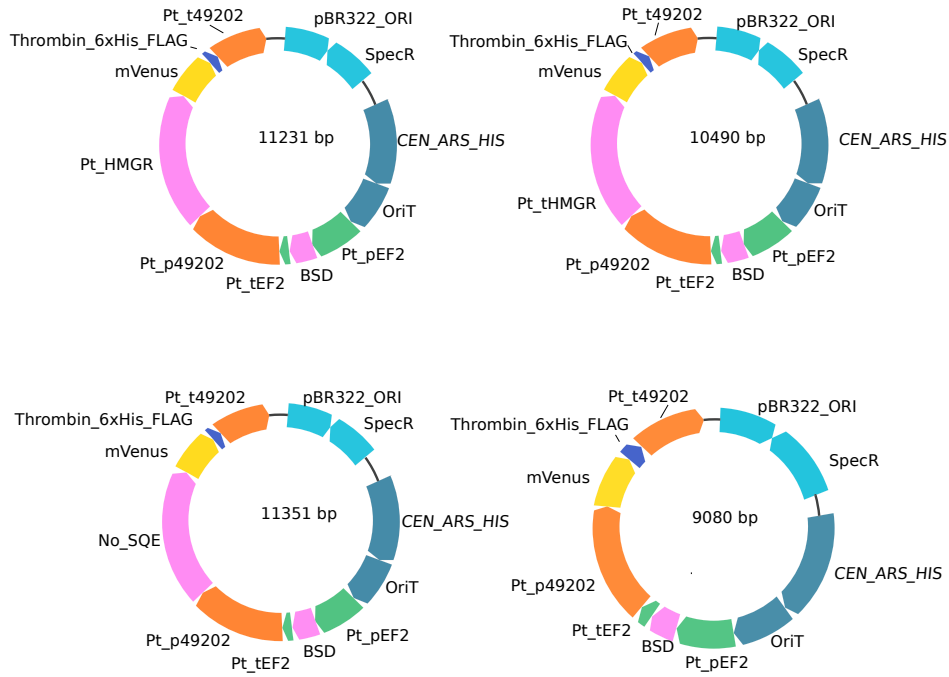
Figure 4.S1: Maps of plasmids used for *T. pseudonana* transformation. CDS are shown in pink, in green promoters and terminators and fluorescence mVenus gene in yellow.
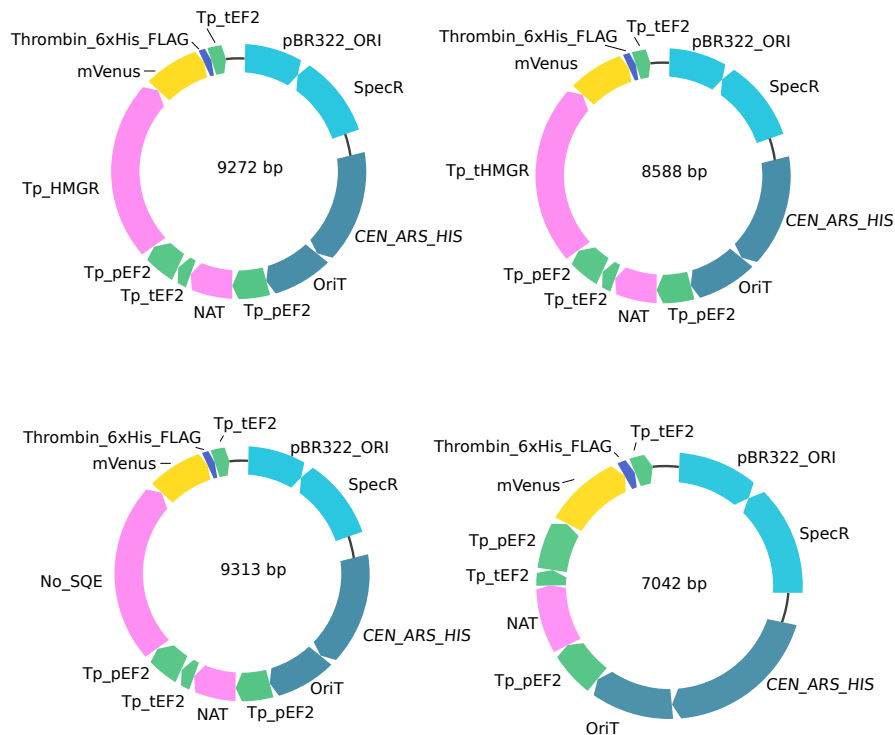
Figure 4.S2: Maps of plasmids used for *P. tricornutum* transformation. CDS are shown in pink, in green promoters and terminators and fluorescence mVenus gene in yellow.
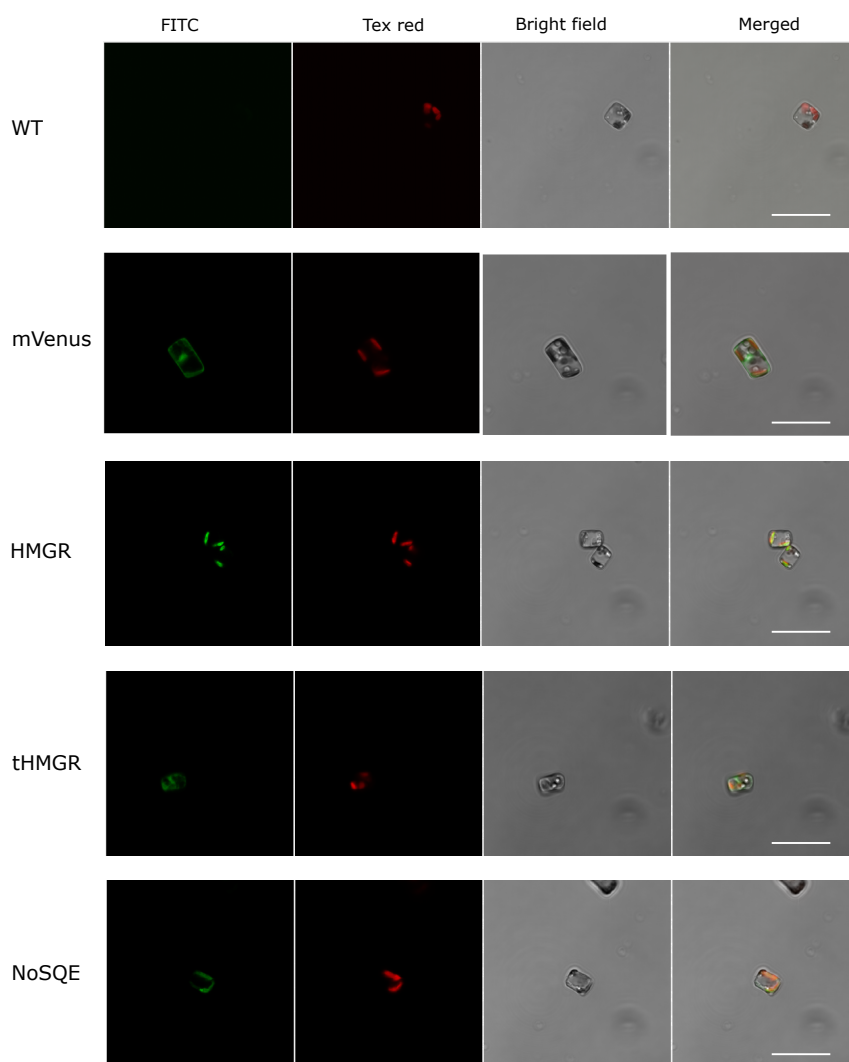
Figure 4.S3: Screening of transformed *T. pseudonana* clones. Colonies were transferred from selection agar plates to liquid medium with 100 µg/mL nourseothricin in 96 well plates, after three days, flow cytometry measurements were taken.
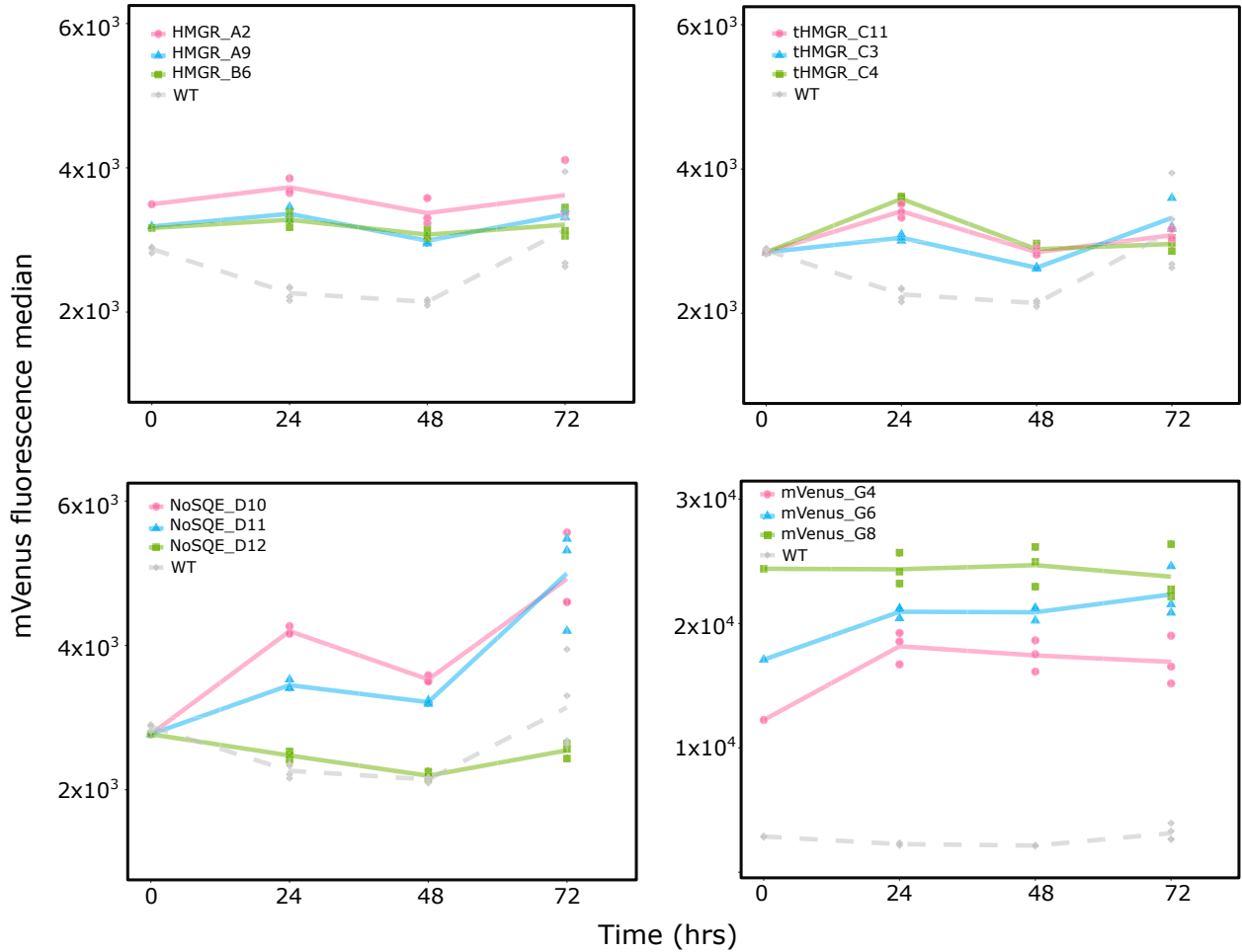
Figure 4.S4: Screening of transformed P. tricornutum clones. Colonies were transferred from selection agar plates to liquid medium with 10 μg/mL blasticidin in 96 well plates, after three days, flow cytometry measurements were taken.
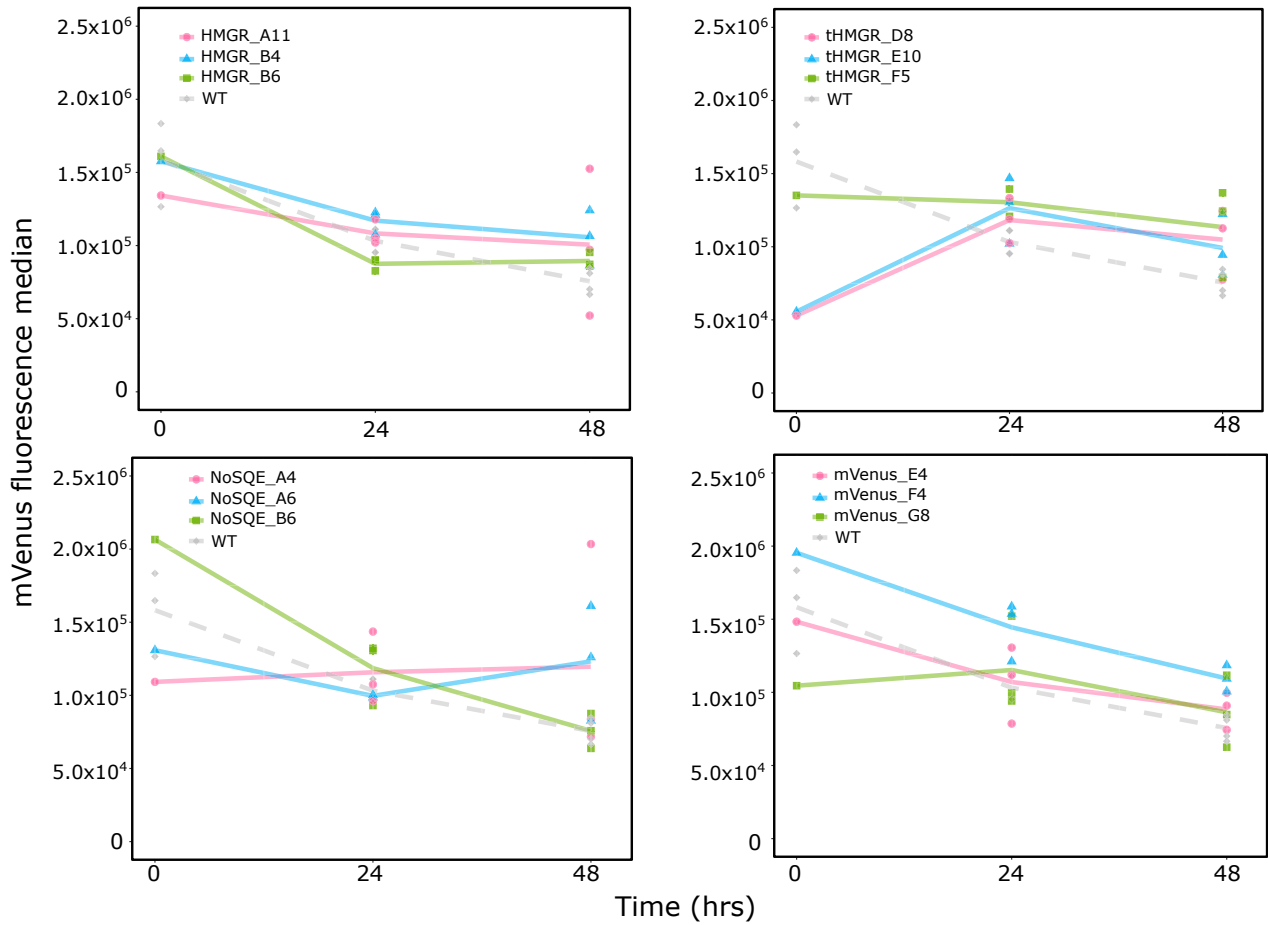
Figure 4.S5: mVenus fluorescence during full scale experiment in *T. pseudonana* transformants (n = 3).
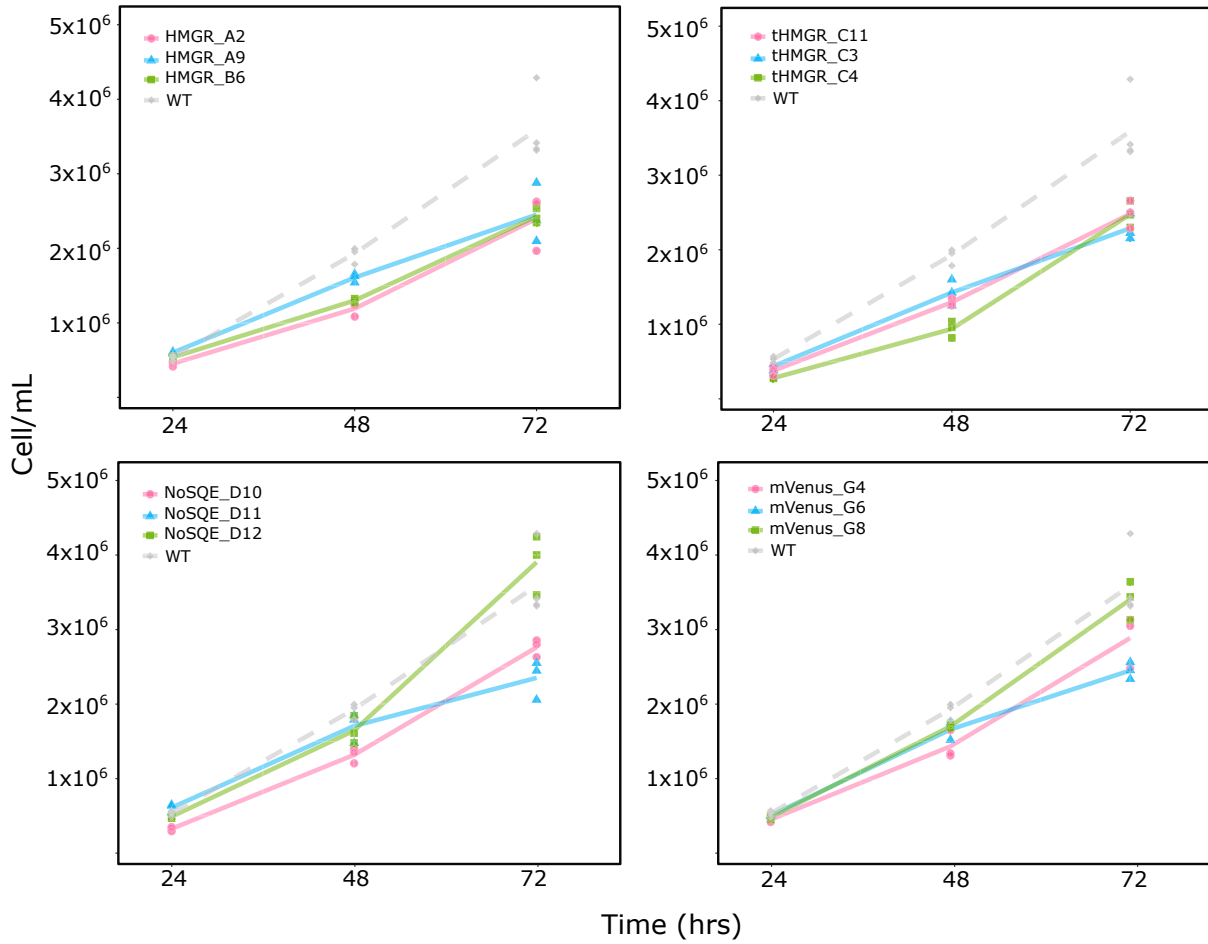
Figure 4.S6: Chlorophyll fluorescence during full scale experiment in *T. pseudonana* transformants (n = 3).
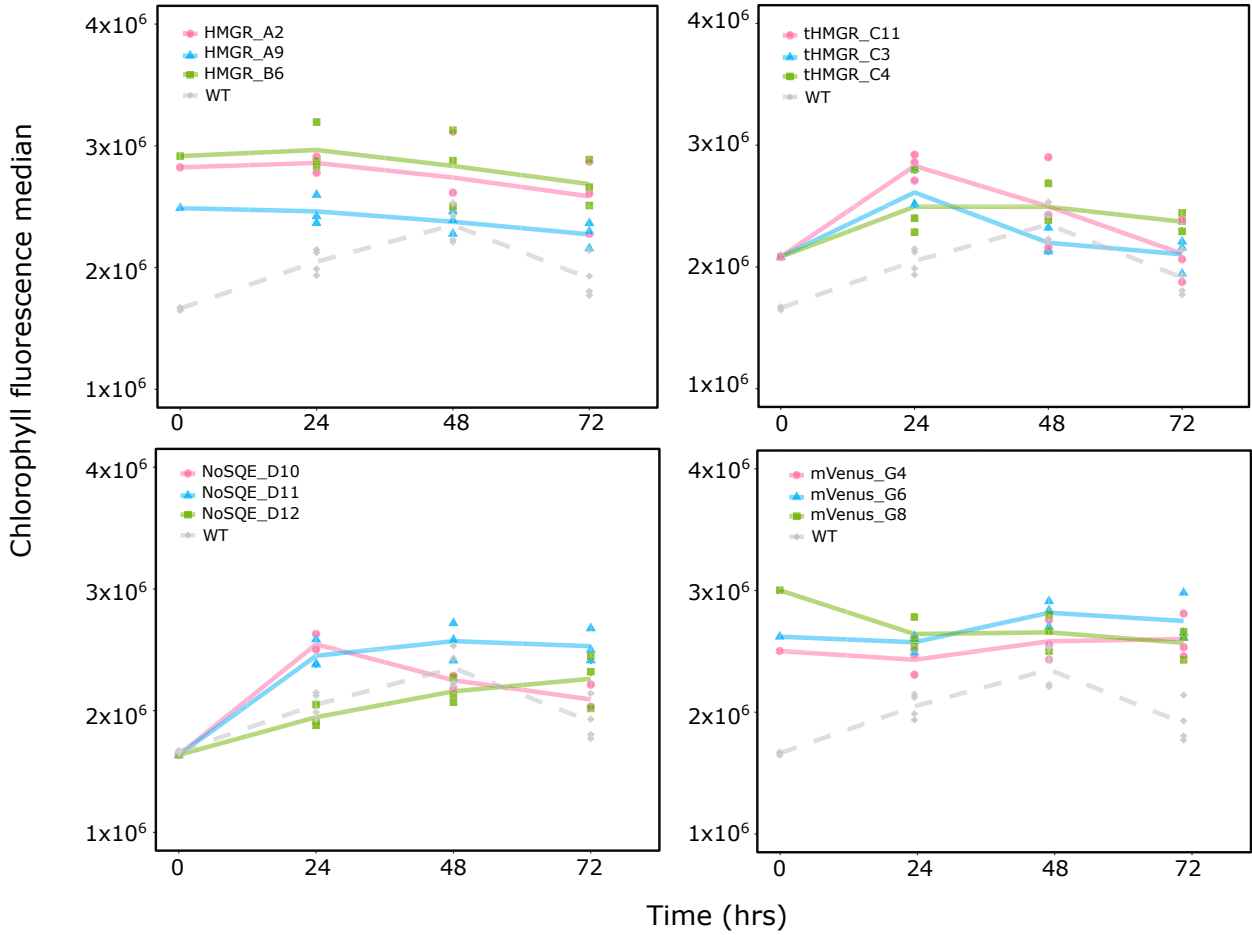
Figure 4.S7: Growth curves during full scale experiment for *T. pseudonana* transformants (n = 3).
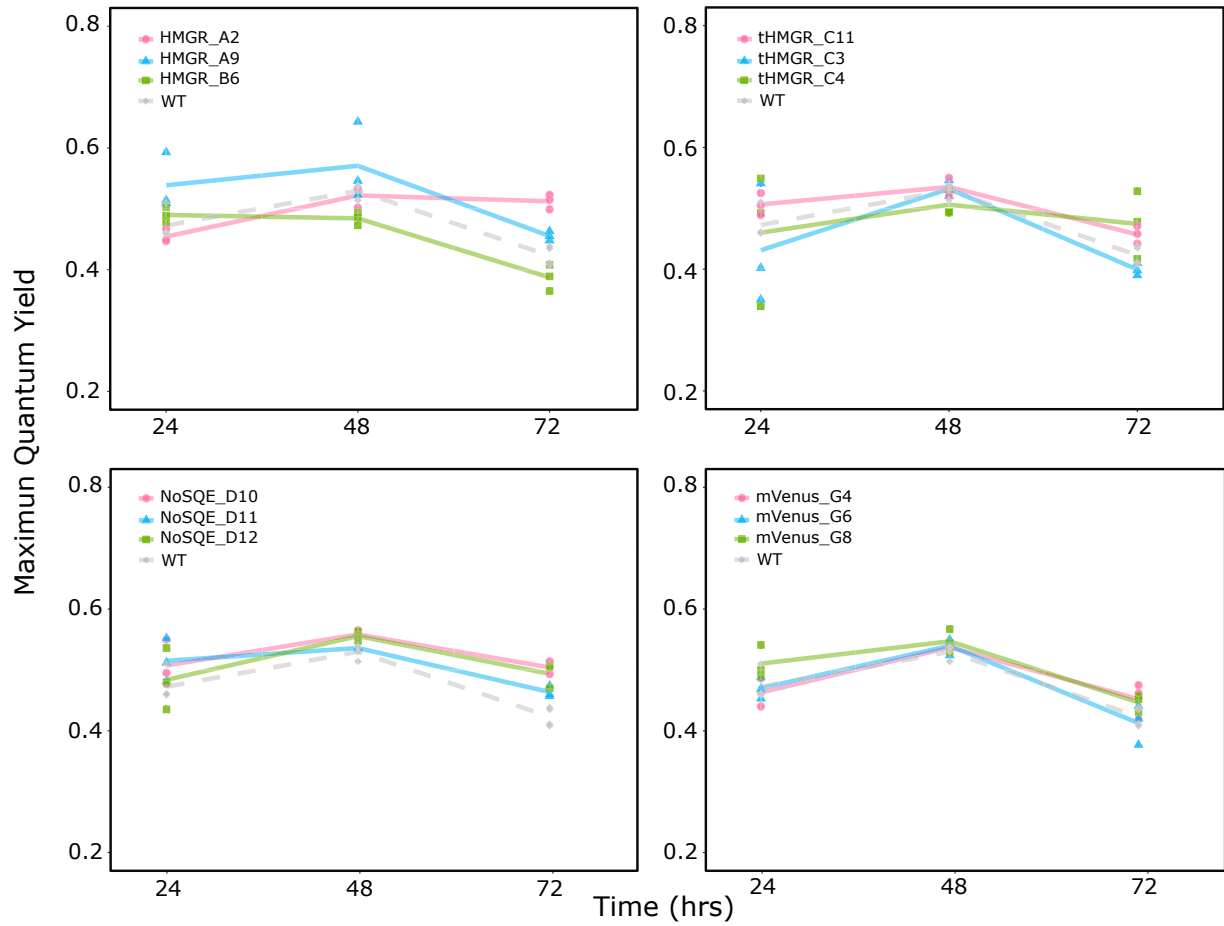
Figure 4.S8: Maximum quantum yield for *T. pseudonana* transformants during full scale experiment (n = 3).
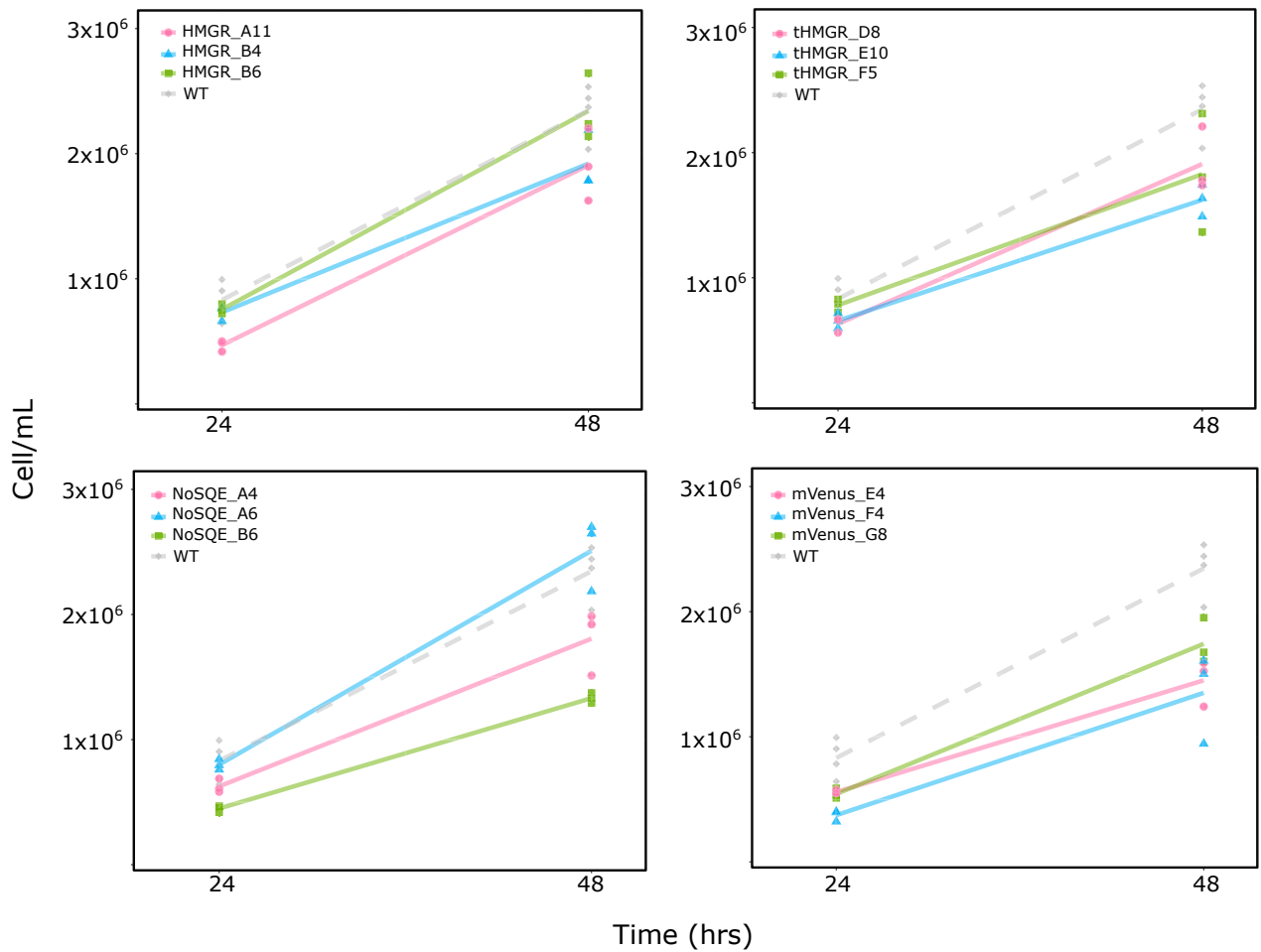
Figure 4.S9: mVenus fluorescence during full scale experiment in *P. tricornutum* transformants (n = 3).
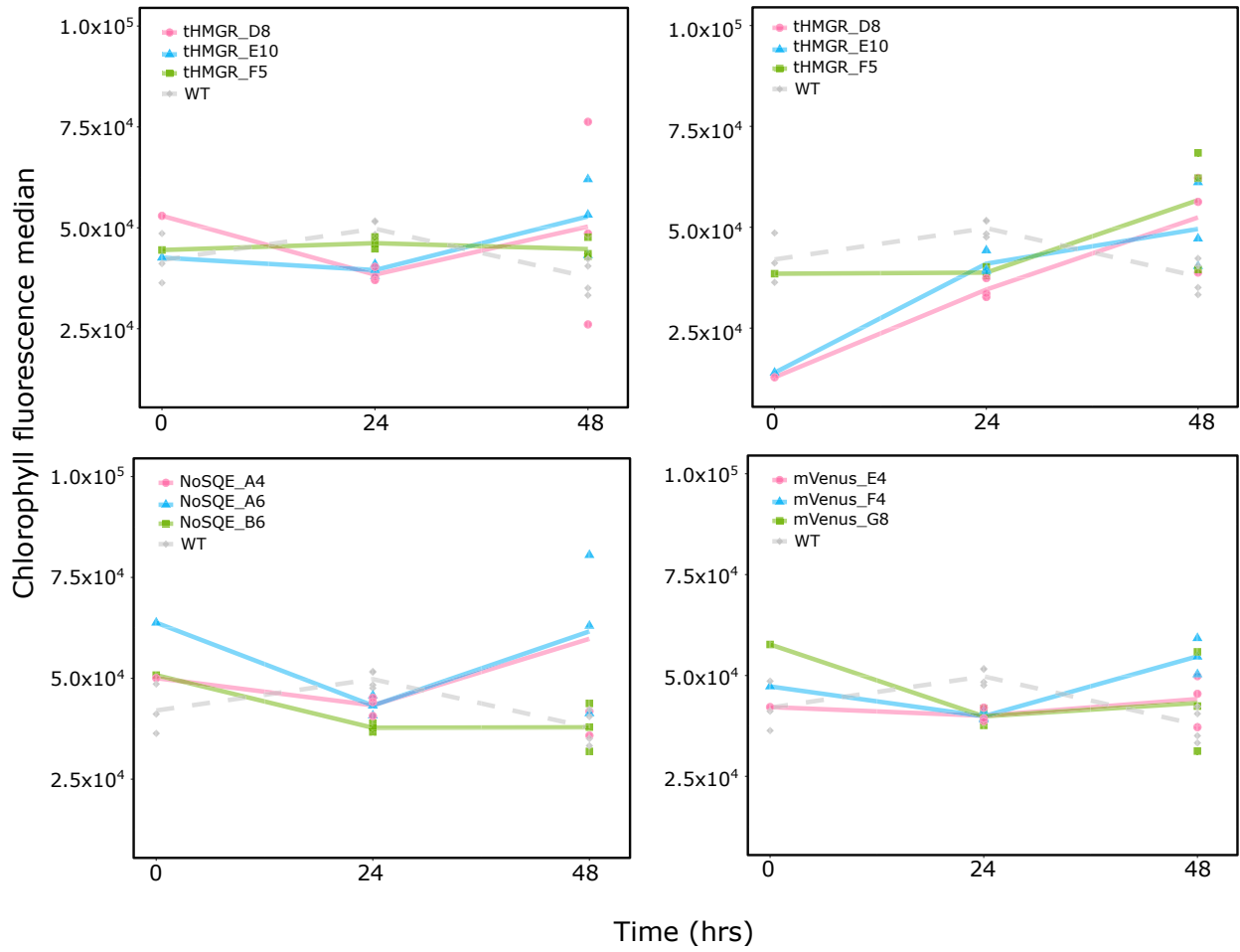
Figure 4.S10: Chlorophyll fluorescence during full scale experiment in *P. tricornutum* transformants (n = 3).
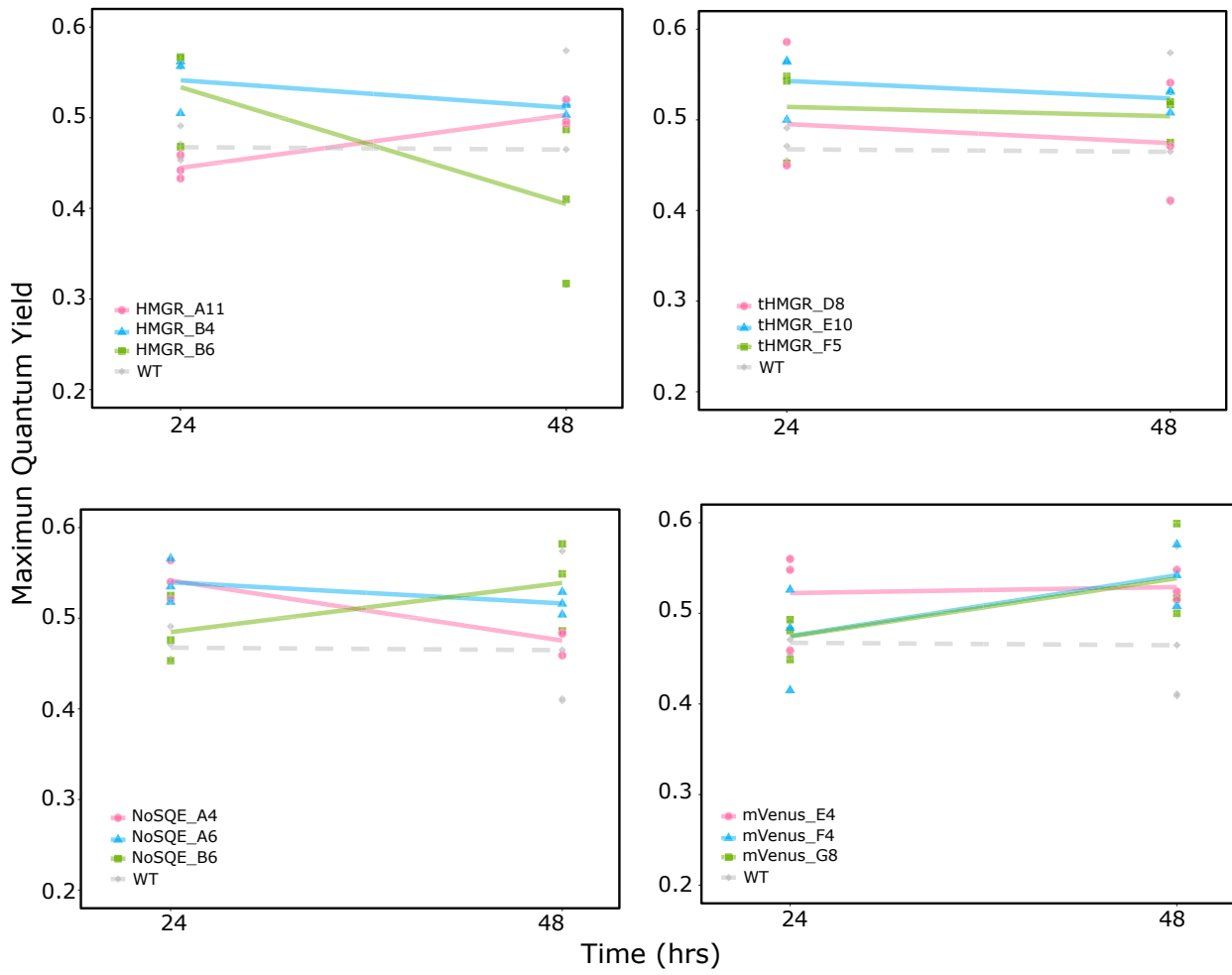
Figure 4.S11: Growth curves during full scale experiment for *P. tricornutum* transformants (n = 3).
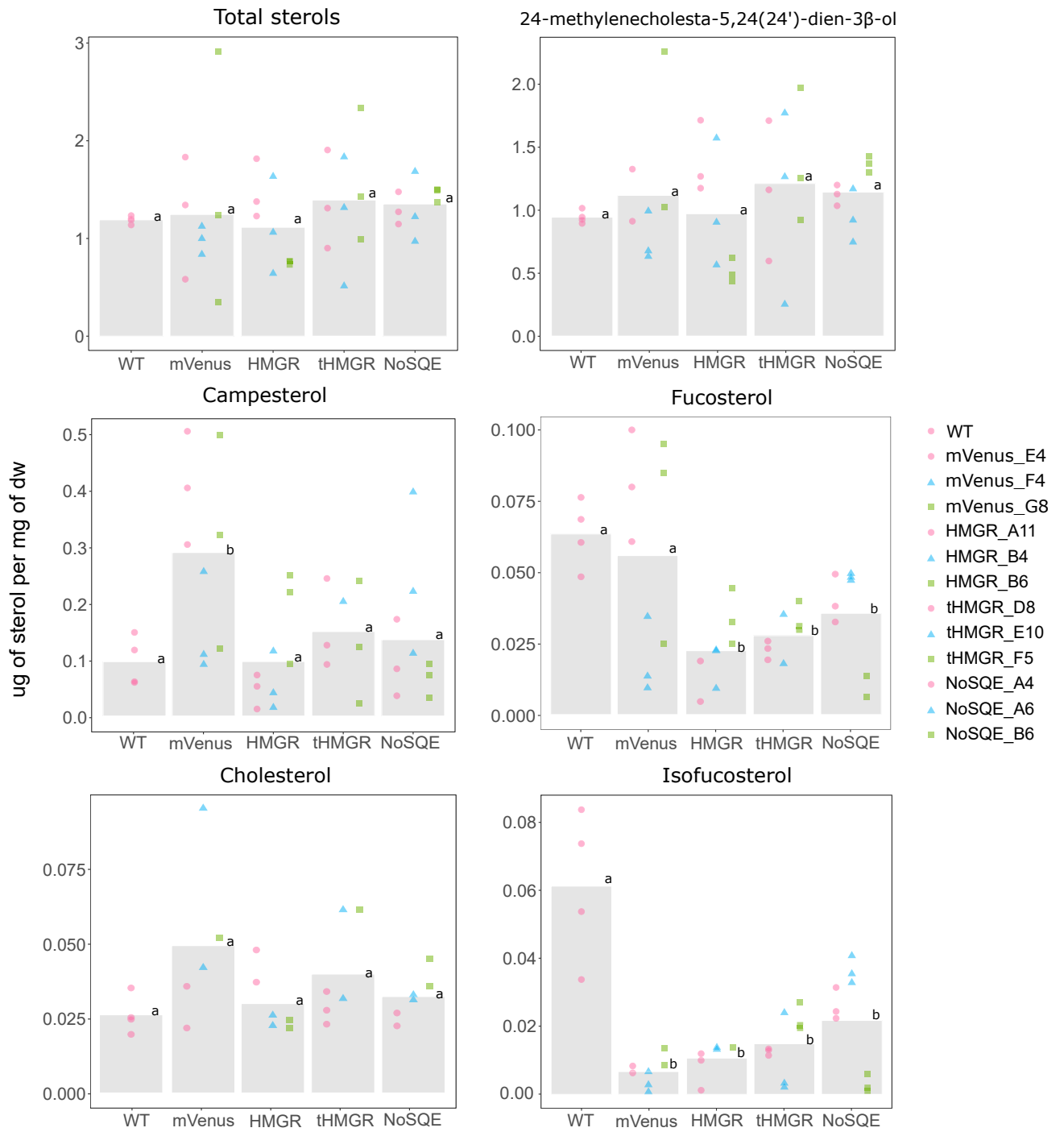
Figure 4.S12: Maximum quantum yield for *P. tricornutum* transformants during full scale experiment (n = 3).

Sequences of genetic parts used to build episomes used in this study

>CEN_ARS_HIS
GCGAGCATCACGTGCTATAAAAATAATTATAATTTAAATTTTTTAATATAAATATATAAATTAAAAATAGAAAGTAAAAAAAGAAA
TTAAAGAAAAAATAGTTTTTGTTTTCCGAAGATGTAAAAGACTCTAGGGGGATCGCCAACAAATACTACCTTTTATCTTGCTCGTC
CTGCTCTCAGGTATTAATGCCGAATTGTTTCATCTTGTCTGTGTAGAAGACCACACACGAAAATCCTGTGATTTTACATTTTACTT
ATCGTTAATCGAATGTATATCTATTTAATCTGCTTTTCTTGTCTAATAAATATATATGTAAAGTACGCTTTTTGTTGAAATTTTTT
AAACCTTTGTTTATTTTTTTTTCTTCATTCCGTAACTCTTCTACCTTCTTTATTTACTTTCTAAAATCCAAATACAAAACATAAAA
ATAAATAAACACAGAGTAAATTCCCAAATTATTCCATCATTAAAAGATACGAGGCGCGTGTAAGTTACAGGCAAGCGATCCTAGTA
CACTCTATATTTTTTTATGCCTCGGTAATGATTTTCATTTTTTTTTTCCACCTAGCGGATGACTCTTTTTTTTTCTTAGCGATTGG
CATTATCACATAATGAATTATACATTATATAAAGTAATGTGATTTCTTCGAAGAATATACTAAAAAATGAGCAGGCAAGATAAACG
AAGGCAAAGATGACAGAGCAGAAAGCCCTAGTAAAGCGTATTACAAATGAAACCAAGATTCAGATTGCGATCTCTTTAAAGGGTGG
TCCCCTAGCGATAGAGCACTCGATCTTCCCAGAAAAAGAGGCAGAAGCAGTAGCAGAACAGGCCACACAATCGCAAGTGATTAACG
TCCACACAGGTATAGGGTTTCTGGACCATATGATACATGCTCTGGCCAAGCATTCCGGCTGGTCGCTAATCGTTGAGTGCATTGGT
GACTTACACATAGACGACCATCACACCACTGAAGACTGCGGGATTGCTCTCGGTCAAGCTTTTAAAGAGGCCCTAGGGGCCGTGCG
TGGAGTAAAAAGGTTTGGATCAGGATTTGCGCCTTTGGATGAGGCACTTTCCAGAGCGGTGGTAGATCTTTCGAACAGGCCGTACG
CAGTTGTCGAACTTGGTTTGCAAAGGGAGAAAGTAGGAGATCTCTCTTGCGAGATGATCCCGCATTTTCTTGAAAGCTTTGCAGAG
GCTAGCAGAATTACCCTCCACGTTGATTGTCTGCGAGGCAAGAATGATCATCACCGTAGTGAGAGTGCGTTCAAGGCTCTTGCGGT
TGCCATAAGAGAAGCCACCTCGCCCAATGGTACCAACGATGTTCCCTCCACCAAAGGTGTTCTTATGTAGTTTTACACAGGAGTCT
GGACTTGACC

> OriT
GATCGTCTTGCCTTGCTCGTCGGTGATGTACTTCACCAGCTCCGCGAAGTCGCcCTTCTTGATGGAGCGCATGGGGACGTGCTTGG
CAATCACGCGCACCCCCCGGCCGTTTTAGCGGCTAAAAAAGTCATGGCTCTGCCCTCGGGCGGACCACGCCCATCATGACCTTGCC
AAGCTCGTCCTGCTTCTCTTCGATCTTCGCCAGCAGGGCGAGGATCGTGGCATCACCGAACCGCGCCGTGCGCGGGTCGTCGGTGA
GCCAGAGTTTCAGCAGGCCGCCCAGGCGGCCCAGGTCGCCATTGATGCGGACCAGCTCGCGGACGTGCTCATAGTCCACGACGCCC
GTGATTTTGTAGCCCTGGCCGACGGCCAGCAGGTAGGCCGACAGGCTCATGCCGGCCGCCGCCGCCTTTTCCTCAATCGCCCTTCG
TTCGTCTGGAAGGCAGTACACCTTGATAGGTGGGCTGCCCTTCCTGGTTGGCTTGGTTTCATCAGCCATCCGCTTGCCCTCATCTG
TTACGCCGGCGGTAGCCGGCCAGCCTCGCAGAGCAGGATTCCCGTTGAGCACCGCCAGGTGCGAATAAGGGACAGTGAAGAAGGAA
CACCCGCTCGCGGGTGGGCCTACTTCACCTATCCTGCCCGGCTGACGCCGTTGGATACACCAAGGAAAGTCTACACGAACCCTTTG
GCAAAATCCTGTATATCGTGCGAAAAAGGATGGATATACCGAAAAAATCGCTATAATGACCGAAGCAGGGTTATGCAGCGGAAGAT

>bsd
ATGGCCAAGCCTTTGTCTCAAGAAGAATCCACCCTCATTGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTCTGAAGACTA
CAGCGTCGCCAGCGCAGCTCTCTCTAGCGACGGCCGCATCTTCACTGGTGTCAATGTATATCATTTTACTGGGGGACCTTGTGCAG
AACTCGTGGTGCTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGCGATCGGAAATGAGAACAGGGGCATC
TTGAGCCCCTGCGGACGGTGCCGACAGGTGCTTCTCGATCTGCATCCTGGGATCAAAGCCATAGTGAAGGACAGTGATGGACAGCC
GACGGCAGTTGGGATTCGTGAATTGCTGCCCTCTGGTTATGTGTGGGAGGGCTC

>nat
ATGACCACTCTTGACGACACGGCTTACCGGTACCGCACCAGTGTCCCGGGGGGACGCCGAGGCCATCGAGGCACTGGATGGGTCCTT
CACCACCGACACCGTCTTCCGCGTCACCGCCACCGGGGACGGCTTCACCCTGCGGGAGGTGCCGGTGGACCCGCCCCTGACCAAGG
TGTTCCCCGACGACGAATCGGACGACGAATCGGACGACGGGGAGGACGGCGACCCGGACTCCCGGACGTTCGTCGCGTACGGGGAC
GACGGCGACCTGGCGGGCTTCGTGGTCGTCTCGTACTCCGGCTGGAACGCCGGCTGACCGTCGAGGACATCGAGGTCGCCCCGGA
GCACCGGGGGCACGGGGTCGGGCGCGCGTTGATGGGGCTCGCGACGGAGTTCGCCCGCGAGCGGGGCGCCGGGCACCTCTGGCTGG
AGGTCACCAACGTCAACGCACCGGCGATCCACGCGTACCGGCGGATGGGGTTCACCCTCTGCGGCCTGGACACCGCCCTGTACGAC
GGCACCGCCTCGGACGGCGAGCAGGCGCTCTACATGAGCATGCCCTGCCCC

>mVenus
ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAG
CGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCT
GGCCCACCCTCGTGACCACCCTGGGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAG
TCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT
CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGT
ACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATC
GAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTA
CCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGA
TCACTCTCGGCATGGACGAGCTG

\>Thrombin_6XHIS_FLAG(DYKDDDDK)
GCGCTGGTCCCTCGCGGTAGCCACCACCACCATCACCACGACTACAAGGATGATGACGATAAGTGA

\>NoSQE
ATGGAGGCATCAACACCAACACAAGGACAAGTGCCACATAGACCATCATTGTCATCAGAGCCAACACCACCAGGATCACCAAACTC
ATCAGTGGGAGTGGGAGGATCAATTGTGGTGGCATCATCATTGTTGTTGGTGTCATCATTGACATTGTGGCAAGCATGGCATCAAC
ATCAACAAACACCATTGGCAACAACAGGAAACGCAGTGAACACAGCAAAGTTGTTGTTGCCAGTGTTGGGAGTGTTGGCATCATTG
TCATTTTTGTGTTACACATCATTGTTTTCATCAGCACAATCACAATCATCATGTGATGAGCCAGATACAAAGATTACATTGGGACA
AAAGTCAACAGATTCATCATTTTCACCATCAACAACAACAACATCATCACAATTGCCATCATCATCAATTTCAAACGCAGAGATGT
TGTTTAGATCAGGAGGAGTGGGAGGAAACAGAATTCCACCAAGACAAGCACCATCAGCAGGACAAACAATTATTGTGGTGGGAGCA
GGAACAGCAGGAGCAGCAGCAGCAGTGGGATTGGCAGAGAGAGGATTTAAGGTGTTGTTGTTTGAGAGAGATATGACATTGCAAGA
TAGAATTGTGGGAGAGTTGTTGCAACCAGGAGGAGTGAGAGCATTGGAGAGATTGGGATTGGATGAGTGTGCAAAGGAGGGAATTG
ATTCAGTGGGAGTGGAGGGATACGTGGTGTTTAGACCACATAAGTTGGAGGTGGAGGGAGGAGAGGAGGAGGTGGTGAAGAACATT
ATGTTGAACTACCCAGCATACGATCCAGCATCAATGAAGGAGTTGTTTGGAGTGATTGATCCATGTGTGTCACCATCATCATCATC
ATCATCATCATCATCATTGTTTTCAATTCCACCATCAACAGAGGAGGTGCATGGAGCAGGAAAGACATTGTCACCACATTCAG
CAACAAAGGCAAGAAAGACAGCAGCAGCAGTGGCAACAGCAGCAACAACAGTGGCACAAGAGCATGAGCCATTGCGTCCAAAGGGA
AGATCATTTCATAACGGAAGATTTGTGCAAAGATTGAGAGAGAGAGCATTGGCAGAGAAGAACGTGACAGTGATTCAAGGAATTGT
GTCAAGATTGGTGCATGAGGAGGATTTGACAGAGGGAGCAAAGGAGGGAAGAGAGAAGGGAAAGAAGCAAAGAGTGGTGGGAGTGG
AGTACAAGATTGAGACAGGAAGAACATTGACAAACGAGGAGGGAGTGGAGGTGCCAGAGAAGGAGACAAGACGTGCATTGGCACCA
TTGACAATTGTGGCAGATGGATTGTGGTCATCATTGAGAACACATGTGGATAAGTGTGCAAAGCCACATCAATTGTCATCATTTGT
GGGATTGATTATTACACATCCATTTATGTCATCACCATTGCCACATGCACATTGTGGACATGTGGTGTTGGCAGAGCCATCACCAG
TGTTGTTGTACCAAATTTCATCAACAGAGACAAGAGTGTTGGTGGATATTCATGGAAAGTTGCCATCAGCATCAACAGGAGAGTTG
GCATCACATTTGTACACACATATTTTGCCACAATTGCCACCATGTTTGCAAGCACCATTTGCAGAGGCAGTGGCAAGAGGAAACTT
TAAGTCAATGCCAAACAGAGCAATGCCAGCAGCAAAGTTGGCAAAGGATGGAGTGATTTTGTTGGGAGATGCATTGAACATGAGAC
ATCCATTGACAGGAGGAGGAATGACAGTGGCATTGAGAGATGTGGAGGCATTGTGTTACTTGTTGGAGGAGCCAACAGCAGATTTG
ACAAACCCAGTGGCATTTAGAGCATTGAAGAAGAGATTTGCAACAGAGAGAAAGTCATGGGCAGCAACAATTAACGTGTTGGCAAA
CGCATTGCATGCAGTGTTTTCAACACCAGATGGAGATTTGACAAGAAAGGATTTGAGAGATGCATGTTTTGATTACTTGTCAATTG
GAGGATACTGTGCAGGAGGACCAATTAACTTGTTGTCAGGATTGACACCAAGACCATCAATTTTTGGCAATGCATTTTTTTTTGGTG
GCATTTCATGGAGTGGGAAAGTACGTGTGGCCATTTCCAACATTGGGAGGAATGGTGAGAATGTACGATATGTTGCATGTGGCATG
TCAAATTATTATGCCATTGTTGACAGTGGAGGGATCAACAGTGTTGGCATCATGGCCAGTGAGAACAGCAATTGGAGTGGTGTTTC
CATACGGAAAGTCAGGT

\>Pt_EF2_promoter
TCCATTTTGACATGTTTCCTAGCTAGAAGGAACAGCTATTCATATGAATTTTCCGAGGTCACGCTGTGATGGATCGATTCGTTCCC
GTCACAGACGCAAGCAGCGAGGTGATGGCTGGACGGGCACGCGTCACAATTTAGAACAACCCGGACTCTTGTGACGCGAGCTCCGG
GTGCGGAAAACCCTACGGGGTGAAAGGAGCCGCAGCTTCCACCCATTTACTTGACTGATAACCCAGTACACACAACGCTACCGTGA
CGGAAATGCCGCGCGCTGTGAAGCCGTGGTGAATCTTTCGCTGCTGCTCGTTCTCTCCACACACAGAATTGCACGGCTTGACGATT
AGCAAAAAGCAGAAGAAGCAGTGTCAAGATGGTGAGTGTCTTGAGTACAAAAGTAGCGCACGGTAGTGCAGAATACGGCTGTGTAG
ACTATGCGGCGTCGGAGTAGCCGAGATCCCGTAAAAAACGGTAGGAATCGGAGGAATCTCAATGAGTGATCGTGCTGCGTACGCCG
CTTCCGTTGGGGACCATACGATCGCGTAGTTCTTGTTAGCAACGACATGCCTTTACGTCGCGAAGGTCTTTCAAGTAACTAAAACA
ACGGAATATGTCAAGATTACCTCGATTTGATGGATTCGAATGAGGATCGGGATCAGGAGTTCTATTGACACTAGTTGATTTCCTCA
ATGCGAATGCCGCTCAATGTCCCCCTTTCTCATCTATGACTGTTTACTTCCCTTAGGTGAACTTTACAGTTGAACAG

\>Pt_EF2_terminator
ACAGAAAAACAGACTCATAGGGTACTTCGAACATAAGAAAAGATTCATTACATGACGGATTTCAGTTTCGTATTCGTATAGGTTTA
TTGAGTTGATGAACATGGCGGTCAGATATGCCAATCGCTTTTCTTGGTACGAGTTTGCTAGG

\>Pt_49202_promoter
TCCAAACGGTACGTCAGATCCCACGTCACTACCGCACCCGGTTATTCCAGTGCGTAAGCTCGTAGTGGAAGAACAGGCTGCGGTAG
TGACGTGGAATTTGACACTTTCCAATCTTCAAGACCTCTAGCCAGCAAAGACTTCTTGAATCTGCAAGGATATTCTATTTAGATAT
AAGTATATACTTGACAGTGAGGGGGGTCGTGACCGTCTCTCGCTGGGATAGAGAAGATGGCAACCATTGAATTGAGAACTCGAAAC
TTGAGCTTGGTTGGATCTTGATCATTCTGAATTTGCAGGATAGGTTAACAGTAAAAGTCAATCTTGCGCTCGCCATCTACCAGAAA
GCACGTGGGTGCCCAGATCCTCCCAACGAGCGAACGGACGGATGCTAGTCGGGGGGATCTCTGTGGGCCATTCAAACGAAAGGCAG
ATGGCATCGTTCCAAGATCTATTGGTCATTGGGCTTCTCATACAGTCATTGACACCACGATTGAAGACAAATTTTTGTTGACAGTG
AAATCGTTGCGGACGCGCTCCGATTCAAAAGTTTCCCTACTCGGTCGCGTCAACCGACCTTCCCCTCGTGAGTTTACAGCTTGATA
CGCCCTATTTGGCTTCCTATGTAAAGTCTCCAGTGATCATATTAAAGTCATTGGTCATGACACTCTCCTTGTTCAAAATTTGGTAC
CATACAATGCGCGGAGCAGTCAGTTCTCATTTCCAGCGATCGTTAAAACAATTAGAGAAATGCCTTACTTCTCTCTTTATATCTGT
ATTACTCAGAGTCAGTGAGTAGTACGAATTTCCAACGTCCTAGCTCCAATTTATGGTTTCAAACAATCCGAATTCGTTGGTCCACA
GTCAAGGAATTCGCAAGAATCCCTATTTCACTAGGCGTGCACTATGGGGTCGCTTTTGAGGTTCTCATTTGCGGCGATCGTTAAAA
CGATTACAGAAATGCTACACTTCGCTCTTTATATATATTACTCATTGAGCATCAGGAATGTCCAATGTCGTAACTCCGATTCATGG
TTTCAAACAATCCGAATTCGTTGGTCCGAGGAATCCGCAAAAAGTAATTGTAGTCTATGTACGACCGAGTCGCTGGTGAGGGTTTT
TTGTCGTTGGATTGCTCGAGCGAATTCTCCGAGATTGTACAGAAAGCTCAACTTTCGGAAGGCATCTATGTGGTCCAACCAATCTA

TTCAAATATATCAATTGTTAAGAGCAGAGTTCGATCCTCACTATGTGGAGGGGCATGGTGAGATTACGGAGTGCCGAGACATAAGA
GCTCCCTGCAATCAAGTTGATGGACATGCTTCGAAAATGCCATGACTGAGAACGAGAACGCCCAGATGCTCCGCGGCACGAGAACT
CGATGAACACGAACCGAATCCAACGCCAAAAACGCGTTGTCCAAAACATTGCAACTCTCTCTGATAAAAACCAAAACGCCGCAAAT
TGCACCAGCTTTTTTGGTTGTCCTACGCCTCTTTCCGCACCACC

>Pt_49202_terminator
TTCGATCGACGAGCTTACCTTCACGGCCACCGTTCCCGACCCCGACGTTGTTCAGGCGATCATTACGGGTGGCAGCGCCGCTCCGG
GAGCCACCGGTTTGTACGGCATTGCCACCGCAGATGCCGAGTGCCCGGCCAACGAGTGCATCCATGTGGACGTGATTGTGTACGCC
ATTTTCGCGAATGAAGCCACGACCAACTTGATTGTGGAAATCAGCGGCGACGCGGTCATGTCTGTGGGACGTACTGCCCGCAAGTT
GCGTGCGAAAGTCCAGCCCACACGCGAGTTGCAGGAGACGCTCCAGATGCAAGCCTTTCGTAACGAGATTGTCCTCCCCCCTTTGC
CTGGTTCGGGGGGCCACGTTGGCCAACGTGTGCGCGGCGGCGGTGCCCTTCGCCGTGGCCGTGGTGGCCGGTGCGCTCGCTCTTTAA
AAAGTGAGAGTTTGTGTGTGTGAAAGACGCGTCGCGAGACGCTCCTTATTTCCGTCGGCATGTGTCTTGTCGTGGTCCTATGTGAT
CCCCGGGTCGGTTCACGGGAGTTGCCGAAGTAGAGAATAAGAATTATATAACTGTACAGAGATACAGACATGTGCTGCTATTACGT
CCAACAGTAGGTATACACCTCCCGGGCAATAGAGTGGGCCACATCCATACACACACACCCGAGTTGGTAGATTGATTTCGGTTCGT
CGAATGCCAAGAATGTGCTTCCTGAAGAAAATTTGCTCCTTCCCACTCCGTCCGTACGACCATCCGCCCGTGTACAGCGCGGCAGA
AGTAGAAACGAGTACCAAGGCTCTGTGTGTGTGTGTGTACCGTTTGGACGCGTCGGCAAAGGTACGGCAGTCCTCCGGGTGTGTCC
CGTACTCGTGGTACCAACGAATCTGCACAATTGTGCCGTTCGGAGGACACGCCTTCCCGGATGCATGCATCCGGAAGACCGTCGAT
GCGAT

>Pt_HMGR
ATGACGGTGACTATCAGCAGTAGTATTAGTAGTAGTATTGGTAGTAGTACGACTGCCCCAACACTGGGAATGCAGGTGGATGCCCT
TGTCCAACAAATGGATAACTTGAGCAGTACGCAGTTGTACGGATTGATTGTTGGTTTGACCATTCTCGTCTCCTTTGTGTTGCTCG
GGTCTTCCGCGGACATTCCCGTTTCTCTTCGAGACACTACCAAAGAGACGTCGTCCTCCTCCTCCACACCCCGCAAGACAACCACG
GTGTCGTCGTCGTCGAATCGGGGACCGGAGCCGCGCTGGCACATCTTTACTTACGTCAATTACGCCATTGTGGCGTGCTTTGTCGC
CTCCGTTGCAGAATTTGGACGCAACGCTTCGGCGTACCTTGCTGCCGACGACAACGTCGTGCTATACTTTCTCGTCGCTTGGAGCG
TCTTTTTTGTGCTACTTTTTTGGTTTCTTTGGTGTCTCCTTTGTCCACGACGCGGACGCTGCGGTCGCGAGTCCCACACCAACCAAG
CCCACGGTCCGGTACGTCCCTGCCTTGATCCATTCGGCCTTGTCTCGAGATATACATGTGTGTGCGTGTGTATACTTGTGCTATGC
GTACTTGTGGTGTGTGCACATGCGTATACACCAGTACTTATACGTACCCGACCGAGTTTCCTTCGTTCGTTCCACCTCGTCCCGGC
CCACTTTTCGTGTTCCGGTGACTCCGTCTCACCCTCCTTGCTTATTGGTTCGCAGCGACTCTATTTCCACCAAGACCAGCGCGTCC
AACACTCTTTCCGCTCGGCATCCTCCCGCCCCGTCGGCACCCGTATGCAGTGATCCCTCCAGTTTCACACCAATCCATACCAGCAA
GACCAACATTACGACACTAGATAACGCTGCCATCTGTCAACTGGTGCTCACCAATCAGATCAAGGACCACGAACTCGAAAAACGTC
TCGACGCGCACCGCGCCGTGCAGGTCCGCCGTTTAGTCGTCGCCCACAAACTCGACACCTTGGAACACATCAACGCGCACGCTCTC
GACAATCTGCCCTCGGAACCTTCCCTCGACTACACCCGCGTACACGGCGCCAATTGCGAAATCGTTGTGGGCTACGTACCCCTCCC
GGTGGGACTCGTCGGTCCCCTCACGGTCAATGGCGAAACCGTATACGTGCCCATGGCCACCACCGAAGGATGTCTCGTGGCCTCCA
CGAACCGCGGTGCCAAGGCCATCACGGCCGGGGGTGGAGCGACCGCCGTCCTCCTCCGGGACGGTATTACTCGGGCGCCCTGCGTG
CGCATGCCCTCGGCAGCGCAAGCCGCTCATCTCAAATTGTGGTGTGAAACGCCCCAACACTTTAGTACCCTTAAACGAGCCTTTGA
AAGTACCACCTCGTTCGGGAAACTTCTCAAGTGCGCCCCCACAGTCGCTGGACGTAACGTGTATTTGCGCTTGACCTGCTTTTCTG
GAGACGCCATGGGCATGAACATGGTTTCCAAGGGCAGTTTGGCCGTCATTGAAACCCTTCAACAAGAATTTCCGGAACTCGTCCTC
GTCGCCTTGTCTGGCAACATGTGCACCGATAAGAAAGCCGCCGCTACCAACTGGCTGGAAGGACGAGGAAAATCGATCGTCGTGGA
AGCGACCATTCCCAAGGACGTCGTCACAAACACCCTCAAGACTACGGTCCACAGCATGGTGCAAACCAATTTGCACAAGAATTTGA
TCGGTTCGGCCATGGCCGGAGCTTTGGGCGGATTTAACGCCCACGCCTCCAACATTGTTACCGCGGTCTTTTTGGCCACGGGACAG
GACCCCGCGCAAAACGTGGAGTCCAGTAATTGCATTACCCTGTTGGAAGAAACGGAAGAAGGCGACTTGTGGATTTCCTGCACCAT
GCCCAGCATTGAGGTCGGTACGGTCGGTGGCGGAACTTCCTTACCCGCCCAAGCGGCCTGTTTGCAGGCCATGGGTGTCCGGGGCG
GAGGCGCCACTCCCGGCGCCCACGCCCAAAAACTGGCACAAATCGTCGCGTCCGCTACTTTGGCCGGCGAACTTTCGCTGCTGGCC
GCACTCGCCGCGAATACCCTCGTACAGGCCCACATGCAACACAACCGCAAACCCGCCGCCAAG

>Pt_tHMGR
ATGGACTCTATTTCCACCAAGACCAGCGCGTCCAACACTCTTTCCGCTCGGCATCCTCCCGCCCCGTCGGCACCCGTATGCAGTGA
TCCCTCCAGTTTCACACCAATCCATACCAGCAAGACCAACATTACGACACTAGATAACGCTGCCATCTGTCAACTGGTGCTCACCA
ATCAGATCAAGGACCACGAACTCGAAAAACGTCTCGACGCGCACCGCGCCGTGCAGGTCCGCCGTTTAGTCGTCGCCCACAAACTC
GACACCTTGGAACACATCAACGCGCACGCTCTCGACAATCTGCCCTCGGAACCTTCCCTCGACTACACCCGCGTACACGGCGCCAA
TTGCGAAATCGTTGTGGGCTACGTACCCCTCCCGGTGGGACTCGTCGGTCCCCTCACGGTCAATGGCGAAACCGTATACGTGCCCA
TGGCCACCACCGAAGGATGTCTCGTGGCCTCCACGAACCGCGGTGCCAAGGCCATCACGGCCGGGGGTGGAGCGACCGCCGTCCTC
CTCCGGGACGGTATTACTCGGGCGCCCTGCGTGCGCATGCCCTCGGCAGCGCAAGCCGCTCATCTCAAATTGTGGTGTGAAACGCC
CCAACACTTTAGTACCCTTAAACGAGCCTTTGAAAGTACCACCTCGTTCGGGAAACTTCTCAAGTGCGCCCCCACAGTCGCTGGAC
GTAACGTGTATTTGCGCTTGACCTGCTTTTCTGGAGACGCCATGGGCATGAACATGGTTTCCAAGGGCAGTTTGGCCGTCATTGAA
ACCCTTCAACAAGAATTTCCGGAACTCGTCCTCGTCGCCTTGTCTGGCAACATGTGCACCGATAAGAAAGCCGCCGCTACCAACTG
GCTGGAAGGACGAGGAAAATCGATCGTCGTGGAAGCGACCATTCCCAAGGACGTCGTCACAAACACCCTCAAGACTACGGTCCACA
GCATGGTGCAAACCAATTTGCACAAGAATTTGATCGGTTCGGCCATGGCCGGAGCTTTGGGCGGATTTAACGCCCACGCCTCCAAC
ATTGTTACCGCGGTCTTTTTGGCCACGGGACAGGACCCCGCGCAAAACGTGGAGTCCAGTAATTGCATTACCCTGTTGGAAGAAAC
GGAAGAAGGCGACTTGTGGATTTCCTGCACCATGCCCAGCATTGAGGTCGGTACGGTCGGTGGCGGAACTTCCTTACCCGCCCAAG
CGGCCTGTTTGCAGGCCATGGGTGTCCGGGGCGGAGGCGCCACTCCCGGCGCCCACGCCCAAAAACTGGCACAAATCGTCGCGTCC

GCTACTTTGGCCGGCGAACTTTCGCTGCTGGCCGCACTCGCCGCGAATACCCTCGTACAGGCCCACATGCAACACAACCGCAAACC
CGCCGCCAAGTCAGGT

>Tp_EF2_promoter
AGTGTGCAATGCAGTCAATTCAATAGATATGTAAAATGATGATGTAAGGGAAACAATCAAGAACAGATTTCAGATCAAACTGAATG
AACTCTTGCACATTGACACTTGACAAACACATAGTAGTGTCCCGACAATCCCAATTGAAAAACATAGTAAGCGATAGACTCTCTTT
CACTCTGCTTTTCTCTGCTCTCTTTTTTTGACGGTTGACAGACATACTTCCACGTGGCCGACGTAACTCGTTGAGCGGTTTGTCAC
TATATGTTCATGTCTCGAAAAGATGACGGGACACCAAGATTCCTGCTTGCTTCTACCCGTTGACGCCGCACGCACCAACAAAGACA
TCAACGCCCGCCTCCACCCAGCCACTTTACAAAGTCCTTCTTTTCAACACGCTAGTTAGCACCCCCGCGATTAATTAAAGAGAAAA
GAACGTCAAG

>Tp_EF2_terminator
ATATCTTCTTGCAACAATGGTAGCGCGATTGATGAGATAATAGATAAAAAAATCCACTGTGAATGTTTTTCAATGAGTCTACAATA
GTGACTTGACAGGTCGTATAGAGTGAGACATTAGTTATCTCTAACCAACCCTGCT

>Tp_HMGR
ATGGCGGCAGCAGCAGCACCAACCATAGGCCAACGCCTAGACACACTCCTCGCCGCCCTCTCCAACATTGGTGACATGAACCACCT
CCTCTCAAACTATCAGCCATCCACCTCGCAGATCTACTCCTTGGTCATCGTCCTCAGTGTGGCATTTTCATTCCGATTACTAAACA
GCGGTGATGATTTGGGCAGCAAATTGTCGTCGTCCACTGTGGAACAGGCGAAGCGGGGTGAGAAGGGAACTTCGTCGAAGTACGCG
AAGGCCAACAACAAAAACAACAAACAATCCTACGATTTCCCACAGCCCAAATGGCACATTCTCAAACTAACCAACTACGTCGTCAC
CACCCTCTTTCTCCTCTCCATAATCACCTTCCTCTCAAATGCTAGCGTGTATCTCAATGACAACACGGCATTAATGACTTTTTTGG
GTGTTTGGTCTGGATTGTTGTGTTATTTCTTTGGATTTTTTTGGCATTTCGTTTGTGGAGTTGGATGATCTTGTGGTGACTGATGCT
GATGGGCAACAACGGCAGCAGGGAGTTGTTACTCAGCCGCAGAAGCAGGCTAGTTGCAAATCTAGAAAGGTGGAGGTGAGTGTACT
GATATCATGTTTGATGGCACATGAGATAATGTGCTACTCAACCATTCTTGACTAAACAATTCTTATCCATCCATCCATCCAAAGCA
GCCCCAACGACCCACCCGTCAAAGTCATCAGTCTCCATCCACCAGCATCCTCTACACCCGTCTGTTCGGATCCCATCAAAAGCAAG
TCTCCCACCAACACAGGAATCCCCACAAACGTCAAAGATTTATCCAACGAAGAAATTGCAACCCTCGTCCTCCAAGACAAGATTAA
AGACCACCAGTTGGAGAAATTATTAGATCCTCATCGTGCAGTTGCCGTAAGGAGACTCAAGTTTGATGCCCTCCTATCCTCACTTG
GCAAAACTACTGGAGAGGATGATAAGAAGGGGGGAGTGCTATCCGAACTACCCCACGAACACGATTTAGACTACAAACGAGTCCTC
GGTGCGAATTGCGAGATTGTAGTGGGGTATATTCCTATTCCTGTGGGAATTGTGGGACCGGTGACGTTGAATGGGGAGAGTGTGTA
TATTCCCATGGCGACTACTGAGGGGTTGGTGAGTGAATGTTGCTTGCATTTGATTTATTTGTTTACATTTGGTTGAGGAACAATGCT
CATTGCACCTACATCTCAATCTTTCATTGCTAACTTCTGCCACCATTCTTTCCTCCACACCACAGTCTCGTAGCAAGCACCAACCG
TGGATGCAAAGCCATCACCCAAGGATCCGGGGCCCGCAGCACCATCCTCCGCGACGGCATCACCCGCGCCCCCTGCGTTCGCCTCC
CCTCCGCACACGAAGCAGCTCAAGTTCACCTTTGGATTGAAGAGGCGGACAACTTTGCAAAGCTCAAAGAGGCGTTTGAGTCTACC
ACCAGTTTTGGCAAGTTGATTGAAGCTAGTCCGACTGTGGCCGGGAGAAATGTGTACATTCGCCTCCGTTGTTTCTCGGGAGATGC
AATGGGAATGAACATGATTTCCAAAGGAAGTCTCGCCGTCATTGAATGTCTTCGGGAGCAATTCCCCCAACTGTCTCTCGTTGCGT
TGAGTGGCAACATGTGCACGGATAAGAAGGCGGCCGCGATGAATTGGATCGAGGGACGTGGAAAAAGTGTCGTCATTGAAGCTACC
ATTCCCAAAGATGTGGTGCGATCTACTCTCAAAACGTCGGTCCCTGCCATTGTGGAAGCAAATGTGAATAAGAACTTGATTGGAAG
TGCCATGGCCGGGACAGTTGGAGGATTTAACGCCCATGCGGCCAACAATGTCACTGCGGTGTTCCTCGCTACGGGTCAGGATCCTG
CCCAGAATGTAGAGAGTTCCAATTGCATCACTCTTATGGAAGTGTCACCTGAGGGAGACTTGTGGATATCATGCACCATGCCCAGT
ATCGAGGTTGGAACAGTCGGTGGAGGAACTGGACTTTCTGCTCAGTCGGCGTGTCTTCGTGCCATTGGTGTCAAGGGAGGGGGAGA
GAATCCTGGAGATAACGCGAGGCAGTTGGCTCATGTGGTGGCTTGTGCTACCATGGCAGGGGAGTTGAGTTTGATGGCGGCATTGG
CGTCTAACTCGTTGGTTGCGGCTCATATGCAACATAATCGCAAGCCTGCTTCAAAG

>Tp_tHMGR
ATGAGCCCCAACGACCCACCCGTCAAAGTCATCAGTCTCCATCCACCAGCATCCTCTACACCCGTCTGTTCGGATCCCATCAAAAG
CAAGTCTCCCACCAACACAGGAATCCCCACAAACGTCAAAGATTTATCCAACGAAGAAATTGCAACCCTCGTCCTCCAAGACAAGA
TTAAAGACCACCAGTTGGAGAAATTATTAGATCCTCATCGTGCAGTTGCCGTAAGGAGACTCAAGTTTGATGCCCTCCTATCCTCA
CTTGGCAAAACTACTGGAGAGGATGATAAGAAGGGGGGAGTGCTATCCGAACTACCCCACGAACACGATTTAGACTACAAACGAGT
CCTCGGTGCGAATTGCGAGATTGTAGTGGGGTATATTCCTATTCCTGTGGGAATTGTGGGACCGGTGACGTTGAATGGGGAGAGTG
TGTATATTCCCATGGCGACTACTGAGGGGTTGGTGAGTGAATGTTGCTTGCATTTGATTTATTTGTTTACATTTGGTTGAGGAACAA
TGCTCATTGCACCTACATCTCAATCTTTCATTGCTAACTTCTGCCACCATTCTTTCCTCCACACCACAGTCTCGTAGCAAGCACCA
ACCGTGGATGCAAAGCCATCACCCAAGGATCCGGGGCCCGCAGCACCATCCTCCGCGACGGCATCACCCGCGCCCCCTGCGTTCGC
CTCCCCTCCGCACACGAAGCAGCTCAAGTTCACCTTTGGATTGAAGAGGCGGACAACTTTGCAAAGCTCAAAGAGGCGTTTGAGTC
TACCACCAGTTTTGGCAAGTTGATTGAAGCTAGTCCGACTGTGGCCGGGAGAAATGTGTACATTCGCCTCCGTTGTTTCTCGGGAG
ATGCAATGGGAATGAACATGATTTCCAAAGGAAGTCTCGCCGTCATTGAATGTCTTCGGGAGCAATTCCCCCAACTGTCTCTCGTT
GCGTTGAGTGGCAACATGTGCACGGATAAGAAGGCGGCCGCGATGAATTGGATCGAGGGACGTGGAAAAAGTGTCGTCATTGAAGC
TACCATTCCCAAAGATGTGGTGCGATCTACTCTCAAAACGTCGGTCCCTGCCATTGTGGAAGCAAATGTGAATAAGAACTTGATTG
GAAGTGCCATGGCCGGGACAGTTGGAGGATTTAACGCCCATGCGGCCAACAATGTCACTGCGGTGTTCCTCGCTACGGGTCAGGAT
CCTGCCCAGAATGTAGAGAGTTCCAATTGCATCACTCTTATGGAAGTGTCACCTGAGGGAGACTTGTGGATATCATGCACCATGCC
CAGTATCGAGGTTGGAACAGTCGGTGGAGGAACTGGACTTTCTGCTCAGTCGGCGTGTCTTCGTGCCATTGGTGTCAAGGGAGGGG
GAGAGAATCCTGGAGATAACGCGAGGCAGTTGGCTCATGTGGTGGCTTGTGCTACCATGGCAGGGGAGTTGAGTTTGATGGCGGCA

TTGGCGTCTAACTCGTTGGTTGCGGCTCATATGCAACATAATCGCAAGCCTGCTTCAAAG

HMGR sequences for phylogenetic analysis

>Fragilariopsis_cylindrus
```
MGMNMVSKGSLAVIETLQKEFPSLELVALSGNFCTDKKAAATNWLMGRGKSVVVEAIIPKDVVQKTLKTTVKALVHTNINKNLIGS
AMAGAIGGFNAHASNIVTAIFLATGQDPAQNVESSNCITLMEETDDEDLWICCTMPSIEVGTVGGGTSLPAQAACLEAIGCRGGAS
TPGDNARKLATVVAAATLAGELSLLAALAANTLVQAHMVHNRKPATKN
```

>Thalassiosira_pseudonana
```
MAAAAAPTIGQRLDTLLAALSNIGDMNHLLSNYQPSTSQIYSLVIVLSVAFSFRLLNSGDDLGSKLSSSTVEQAKRGEKGTSSKYA
KANNKNNKQSYDFPQPKWHILKLTNYVVTTLFLLSIITFLSNASVYLNDNTALMTFLGVWSGLLCYFFGFFGISFVELDDLVVTDA
DGQQRQQGVVTQPQKQASCKSRKVESPNDPPVKVISLHPPASSTPVCSDPIKSKSPTNTGIPTNVKDLSNEEIATLVLQDKIKDHQ
LEKLLDPHRAVAVRRLKFDALLSSLGKTTGEDDKKGGVLSELPHEHDLDYKRVLGANCEIVVGYIPIPVGIVGPVTLNGESVYIPM
ATTEGLVASTNRGCKAITQGSGARSTILRDGITRAPCVRLPSAHEAAQVHLWIEEADNFAKLKEAFESTTSFGKLIEASPTVAGRN
VYIRLRCFSGDAMGMNMISKGSLAVIECLREQFPQLSLVALSGNMCTDKKAAAMNWIEGRGKSVVIEATIPKDVVRSTLKTSVPAI
VEANVNKNLIGSAMAGTVGGFNAHAANNVTAVFLATGQDPAQNVESSNCITLMEVSPEGDLWISCTMPSIEVGTVGGGTGLSAQSA
CLRAIGVKGGGENPGDNARQLAHVVACATMAGELSLMAALASNSLVAAHMQHNRKPASK
```

>Thalassiosira_pseudonana_tHMG
```
SPNDPPVKVISLHPPASSTPVCSDPIKSKSPTNTGIPTNVKDLSNEEIATLVLQDKIKDHQLEKLLDPHRAVAVRRLKFDALLSSL
GKTTGEDDKKGGVLSELPHEHDLDYKRVLGANCEIVVGYIPIPVGIVGPVTLNGESVYIPMATTEGLVASTNRGCKAITQGSGARS
TILRDGITRAPCVRLPSAHEAAQVHLWIEEADNFAKLKEAFESTTSFGKLIEASPTVAGRNVYIRLRCFSGDAMGMNMISKGSLAV
IECLREQFPQLSLVALSGNMCTDKKAAAMNWIEGRGKSVVIEATIPKDVVRSTLKTSVPAIVEANVNKNLIGSAMAGTVGGFNAHA
ANNVTAVFLATGQDPAQNVESSNCITLMEVSPEGDLWISCTMPSIEVGTVGGGTGLSAQSACLRAIGVKGGGENPGDNARQLAHVV
ACATMAGELSLMAALASNSLVAAHMQHNRKPASK
```

>Phaeodactylum_tricornutum
```
MTVTISSSISSSIGSSTTAPTLGMQVDALVQQMDNLSSTQLYGLIVGLTILVSFVLLGSSADIPVSLRDTTKETSSSSSTPRKTTT
VSSSSNRGPEPRWHIFTYVNYAIVACFVASVAEFGRNASAYLAADDNVVLYFLVAWSVFLCYFFGFFGVSFVHDADAAVASPTPTK
PTVRDSISTKTSASNTLSARHPPAPSAPVCSDPSSFTPIHTSKTNITTLDNAAICQLVLTNQIKDHELEKRLDAHRAVQVRRLVVA
HKLDTLEHINAHALDNLPSEPSLDYTRVHGANCEIVVGYVPLPVGLVGPLTVNGETVYVPMATTEGCLVASTNRGAKAITAGGGAT
AVLLRDGITRAPCVRMPSAAQAAHLKLWCETPQHFSTLKRAFESTTSFGKLLKCAPTVAGRNVYLRLTCFSGDAMGMNMVSKGSLA
VIETLQQEFPELVLVALSGNMCTDKKAAATNWLEGRGKSIVVEATIPKDVVTNTLKTTVHSMVQTNLHKNLIGSAMAGALGGFNAH
ASNIVTAVFLATGQDPAQNVESSNCITLLEETEEGDLWISCTMPSIEVGTVGGGTSLPAQAACLQAMGVRGGGATPGAHAQKLAQI
VASATLAGELSLLAALAANTLVQAHMQHNRKPAAK
```

>Phaeodactylum_tricornutum_tHMGR
```
DSISTKTSASNTLSARHPPAPSAPVCSDPSSFTPIHTSKTNITTLDNAAICQLVLTNQIKDHELEKRLDAHRAVQVRRLVVAHKLD
TLEHINAHALDNLPSEPSLDYTRVHGANCEIVVGYVPLPVGLVGPLTVNGETVYVPMATTEGCLVASTNRGAKAITAGGGATAVLL
RDGITRAPCVRMPSAAQAAHLKLWCETPQHFSTLKRAFESTTSFGKLLKCAPTVAGRNVYLRLTCFSGDAMGMNMVSKGSLAVIET
LQQEFPELVLVALSGNMCTDKKAAATNWLEGRGKSIVVEATIPKDVVTNTLKTTVHSMVQTNLHKNLIGSAMAGALGGFNAHASNI
VTAVFLATGQDPAQNVESSNCITLLEETEEGDLWISCTMPSIEVGTVGGGTSLPAQAACLQAMGVRGGGATPGAHAQKLAQIVASA
TLAGELSLLAALAANTLVQAHMQHNRKPAAK
```

>Fistulifera_solaris
```
MESLRLQLSQALSSFVSVSSNQLYGVIVGVTVFGCWGILSLAEQPELPEWQKTAVQRVPRQKEGPEPQWHLFRWINMLVVVAFACS
VAEFSLNATAYDSNVLCQFLMAWSLLLCYFFGFFGVMFVHDLEDSETVSTPVTKSSKIHPPASSTPICSDAIQSKQNKLLIIEDIK
DKKDAEIAQLVFTNQIKDHELEKQLDCHRAVTVRRLVVEQKLASLGHSGALEGLPYETDLDYSRVHGANCEIVVGYVPLPVGIIGP
LTINGESFYIPMATTEGCLVASTNRGAKAITEGGGALAQIVRDGITRAPCIRMPTAMAAAQLKLWCETPENFVQMKAAFESTTSFG
KLHECKATIAGKNVYLRLVCFSGDAMGMNMVSKGTLAVIECLRKEFPELSLVALSGNMCTDKKAAATNWLEGRGKSVVVEATIPKD
VVRKTLKTTVPAIVETNLNKNLVGSAMAGVIGGFNAHASNIVSAIFLATGQDPAQNVESSNCITLMEETESGDLWISCTMPSIEVG
TVGGGTSLAAQAACLKALGCQGGGANPGDNAQRLAKVIASATMAGELSLLAALAANTCYLLTDGFWRTNINSPIPSFFSVVAAHM
QHNRKPTK
```

>Pseudonitzschia_multiseries
```
MAETMGTMPTLGMRLDNMIAHIDQIPSTQIYGLIVAATVGLCVVLLGTGNSNLEIHHQRQMRLQQQSQNDDLKKPKMAASARSGKQ
PKWHLFKWINFVAVGGFLYSVFTFCSNASRYLHHESQGVLVQFLVGWSIFLMYFFGFFGVSLIHEDIPREEEESASIPSDPSSFKT
SKASSSIPENLKELGNDEIAALVLDNKVKDHMLEKLLDPFRAVTVRRIACNQKLAAVLGRNNNDTNVLDKLPSEPSLDYSRVFGAN
CEIVVGYVPLPVGLVGPLTINDESVYVPMATTEGCLVASSNRGAKAITQGGGARARIVRDGITRAPCLRMNSAMEAADLKIWCELP
QNFAILKQAFESTTSFGKLLECNPTVAGKNVYLRLVCFSGDAMGMNMVSKGSLAVIETLQKQFPTCQLIALSGNMCTDKKAAATNW
LLGRGKSVVVECVIPKEVVRTTLKTTVAALVHTNLNKNLIGSAMAGAIGGFNAHASNIVTAVFLATGQDPAQNVESSNCITLMEEQ
DDGDLWMCCTMPSIEVGTVGGGTSLPAQAACLEAIGCKGGGSTPGANAKKLATVVAAATMAGELSLLAALAANTLVQAHMVHNRKP
AAKK
```

>Thalassiosira_oceanica
MMAGTATIGQRLDALVATAQANTDGSAALFALVIGISVSFSFYLLNSGSGSTASAEMTKCGAQLSTCPKDNRTAIQKPRSPEDPPE
PNWNALKLTNCVAATGFLLSVLKFASNASAYLNDSTSLLQFLSIWSVFLLYFFGFFGIALVDLDGLVEEQPSPAPAAKSRQPISPS
KKIAKTVDVEPSPPVKVVNVHPPAKSAPVCSDPSSLKRHASSVPSNLKEMPNEEIAALVLQDKIKDHQLEKLLDPHRAVEVRRLKF
DAKLKSLGNGGALNDLPHKHDLDYKRVLGANCEIVVGYVPIPVGLAGPITLNGESVYIPMATTEGCLVASTNRGCKAITQGSGAFS
TILKDGITRAPCCRLPSAQQAAEVAIWIETPENFAELKRAFESTTSFGKLIEARPTVAGRNVYIRLKCFAGDAMGMNMISKGSLAV
IECLRSIFPALSLLALSGNVCTDKKAAAMNWIEGRGKSVVIEATIPKDVVSKTLKTSVKAIVDANVNKNLIGSAMASVVGGFNAHA
ANNVTAVFLATGQDPAQNVESSNCITIMEETSDGNLWVSCTMPSIEVGTVGGGTGLPAQAAACLRAIGVKGGGEKPGQNAKQLAHVV
AAATLAGELSLMAALASNSLVAAHMAHNRKPPTASKVEKAFDEVQHGFDDVWLSGQQQHGKKLLDRCSHPHKNWLLGLLIKLQARL
RSRLTVAVGLAPRPQDHRYFIGITRYAPLGPIIGISSALRVMRLWDQSRPKQIWQRFLTDPTQTPYTIHDQLLGLTLVLKSAKNLQ
KSEMTPAAERRGSINRGYRITDSTCGFCSTEVCEWITDRLDHDNWDTCSGQRQGASWTNNTAIQSLVVGGQEPARAPLYLHYVPRL
YKAHARSKITGVVCFRLQIDPLSYASWR

>Extubocellulus_spinifer
MTTVPTLGERLDALVDRIDSIPSTQLYIAAVVTSVVFSCMLLNSGGSNGVVDPAKIPDNIGRPPSVRRKRRDGSNPEDEPEPRWHI
LTILNYAAVTSFTSSVLYFATDASRFLADSSSLLKFITGWSVFLCYFFGFFGISFVDADGLIAREEHEAAARQQQQRMQKQHSLGI
SSPPSKSKSKLHSAACSVPVSSDGASTKKRSSSSSCTPVDVKELADEKIASLVLTGQIKDHTLEKILGCDRAVTVRRLAFDCKLST
LDRGGALSDLPSGPSLDYDRVFGANCEVVVGYIPIPVGMVGPLTLNDETFFVPMSTTEGCLVASTNRGCKAISQGTGAISTVIRDG
ITRAPCVRMKSAKEAADLKIWCEQKDNFALLKKEFESTTSFGKLLSVSPIVAGKNVYIRLQCFSGDAMGMNMVSKGSLAVIDLLKS
TFPTLELVALSGNVCCDKKSGAINWIEGRGKSVVVEATIPLEVVRSTLKTNVKTIVHVNIQKNLIGSAMASTVGGFNAHASNIVTA
IFLATGQDPAQNVESSNCITLMEETDEGDLWISCTMPSIEVGTVGGGTSLPAQSACLKAIGCYGSGDTPGQNAKKLASVVAATTMA
GELSLLAALASNTLVQAHMAHNRKPATSK

>Attheya_septentrionalis
MTNTVASVPTVGERLDALIDTIDSLPSTHLYGAVVAVSVVFSFALLNWGGPSNHTREDAPQKRRLSWEMRRSAASASAVSFKSGET
EPQPKWHILRVLNYAAVGSFAASVVIFCRGASVYLNDPATLLQFLVGWSLFLCYFFGFFGISFVDPDGMMMSYDPQQSQANNDPQD
STNCKSSSAKKPLHPPAAATPVCSDGKASCKAPVVAAAALSSSAIKEMSDEQVAQLVVSNQVKDHMLEKVLDPHRAVAVRRLAFDT
KLASLDRGGALDELPSGPSLDYARVFGANCEIVVGYVPLPVGMVGPLTLNNETVYVPMATTEGCLVASTNRGAKAICQGGGAHSCI
VRDGITRAPCVRMKTAMEAAKLKNWCEEPQHFAKLKAAFESTTSFGKLEQCNPTVAGRNVYLRLRCFSGDAMGMNMVSKGSLAVID
LLKSEFPTLFLVALSGNMCTDKKPAAINWIEGRGKSVVVEAIIPKDVVRQTLKTTVQSICSTNIHKNLIGSAMAGAMGGFNAHAAN
IVTAIFLATGQDPAQNVESSNCITLMEETEEGNLWISCTMPSIEVGTVGGGTSLPAQSACLKAIGCKGGGAKPGQNAQQLAHVVAA
ATMAGELSLLAALAANTLVAAHMQHNRKPTTTSK

>Craspedostauros_australis
MNMVSKGSLAVIKLLEERFENMQLVALSGNMCTDKKAAATNWLEGRGKSVVVECVIPQSIVQKTLKTTVKALIHTNINKNLIGSAM
AGALGGFNAHASNIVAAVFLATGQDPAQVVESANCITLMEENDNGDLWMSCTMPSIEVGTVGGGTSLAAQAACLDAVGCRGGGKNH
PGENAQKLSQVVAAATMAGELSLLSALAANTLVQAHMKHNRKAPTKK

>Thalassiosira_weissflogii
MAHPSSYERVSSPPPTIGKHIDAVIAKIDSLPSSYVYTLAIAASLTFSFYLLNSGGGSNSMQMEPHGDVDRASAIKKPKQKSRRSL
DEPQPKWHILKLTNYVATSGFLLSVLKFVSNASMYLNDSTSLLQFLTIWSIFLCYFFGFFGISFIELDDFEQESHPQHQAPVSAIN
VSHAKHCNAVKSGSLEPPVKVINVHPHSTSAPVCSDPASFRKPSSTIPGNLKDMSNEEIASLVLDDKIKDHQLEKLLDPHRAVEVR
RLKFDAQLQSLGRGGALAELPHKHDLDYKRVLGANCEIVVGYVPIPVGMAGPITLNGESVYIPMATTEGCLVASTNRGCKAITQGS
GAFSTIVRDGITRAPLVRLPSAREAAEVLLWIEDENNFKILKEAFESTTSFGKLIEARPTVAGRNVYIRLRCFSGDAMGMNMISKG
SLAVIECLRRKFPQLSLAALSGNMCTDKKAAAINWIEGRGKSVVIEATIPKEVVRSTLKTSVPAIVEANTNKNLIGSAMAATVGGF
NAHAANNVTAVFLATGQDPAQNVESSNCITLMEETSDGDLWISCTMPSIEVGTVGGGTGLPAQSACLKAIGVKGGGDSPGDNSKKL
AHVVAAATMAGELSLMAALASNSLVAAHMAHNRKPVSK

>Odontella_Sinensis
MASSSVAYMTVGQRLDSFIESIDPSLSPTTKLYACVVCASVGFSFLLLNGRGSNVGGPDASGGPKSASKNESSSSGPSRRSGGGVS
AASASSGGGRREPKWGVLKALNVVAALAFLISVIRFASDASRHMSDSTSLLKFMSIWSALLCYFFGFFGISFVDAEGFSVAGEETK
TTEKTTKKAAGEEEDDDEDGGPGKKTAPAGSRSIPEKPQHANLHPPAPSAPVCSDPTSRAPITKKALPPNLQGMPDNEIAALVLSG
VVKDHQLEKLLSPARAVPVRRIVFERKLSSLNRGGALDELPHEPSLDYGRVHGANCEIVVGYVPLPCGVIGPLTLDGETVYVPMAT
TEGCLVASANRGCKAISQGGGARSAVIRDGITRAPCVRMRSAMEAAELKLWCEDPANFARLKSAFESTTSFGKLLGAHPTVAGRNV
YLRLRCFSGDAMGMNMVSKGSLAVIELLRSTFPSLELVALSGNMCTDKKAAATNWIEGRGKSVVVEATIPKDVVRATLKTTVPAMV
NTNLQKNLVGSAMAGALGGFNAHAANIVTAIFLATGQDPAQNVESSNCITLLEETPTGDLWISCTMPSVEVGTVGGGTGLPAQASC
LRAIGCKGGGANPGDNARKLARVVAASVLAGELSLLAALAANTLVQAHMIHNRKKPTAGK

>Rhizosolenia_setigera
XNNPHFSATMTETIGQRXDALLESIDSVPSLQLYAAIVTSVVFSFVILNTGSHNAGLFQNNEQQRTISSSQRKGSPTMKSNTANYT
NTSFDRTKDNDQNQPQIKWYLLKTLNFAVAIGFVISITKFSSDASTYITDSISILKFLCAWSLCLCYFLGXFGMSFVDSEDLCSVP
TAPPAGATSPVAAVHKKQVSKAASFSRPIHPPAMSTHPVCGTGEISSPTTTASLPDDIQSLSNEVAASLVVSNKIKDHTLEKHFGP
FRAVXXRRLAFETKLASLGHSGALDDLPSGPSLDYKRVFGANCEIVVGYVPIPVGMCGPITLNGESVYIPMATTEGCLVASTNRGC
KAISAGSGATSVILRDGITRAPCLRMKSAKEAADLKLWCEEQENFLLLKXAFESTTSFGKLLSAEPTVSGKNVYXRLRCFSGDAMG
MNMISKGSLAVVEXLRELFPTLSLIALSGNMCTDKKAAATNWIEGRGKSVVIEATIPKDVVTKILKTDVKSIVSTXLQKNLIGSAM
AGALGGFNAHAANNVTAIFLATGQDPAQNVESSNCITLLEETEDGDLWICCTMPSIEVGTVGGGTGLPAQSSCLKMIGCKGGGETP
GDNAKQLAHVVAAGTMAGELSLLAALASNTLVAAHMQHNRKPAQKK

>Coscinodiscus_wailesii
MTHAPANQHTTLGMRIDAVLQEIESISPTVLYSVILVASILFSMILLNTGGGRSTEPPRLRDDVTKPSRKIQRFSDEPQPKWYILK
FCNYAAVIGFAFSIFKFAIDATTYINDSTLLMKFTVCWSLFMCYFFGFFGISFVDADEIEQRAVAGGFDGREGREVERKGILKGSI
TRKEHEYPIHSPPQSIHPVCNPNNADIAKKQLPDKSPSSTRISDIKSLSDVEIASLVTSNKIKDHQLEKLLDPRRAVDVRRIVVDS
KLSSLGRGGALADLPSSPSLDYSRVFGANCEIVVGYVPIPVGMVGPLTLNGETVYIPMATTEGCLVASTNRGCKAISAGSGAHSSI
VRDGITRAPCLRLNSTREAADLMLWCEDESNFKLLKQAFESTTSFGKLLRVDPTIAGRNVYLRLKCFSGDAMGMNMVSKGSLAVVD
YLRTVFPTITLVALSGNVCTDKKAAAINWIEGRGKSVVVEATIPQHVVRSTLKTTVEAIVSTNIQKNLIGSAMAGTVGGFNAHASN
IVTAVFIATGQDPAQNVESSNCMTLMEETPTGDLWISCTMPSIEVGTVGGGTGLPAQAACLKAMGCQGGAADKPGANAKQLAHVVA
SATMAGELSLMAALAANTLVQAHMQHNRKPTIGKKA

>Skeletonema_marinoi
MTFSEPTTIGMRIDAVIASLESLPHTNPEILYASLIGASLMFSFLVLNSGGGSSMALPDGGDADSLNKPKKISNAKPVSSSDEPQPK
WHILRLTNYVVTFGFLLSVLKFASNATTYLDDSTSLLQFLIIWSISMCYFFGFFGISFIELDELESEPQNHQHQMLRKPMPTEESA
EPPVKVVKMHPPASSAPVCTDPTSFKKPSPSIPSNIKELSNGEIASLVLQDKVKDHQLEKLLDPHRAVEVRRLKVDNQLDSLGRGG
ALAELPHKHDLDYKRVLGANCEIVVGYIPIPVGMAGPITLNGESVYIPMATTEGCLVASTNRGCKAITQGSGAVSSILRDGITRAP
CVRLPSAKEASEVFLWIERPENFQKLKEVFESTTSFGKLLEARPSVAGKNVYIRLRCFAGDAMGMNNMISKGSLAVIECLRQVFPNI
SLVALSGNMCTDKKAAAINWIEGRGKSVVIEATIPKGVVRSTLKTSVPAIVEANVNKNLIGSAMAGVVGGFNAHASNNVTAVFLAT
GQDPAQNVESANCITLMEETPEGDLWISCTMPSIEVGTVGGGTGLPAQAACLKAIGVKGGGENPGDNAKQLAHVVAAATMAGELSL
MAALASNSLVAAHMTHNRKPASK

>Skeletonema_menzelii
MTSEATTIGMRIDAVIASLESLPHTNPEILYASVIGASLLFSFLVLNSGGGSSMALPDGGDADSLNKPKISKAKLVSPSDEPQPKWH
ILRLTNYIVTFGFLLSVLKFASNATTYLDDSTSLLQFVIIWSISMCYFFGFFGISFIELDELENEPQNHQHQILRKPMPTKESADP
PVKVVKMHPPASSAPVCIDPTSFKKPSPSIPSNIKELSNSEIASLVLQDKVKDHQLEKLLDPHRAVEVRRLKVDVQLDSLGRGGAL
AELPHKHDLDYKRVLGANCEIVVGYVPIPVGMAGPITLNGESIYIPMATTEGCLVASTNRGCKAISQGSGAVSSILRDGITRAPCV
RLPSAKEASEVHLWIESPENFQKLKEVFESTTSFGKLLEARPSVAGKNVYIRLRCFAGDAMGMNNMISKGSLAVIECLRQRFPNLSL
VALSGNMCTDKKAAAINWIEGRGKSVVIEATIPKGVVRSTLKTSVPAIVEANLNKNLIGSAMAGVVGGFNAHASNNVTAVFLATGQ
DPAQNVESANCITLMEETPEGDLWISCTMPSIEVGTVGGGTGLPAQAACLKAIGVKGGGENPGDNAKQLAHVVAAATMAGELSLMA
ALASNSLVAAHMTHNRKPASK

>Thalassiosira_miniscula
MPSSPTAPPPPPTLGMQIDAVIASLESISTPQLYGTVIAISLAFSFYLLNSGGGSSSPALQMMTNDTKEQQDADADSLRKPARAAA
TSSSSSSESSSEDRPEPRWHILRFTNYVAALGFLVSVLQFASNASAYLNDSNSLLQFLTVWSLFVVYFVGFFGISFVELDDFVDGA
QISSSVARQRAQQQQQQQQQQQSQNMTVAASQAKTPGRKSHKTDGSDSPPIKVVNVHPPATSAPVCSDPKSFKKPSAASIPPNLKE
LPNSEIASLVLQDKIKDHQLEKLLDPHRAVQVRRLKFDARLSALGNGGALDELPHRHDLEYKRVLGANCEIVVGYVPIPVGMAGPV
TLNGESVYIPMATTEGCLVASTNRGCKAISQGSGASSTILKDGITRAPCVRLPSAKEAAEVALWIETPENFATLKEAFESTTSFGK
LLSARPTVAGKNVYVRLRCFSGDAMGMNNMISKGSLAVIECLRQIFPALSLVALSGNMCTDKKAAAMNWIEGRGKSVVIEATIPADV
VRSTLKTSVTAIVEANLNKNLIGSAMAGTVGGFNAHAANNVTAVFLATGQDPAQNVESSNCITLMEETPEGDLWMSCTMPSIEVGT
VGGGTGLPAQAACLRAIGAKGGGENPGDNARRLAHVVAAATMAGELSLMAALASNSLVAAHMAHNRKPASK

>Thalassiosira_punctigera
XDTSELKGSKSSNRSPIXPSQNSESMGDATSAPPTVGMQIDAAIAQLENVSTTQLYAIVIAASLAFSFYLLNSGGGSSALTIMNEK
DGDADSLRKPEARKRKDRPRASTSSDEPEPRWHILRITNYVVAAGFLLSVLQFASNASTYLNDSTSLLQFLSVWSIFLCYFFGFFG
ISFIELDDFVDQQPSQQQQQPSHKLKKAESISPDPPVKVVNVHPPSSSAPVCTDPTSFKKPPSNVPGNLKDLPNSEIASFVLQDKI
KDHQLEKLLDPHRAVEVRRLKFDAKLDSLGRGGALADLPHKHDLDYKRVLGANCEIVVGYLPIPVGLAGPITLNGESVYIPMATTE
GCLVASTNRGCKAISQGSGASSTILKDGITRAPCVRLPSAKEAAKVALWIGTPENFLTLKKAFESTTSFGKLLDATPTVAGRNVYI
RIRCFSGDAMGMNMISKGSLAVIETLKKVFPDLSLLALSGNMCTDKKAAALNWIEGRGKSVVIEATIPKDVVRSTLKTSVQAIVEA
NVNKNLIGSAMAGTVGGFNAHAANNVTAVFIATGQDPAQNVESSNCITLMEETSEGDLWISCTMPSIEVGTVGGGTGLPAQSACLK
VIGVKGGGENPGDNARQLAHVVAAATMAGELSLMAALASNSLVAAHMSHNRKPTSK

>Ditylum_brightwellii
KMATRLTETTDVTIGMRLDELISSIDTIPSTQLYMAVVVVSVIFSFLLLNSGGTSIEQEAPPSSPPSTALKKITRQGPEPKWHILK
ILNYAAVSSFLVSVGCFASDASRHMNDSSSLFKFLFGWSFFLCYFFGFFGISFVDADEIMEGQNFESSERSTSSKSKAVHPPAEST
PVCTVPSSFEKPSLSKTDLQSKSNEELSNMVLTGQIKDHQLEKLLDHHRAVDVRRLAFDSKLSSVGCGGALTDLPSGPSLDYSRVF
GANCEIVVGYVPLPVGMVGPLTLNGQSVYIPMATTEGCLVASTNRGCKAITQGSGATSVILRDGITRAPCVRMNSAKEAAELKLWC
EMPANFATLKHAFESTTGFGKLLSVEPTVAGKNAYLRLQCFAGDAMGMNMVSKGSLAVIDLLKSVFPTLVLVALSGNMCTDKKAAA
TNWIHGRGKSVVVEAIIPQHVVRTTLKTTVQAIVQTNIHKNLIGSAMAGAIGGFNAHAANNVTAVFLATGQDPAQNVESSNCITLL
EETEDGDLWICCTMPSIEVGTVGGGTSLPAQSACLKAIGCKGGGVNPGDNAKQLAHVVAAATMAGELSLLAALAANTLVQAHMTHN
RKPTTAKST

>Thalassiothrix_antarctica
MTSTTPTTLGMKLDTFLSNMDSLSSTELYGLIVVITVIISFSLLGYGGNQDMEEKELRLPSPKRSSLQKQQKSGKQPKWFVLKILN
YLAVGSFLASVGLFSWNAKLYLAKEDSSALIRFLIGWSIFLCYFFGFFGISFVDHDDFVVDVNGATTTVDNSNNDKNSSIISKPRK
MLHEPAPSAPLCSDPPMKKSSSSAVPKKEDIRAMNDASIAQLVLTKKMKDHELEKCLDASRAVVVRRMIMAEKLSSLGLGQNCLQN
LPSNPSLDYSRVYGANCEVVVGYVPLPVGMVGPLTLNGESVYIPMATTEGCLVASTQRGVKAISSGKQGAICTILKDGITRAPCLR
LPSATEAAQLAVWVSDPNNFQILKKAFESTTSFGKLLSAKTTVAGKNAYLRMVCFSGDAMGMNMVSKGSLAVIDLLRTHFPRIKLI
ALSGNMCTDKKAAATNWLEGRGKSVVVEAVIPKSVVEKTLKTTVKALVHTNIQKNLVGSAMAGALGGFNAHAANIVTAIFLATGQD
PAQNVESSQCITLMEETEDGQGDLWISCTMPCLEVGTVGGGTSLPAQRACLECMGVAGGNRETPGANAKKLATMVAGATMAGELSL
LAALAANTLVQAHMKHNRKPAVAKK

>Cylindrotheca_closterium
MASVISNYIDELSSTQLYTAIVGATVTLCVVLLGPAAGGNENLPNNMLGTTTTSNDNNKKNAQQKQQPKWYLFKYINVGVFCMFVS
SVAIFLWDASTYIHNGDRMTQFLVGWSLCLCYFFGFFGVFFIHQDLLMNQNQNQQGDSSDIAENKQEMVPKTKEVKSVSPPPKAKV
VHEPAACAPVCSDPASFNSSSSSNKKAAISKDTDITGMTNDEIAQLVLDLQVKDHELEKRLDPFRAVTVRRMIAQHKLQSVHSASS
SAFLEDLPEGPSLDYSKVFGTNCEMVVGYVPLPVGMVGPLTLNGESVYFPMATTEGCLVASTQRGAKAISQGTLGAQALIVKDGIT
RAPCVRMASAMEAAQLKLWCQEAANLQVLKEAFESTTSFGKLLECHATVAGKNVYLRLVCFSGDAMGMNMVSKGSLKVIETLQNIF
PNLELVALSGNMCTDKKAAATNWLHGRGKSVVVEAVIPKDVVKKTLKTTVPALVSTNQNKNLIGSAMAGAMGGFNAHASNIVTAIF
LATGQDPAQNVESANCITLMEQVEETGDLWISCTMPSIEVGTVGGGTTLPAQAACLEILGCKGGSRENPGKNAQKLALVVAAATMA
GELSLMSALAANTLVQAHMKHNRKKTV

>Nitzschia_punctata
MAATFDTMPTTIGMKLDAMIAEIDNLSSTQLYGVIVAATVVLCVVLLGTGHSNLDLQHSNNNDPLLKKQPAVAPSGNIKQPRWHIF
KWINYLAVAAFLWSVCTFCLNASQYLHHEESQGVLVKFLLGWSVFLLYFFGFFGVSLIHEDIPKEEAGAASSAANRLSSVSQSNKT
KTVETSSKNKALHGAAACTPVCSDPSSFKVSSSVPSNIKELADEDVADLVLKNKVKDHELEKRLDPFRAVTVRRMTANRKLASVLP
QNKPNVLDKLPATPSLDYSKVHGANCEIVVGYVPLPVGLVGPLSLNNETVYVPMATTEGCLVASTNRGAKAITQGGGAQARIVRDG
ITRAPCVRMESAMGAADLKVWCEKPENFARLKQAFEGTTSFGKLQACHPTVAGKNVYLRLVCFSGDAMGMNMVSKGSLAVIETLQK
EFPSLQLVALSGNMCTDKKAAATNWLQGRGKSVVVEATIPKDVVRTTLKTTVAALVHTNMHKNLIGSAMAGSLGGFNAHASNIVTA
VFLATGQDPAQNVESSNCITLMEETDEGDLWISCTMPSIEVGTVGGGTSLEAQAACLEAIGCKGGGATPGENAKKLATVVAAATMA
GELSLLAALAANTLVQAHMTHNRKSNKK

>Nitzschia_sp.
VLSQVDELSSTQLYGLIVAATVTLCVFLLGTGSNTDVEFAATNSKMKTGTNDDGLKKPNASTNNTASTTRQPRWHIFKWVNYLASA
AFLSSVGMFCMNASQYLHHESKGVLVQFLVGWSVFLMYFFGFFGISLIHEDIPKDNQDDDQGIEQPRNTTTRYAAGSSSTTAAAKKQ
ALHTAAPSAPVCSDPASFKASSSSSVPSNLKELPDEDIANLVLMNKLKDHELEKRLDPFRAVTVRRLVVNQKLSTLLTDGKTNQPV
TNVLEKLPSTPSLDYSRVFGANCEIVVGYVPLPVGLVGPLSLNDETVYVPMATTEGCLVASTNRGAKAITQGGGAQARIVRDGITR
APCVRMASAMEAAADLKVWCEEPQNFAVLKQAFESTTSFGKLQACNPTVAGKNVYLRLVCFSGDAMGMNMVSKGSLAVIETLQAKFP
SLQLVALSGNMCTDKKAAATNWLHGRGKSVVVEAIIPKDVXPWNSQNNSLSLNDETVYVPMATTEGCLVASTNRGAKAITQGGGAQ
ARIVRDGITRAPCVRMASAMEAAADLKVWCEEPQNFALLKQAFESTTSFGKLQACNPTVAGKNVYLRLVCFSGDAMGMNMVSKGSLA
VIETLQAKFPSLQLVALSGNMCTDKKAAATNWLHGRGKSVVVEAIIPKDVVRGTLKTTVDALIFTNTNKNLIGSAMAGSIGGFNAH
ASNIVTAVFLATGQDPAQNVESSNCITLMEPTDDGDLWISCTMPSIEVGTVGGGTSLPAQSACLEAIGCKGGGATAGENAQKLARV
VAAATMAGELSLLAALAANTLVQAHMQHNRKPAPKKELTLTPSFVQNNTPIVPFLRKAHLERTYRSVLGCSQIVIRKSRAPCVRMA
SAMEAAADLKVWCEEPQNFAVLKQAFESTTSFGKLQACNPTVAGKNVYLRLVCFSGDAMGMNMVSKGSLAVIETLQAKFPSLQLVAL
SGNMCTDKKAAATNWLHGRGKSVVVEAIIPKDVVRGTLKTTVDALVHTNTNKNLIGSAMAGSIGGFNAHASNIVTAVFLATGQDPA
QNVESSNCITLMEPTDDGDLWISCTMPSIEVGTVGGGTSLPAQSACLEAIGCKGGGATPGENAQKLATVVAAATMAGELSLLAALA
ANTLVQAHMQHNRKPAPKK

>Pseudonitzschia_pungens
MAQTMGTMATIGTRIDNVIAQFDQLPSTQLYGLIVAATVGLCVVLLGTGNSNFELHHQRQKQLQQNNDDLKKPKMAAMPSGRQPRW
HIFKWINFVAVGAFLCSVFIFCSNASRYLHHESQGVLVQFLVGWSVFLMYFFGFFGVSLIHDDIPREEEESLPSKPQTKSVPTNTS
APKKTVHPPAACTPVCSDPSSFKSSKPSVPENLKELSDDEIAALVLGNKVKDHMLEKLLDPFRAVTVRRIACNQKLASVLGREDKS
NVLEKLPSEPSLDYSRVFGANCEIVVGYVPLPVGLVGPLTINDESVYVPMATTEGCLVASSNRGAKAITQGGGARAKIVRDGITRA
PCLRMNTAMEAADLKIWCEKPQNFAILKQAFESTTSFGKLKECNPTVAGKNVYLRLVCFSGDAMGMNMVSKGSLAVIETLQKYFPT

CQLVALSGNMCTDKKAAATNWLYGRGKSVVVECVIPKEVVRTTLKTTVSALVHTNLNKNLIGSAMAGAIGGFNAHASNIVTAVFLA
TGQDPAQNVESSNCITLMEEEVNGDLWMCCTMPSIEVGTVGGGTSLPAQAACLEAIGCKGGGVTPGANAKKLATVVAAATMAGELS
LLAALAANTLVQAHMAHNRKPASKK

>Saccharomyces_cerevisiae
MPPLFKGLKQMAKPIAYVSRFSAKRPIHIILFSLIISAFAYLSVIQYYFNGWQLDSNSVFETAPNKDSNTLFQECSHYYRDSSLDG
WVSITAHEASELPAPHHYYLLNLNFNSPNETDSIPELANTVFEKDNTKYILQEDLSVSKEISSTDGTKWRLRSDRKSLFDVKTLAY
SLYDVFSENVTQADPFDVLIMVTAYLMMFYTIFGLFNDMRKTGSNFWLSASTVVNSASSLFLALYVTQCILGKEVSALTLFEGLPF
IVVVVGFKHKIKIAQYALEKFERVGLSKRITTDEIVFESVSEEGGRLIQDHLLCIFAFIGCSMYAHQLKTLTNFCILSAFILIFEL
ILTPTFYSAILALRLEMNVIHRSTIIKQTLEEDGVVPSTARIISKAEKKSVSSFLNLSVVVIIMKLSVILLFVFINFYNFGANWVN
DAFNSLYFDKERVSLPDFITSNASENFKEQAIVSVTPLLYYKPIKSYQRIEDMVLLLLRNVSVAIRDRFVSKLVLSALVCSAVINV
YLLNAARIHTSYTADQLVKTEVTKKSFTAPVQKASTPVLTNKTVISGSKVKSLSSAQSSSSGPSSSSEEDDSRDIESLDKKIRPLE
ELEALLSSGNTKQLKNKEVAALVIHGKLPLYALEKKLGDTTRAVAVRRKALSILAEAPVLASDRLPYKNYDYDRVFGACCENVIGY
MPLPVGVIGPLVIDGTSYHIPMATTEGCLVASAMRGCKAINAGGGATTVLTKDGMTRGPVVRFPTLKRSGACKIWLDSEEGQNAIK
KAFNSTSRFARLQHIQTCLAGDLLFMRFRTTTGDAMGMNMISKGVEYSLKQMVEEYGWEDMEVVSVSGNYCTDKKPAAINWIEGRG
KSVVAEATIPGDVVRKVLKSDVSALVELNIAKNLVGSAMAGSVGGFNAHAANLVTAVFLALGQDPAQNVESSNCITLMKEVDGDLR
ISVSMPSIEVGTIGGGTVLEPQGAMLDLLGVRGPHATAPGTNARQLARIVACAVLAGELSLCAALAAGHLVQSHMTHNRKPAEPTK
PNNLDATDINRLKDGSVTCIKS

>Saccharomyces_cerevisiae_tHMG
PVLTNKTVISGSKVKSLSSAQSSSSGPSSSSEEDDSRDIESLDKKIRPLEELEALLSSGNTKQLKNKEVAALVIHGKLPLYALEKK
LGDTTRAVAVRRKALSILAEAPVLASDRLPYKNYDYDRVFGACCENVIGYMPLPVGVIGPLVIDGTSYHIPMATTEGCLVASAMRG
CKAINAGGGATTVLTKDGMTRGPVVRFPTLKRSGACKIWLDSEEGQNAIKKAFNSTSRFARLQHIQTCLAGDLLFMRFRTTTGDAM
GMNMISKGVEYSLKQMVEEYGWEDMEVVSVSGNYCTDKKPAAINWIEGRGKSVVAEATIPGDVVRKVLKSDVSALVELNIAKNLVG
SAMAGSVGGFNAHAANLVTAVFLALGQDPAQNVESSNCITLMKEVDGDLRISVSMPSIEVGTIGGGTVLEPQGAMLDLLGVRGPHA
TAPGTNARQLARIVACAVLAGELSLCAALAAGHLVQSHMTHNRKPAEPTKPNNLDATDINRLKDGSVTCIKS

>Arabidopsis_thaliana
MDLRRRPPKPPVTNNNNSNGSFRSYQPRTSDDDHRRRATTIAPPPKASDALPLPLYLTNAVFFTLFFSVAYYLLHRWRDKIRYNTP
LHVVTITELGAIIALIASFIYLLGFFGIDFVQSFISRASGDAWDLADTIDDDDHRLVTCSPPTPIVSVAKLPNPEPIVTESLPEED
EEIVKSVIDGVIPSYSLESRLGDCKRAASIRREALQRVTGRSIEGLPLDGFDYESILGQCCEMPVGYIQIPVGIAGPLLLDGYEYS
VPMATTEGCLVASTNRGCKAMFISGGATSTVLKDGMTRAPVVRFASARRASELKFFLENPENFDTLAVVFNRSSRFARLQSVKCTI
AGKNAYVRFCCSTGDAMGMNMVSKGVQNVLEYLTDDFPMDVIGISGNFCSDKKPAAVNWIEGRGKSVVCEAVIRGEIVNKVLKTS
VAALVELNMLKNLAGSAVAGSLGGFNAHASNIVSAVFIATGQDPAQNVESSQCITMMEAINDGKDIHISVTMPSIEVGTVGGGTQL
ASQSACLNLLGVKGASTESPGMNARRLATIVAGAVLAGELSLMSAIAAGQLVRSHMKYNRSSRDISGATTTTTTTT

>Homo_sapiens
MLSRLFRMHGLFVASHPWEVIVGTVTLTICMMSMNMFTGNNKICGWNYECPKFEEDVLSSDIIILTITRCIAILYIYFQFQNLRQL
GSKYILGIAGLFTIFSSFVFSTVVIHFLDKELTGLNEALPFFLLLIDLSRASTLAKFALSSNSQDEVRENIARGMAILGPTFTLDA
LVECLVIGVGTMSGVRQLEIMCCFGCMSVLANYFVFMTFFPACVSLVLELSRESREGRPIWQLSHFARVLEEEENKPNPVTQRVKM
IMSLGLVLVHAHSRWIADPSPQNSTADTSKVSLGLDENVSKRIEPSVSLWQFYLSKMISMDIEQVITLSLALLLAVKYIFFEQTET
ESTLSLKNPITSPVVTQKKVPDNCCRREPMLVRNNQKCDSVEEETGINRERKVEVIKPLVAETDTPNRATFVVGNSSLLDTSSVLV
TQEPEIELPREPRPNEECLQILGNAEKGAKFLSDAEIIQLVNAKHIPAYKLETLMETHERGVSIRRQLLSKKLSEPSSLQYLPYRD
YNYSLVMGACCENVIGYMPIPVGVAGPLCLDEKEFQVPMATTEGCLVASTNRGCRAIGLGGGASSRVLADGMTRGPVVRLPRACDS
AEVKAWLETSEGFAVIKEAFDSTSRFARLQKLHTSIAGRNLYIRFQSRSGDAMGMNMISKGTEKALSKLHEYFPEMQILAVSGNYC
TDKKPAAINWIEGRGKSVVCEAVIPAKVVREVLKTTTEAMIEVNINKNLVGSAMAGSIGGYNAHAANIVTAIYIACGQDAAQNVGS
SNCITLMEASGPTNEDLYISCTMPSIEIGTVGGGTNLLPQQACLQMLGVQGACKDNPGENARQLARIVCGTVMAGELSLMAALAAG
HLVKSHMIHNRSKINLQDLQGACTKKTA

>Oryza_sativa
MDVRRGGGGGGRIVGAARRALTWGALPLPMRITNGLAMVSLVLSSCDLLRLCSDRERPLGGREFATVVYLVSLFAHPDAPATTTGDD
DDGQGGSRRARPAAAEPAPMHGHGGGMMEADDEEIVAAVASGALPSHRLESRLGDCRRAARLRREALRRVTGRGVEGLPFDGMDYQ
AILGQCCEMPVGYVQLPVGVAGPLLLDGREYHVPMATTEGCLVASVNRGCRAISASGGAFSVLLRDAMSRAPAVKLPSAMRAAELK
AFAEAPANFELLAAVFNRSSRFGRLQDIRCALAGRNLYMRFSCITGDAMGMNMVSKGVENVLGYLQNVFPDMDVISVSGNYCSDKK
PTAVNWIEGRGKSVVCEAIIKGDVVQKVLKTTVEKLVELNIIKNLAGSAVAGALGGFNAHASNIVTALFIATGQDPAQNVESSQCI
TMLEEVNDGDDLHISVTMPSIEVGTIGGGTCLASQAACLNLLGVKGSNHGSPGANAKRLATIVAGSVLAGELSLLAALASGHLVKS
HMMYNRSSKDVAKAAS

>Solanum_lycopersicum
MDVRRRSEEPVYPSKVFAADEKPLKPHKKQQQQQEDKNTLLIDASDALPLPLYLTTNGLFFTMFFSVMYFLLSRWREKIRNSTPLH
VVTLSELGAIVSLIASVIYLLGFFGIGFVQTFVSRGNNDSWDENDEEFLLKEDSRCGPATTLGCAVPAPPARQIAPMAPPQPSMSM
VEKPAPLITSASSGEDEEIIKSVVQGKIPSYSLESKLGDCKRAASIRKEVMQRITGKSLEGLPLEGFNYESILGQCCEMPIGYVQI

116

PVGIAGPLLLNGKEFSVPMATTEGCLVASTNRGCKAIYASGGATCILLRDGMTRAPCVRFGTAKRAAELKFFVEDPIKFESLANVF
NQSSRFARLQRIQCAIAGKNLYMRLCCSTGDAMGMNNMVSKGVQNVLDYLQNEYPDMDVIGISGNFCSDKKPAAVNWIEGRGKSVVC
EAIITEEVVKKVLKTEVAALVELNMLKNLTGSAMAGALGGFNAHASNIVSAVFIATGQDPAQNIESSHCITMMEAVNDGKDLHISV
TMPSIEVGTVGGGTQLASQSACLNLLGVKGANREAPGSNARLLATVVAGSVLAGELSLMSAISSGQLVNSHMKYNRSTKDVTKASS
>Thalassiosira_rotula
MVVATSAPPTVGMQIDAAIASLETLSTTQLXATVIAVSLXFSFYLLNSGGESSALTMMNESSSDSLRKPKXDRKRMKTQQQQSREE
TDEPEPRWHILRITNYVVATGFTLSVLQFASNASTYLNDSTSLLQFLMVWSLFLCYFFGFFGISFIELDDFVDQQQPSQQQPSQQT
QVAPSRQLKKAESSDPPVKVVNIHPPASSAPVCTDPTSFKKPSSKIPSNLKELPNSEIASLVLQDKIKDHQLEKLLDPHRAVEVRR
LKFDAKLDSLGCGGALTELPHKHALDYKRVLGANCEIVVGYIPIPVGMAGPITLNGESVYIPMATTEGCLVASTNRGCKAISQGSG
ASSTILKDGITRAPCVRLPSAREAAEVALWIGKTXNFLSLKEAFESTTSFGKLLDATPTVAGKNVYIRLRCFSGDAMGMNMISKGS
LAVIDALRQIFPRLSLLALSGNMCTDKKAAAMNWIEGRGKSVVIEATIPQDVVRSTLKTSVRAITEANVNKNLIGSAMAGTVGGFN
AHAANNVTAVFLATGQDPAQNVESSNCITLMEETPEGDLWISCTMPSIEVGTVGGGTGLPAQSACLKAIGVKGGGENPGDNAKQLA
HVVAAATMAGELSLMAALASNSLVAAHMAHNRKPASK

>Pseudonitzschia_delicatissima
CKRTFPSXRSNTDLHPSNKNMAETMTPMPTIGMQLDAMIAKVDALPSTQLYGLIVAVTVGLCVVVLGTGNSNLEMEQQQQDLKKPK
MVVVDGGKQPNWHLFKMINYIAVAAFLYSVFIFCSNASQYLHHGSQGVLAQFLVGWSVFLMYFFGFFGVSLIHDDIPSDTDSVPSQ
PQTKSVAPKPVHKPAACTPVCSDPSSFKSSKSSVPENIKELGDEEIADLVLSNKVKDHMLEKLLDPFRAVKVRRIACNQKLNSVLA
TNSTNVLDKLPSEHDLDYGRVYGANCEIVVGYVPLPVGLVGPLTINDESFYVPMATTEGCLVASSNRGAKAIVQGGGARARIVRDG
ITRAPCLRMNTAMEEAADLKIWCEEPQNFAILKQAFESTTSFGKLQSCNPTVAGKNVYLRLVCFSGDAMGMNMVSKGSLAVIEKLQE
HFPTCQLVALSGNMCTXKKAAATNWLHGRGKSVIVECIIPKEVVRTTLKTTVRALVHTNINKNLIGSAMAGAIGGFNAHASNIVTA
IFLATGQDPAQNVESSNCITLMEEQDNGDLWMCCTMPSIEVGTVGGGTSLPAQAACLEAIGVKGGGVVPGANAKKLATVVAAATMA
GELSLLAALAANTLVQAHMAHNRKPAAKK

>Pseudonitzschia_arenysensis
MAETMTPMPTIGMQLDAMIAQIDALPSTQLYGLIVAATVGLCVVVLGTGNSNLEMEQHDLKKPKMVPVDGGKQPNWQIFKVINYIA
VAAFLYSVFMFCSNASKYLHHESQGVLAQFLVGWSVFLMYFFGFFGVSLIHEDIPSEVDSVPSQPKSVAPTKPVHPPAACAPVCSD
PSSFKAKKSCVPENIKELGDEEIADLVLANKVKDHMLEKLLDPFRAVTVRRIACNRKLNSVHGNTSNVLDKLPSEHALDYSRVYGA
NCEIVVGYVPLPVGLVGPLTINDESFYVPMATTEGCLVASSNRGAKAICQGGGAKARIVRDGITRAPCLRMNSAMEEAADLKIWCEK
PANFAILKKAFESTTSFGKLIECNPTVAGKNVYLRLVCFSGDAMGMNMVSKGSLAVIEKLQEYFPSCQLVALSGNMCTDKKAAATN
WLHGRGKSVVVECIIPKEVVRTTLKTTVAALVHTNVNKNLIGSAMAGAIGGFNAHASNIVTAIFLATGQDPAQNVESSNCITLMEK
QDNGDLWMCCTMPSIEVGTVGGGTSLPAQAACLEAIGCKGGGATPGANAKQLATVVAAATMAGELSLLAALAANTLVQAHMAHNRK
PAAKK

>Fragilariopsis_kerguelensis
MASTEVMPTLGMRLDAMISHIDRLPSSQLYGVIVAATVGLCVVLLGTGNSNLELQQQQRQRQQQMMNVDDLKKPINNRIANVGGKQ
PKWHIFKWINYLAVGAFLWSVYTFCSNASQYLHHESQGVLVQFLVGWSVFLLYFFGFFGISLIHEDIPKEEDETTTMTVSKNIVSR
KPYNASTNSTTTMMKKKNSIHPPASCAPVCSDPTSFKSSGASLSSNSVNLEELADEEIANLVLNNKVKDHELEKRLDPFRAVTVRR
IAFNQKIASVLSDNTNKQNNINNTANVLDKLPSTPSLDYSRVYGANCEIVVGYVPLPVGLVGPLTINDETVYIPMATTEGCLVASS
NRGAKAITQGSGAKARIVRDGITRAPCIRMRSAMEAADLKLWCEEPSNFLILKQAFESTTNFGKLKECNPTVAGKNVYLRLVCFSG
DAMGMNMVSKGSLAVIETLQNEFPSLQLVALSGNMCTDKKAAATNWLNGRGKSVVVEAVIPREVVEKTLKTTVKALVHTNINKNLI
GSAMAGAIGGFNAHASNIVTAIFLATGQDPAQNVESSNCITLIEETDDNDLWISCTMPSIEVGTVGGGTSLEAQSACLEAIGCKGG
GATPGENARKLATVVAAATMAGELSLLAALAANTLVQAHMVHNRKPASNK

>Leptocylindrus_danicus
MAQTADPMTIGTRLDALIASGAEYLNNASQAQVCATLFVSSVAFSFALLNCGKGAHGSSLDPTLKPLYKNTMEPIKTKVKSPDGSR
EPRWYIFKMLNYTAVATFSTSVLHFILYSDVYMNNDAQMMKLMGAWTLFVLYFFGFFGVSFVDTDDHLEGSSPAEDEISEMTADQA
SPSAVVAVAKQMPVKAAAPIHPPAPSHPVCSDKNLTKDALTSSSATPAVKKSSSLSLDLQSMTNEELADLVLTNKMKDHQLETKLN
PTRAVPIRRLVFEKKLASLGHAKSLDELPYEHSLSYERVFGANCEIVVGYVPLPVGMVGPCTLNGESVYIPMATTEGCLVASTNRG
CKAITAGGGAVSTLLRDGITRAPCLRFESAAEAAALALWAQEPHNFAKLKAAFESTTSFGKLLSATPTVAGKNCYLRLKCFSGDAM
GMNMVSKGSLAVVDLLRQHFPTLKLVALSGNMCTDKKPAAINWIEGRGKSVVVEATIPKDIVRTVLKTTVKAIVDTNLQKNLIGSA
MSGSVGGYNAHASNIVTAVFLATGQDPAQNVESSNCITIMEETDDGDLWISCTMPSIEVGTVGGGTSLPAQAACLKAIGVKGGGDI
PGGNARKLAHVVAVATMAGELSLLAALAANTLVQAHMQHNRKPATPAKK

>Chaetoceros neogracile
METQIPTIGMRIDSLISYFDAIPNTSKAFVIICCLFSFFLLNSSPNPAATLNTSTYPGCDVTKAKTSNKGSSACNTVSKSSNCESE
PQPKWHILKLLNVIAVIGLLTSFFWFASNASTYINDSNALLKFMALWSGFLCYFFGFFGISFIDVEELERGSSEEEEKQRTNECQK
SLFVHKQLPVHPPASCTPVCSDNLKISTDKPSKPTTFTPSSNSTNTNQKTDIKSLSNSEIVSMVLSNKIKDHQLEKLLDPYRAVIV
RRATFDQKLSTLNQGDALNDLPYEHDLEYSRVFGANCETVIGYVPLPVGMVGPLTLNGQTVYFPMATTEGCLVASTNRGCKAISQG
SGASSTITRDGITRAPCIRMKSAKEAAELKIWCEIPQNFLTLKAAFESTTSFGKLQSAQVTVAGKNAYVRLCCFSGDAMGMNMVSK
GSLAVIDLLKEKFPTLDLLALSGNMCTDKKAAAINWIEGRGKSVVVESIIPKDVVHATLKTDVKKIVTVNINKNLIGSAMAGSIGG
FNAHASNIVSAVFLATGQDPAQNVESSNCITLMEETEEGDLWVSCTMPSIEVGTVGGGTGLPAQSAGLKIIGCKGGGEQPGANAKK

```
LAHVVASAVMAGGPCFFF
```

>Chaetoceros dichaeta
```
QHQIQNAPNSGMRSSVVNPPPPSLHAPARHTPVCTDPPPNPKPSPTTPLPPPESTSDQAWADMVLTNQVKDHQLEKILNPHRAVVVR
RLIFQSKLGSLATDESCTSPLDELPYQHSLDYTRVHGANCETVVGYIPLPVGMVGPLTLNNQTVYFPMATTEGCLVASTNRGCKAI
SQGSGAISTIVRDGITRAPCLKMRSAREACELKLWCELPSNFCKLKEAFESTTGFGKLQSATTMIAGRNAYVRLCCFSGDAMGMNM
VSKGTLAVVDLLRSIFPTLELVALSGNVCTDKKPSAINWIEGRGKSVVVEATIPKDVVHTTLKTEVKTIVRVNIDKNLIGSALAGS
IGGFNAHAANIVTALFLATGQDPAQNVESSNCMTLMEETDNGDLWISCTMPSIEVGTVGGGTGLPAQKACLKVIGCQGGGVKPGEN
SRQLAHVVASATMAGELSLMAALSANTLVQAHMQHNRKPATTSAVDSKQ*
```

>Chaetoceros affinis
```
MPMEDIQSLSNDEIASLVLTNKIKDHRLEKLTDPDRAVIIRRAVYSHKLATLPTSNSTTSNSPIDALPHLHSLNYSKVYGANCETV
IGYVPLPVGIVGPITLNNQSVYFPMATTEGCLVASTNRGCKAITEGGGAVSVITRDGITRAPCVRMESAKDAALLKLWCEDTTNQS
ENFVQLKKAFESTTSFGKLEAVNVSIAGRNAYIRLRCFSGDAMGMNMVSKGSLAVIDYLKTIFPSLVLVALSGNVCTDKKAAAMNW
IEGRGKSVVVEATIPRDVVVKTLKTTVKAMVSVNINKNLIGSALAGVVGGFNAHASNIVSAVFLATGQDPAQNVESSNCMTLMEET
DSGDLWISCTMPSIEVGTVGGGTGLAAQAACLGIVGCRGGGENPGDNAKQLAHVVAAAVMAGGVSLLGA
```

Chapter 5

General discussion, future perspectives and conclusions

5.1 General conclusions

5.1.1 Species-specific differences in sterol composition are related to environmental conditions and diversification of sterol biosynthesis

Diatoms produce a wide variety of sterol compounds which are unique to each species. Within this diversity lies a set of complex, highly evolved and divergent sterol metabolic pathways and therefore potential value for diatoms to be engineered for the production of high-value products. The details of these divergent pathways are not yet fully characterized.

Through this thesis, I demonstrated that: i) salinity and temperature conditions influence minor sterol levels in diatoms (**Chapter 2**), ii) three diatoms from distinct clades which share a unique core in their phytosterol biosynthesis pathway, and the diversity of sterol types appears to occur downstream of the conserved core (**Chapter 3**), iii) laboratory strains of diatoms can be transformed with metabolic transgenes in order to test the adaptability and re-engineering potential of their sterol metabolites (**Chapter 4**).

I obtained sterol profiles from twelve different diatoms from diverse environments, including four polar diatoms. The results obtained confirmed what has been previously reported: diatoms produce a wide diversity of sterol compounds, different diatom species possess a unique sterol profile in terms of sterol type and abundance, and more interestingly, each diatom shows a unique sterol profile that differs significantly between closely related species (Figs. 2.1 and 2.2, **Chapter 2**) (Rampen et al., 2010). Since sterols are involved in cellular processes that interact with environmental conditions, such as cellular signalling and membrane fluidity (Dufourc 2008), we hypothesized that the diversity of sterol content across diatom species is related to changes in the environment and growth conditions (**Chapter 2**). I found that changes in the relative contents of minor sterols, cholesterol, fucosterol, campesterol and isofucosterol, accompanied shifts in temperature and salinity in three commonly cultivated diatoms. I provide insight into the role that sterol diversity may play in the capabilities of diatoms to adapt and survive under changing environmental conditions. I found that diatoms augmented the levels of minor sterols with different physical effects on membrane cohesion, including branched sterols, suggesting their role in tuning membrane dynamics (**Chapter 2**). These results support a theory that diatoms possess mechanisms to adapt to changes in environmental conditions by modifying their metabolite profile, including sterol composition (Fig. 5.1). This may be similar to the regulation of phytosterols in higher plants, which regulates membrane cohesion to sustain the functional state of membranes during temperature shifts (Beck et al., 2007; Valitova et al., 2016).

Although I observed a response of minor sterol levels to changes in temperature and salinity, which explains the presence of different sterol types in a single species, the fact

that closely related species possess a significantly different sterol profile was not explained. In the rest of this thesis, efforts were focused on elucidating sterol metabolic pathway in diatoms that could provide insights into conservation or diversification of metabolic steps that lead to variations in end-point sterol compounds (**Chapter 3**). I performed comparative metabolic profiling and transcriptomics for three different diatom species in response to inhibitors of enzymes involved in sterol biosynthesis. Interestingly, we found that three divergent diatoms share a unique sterol biosynthesis core that relies on an alternative squalene epoxidase, and the cyclization of 2,3 epoxysqualene into cycloartenol by a conserved oxidosqualene cyclase (**Chapter 3**). The diversification of sterol compounds produced by each species appears to occur in downstream reactions, suggesting adaptive specialization in terminal synthesis pathways.

I not only found this unique sterol biosynthesis core, but provided new knowledge regarding cholesterol synthesis in diatoms. I detected intermediates that suggest a hypothetical route for cholesterol biosynthesis that occurs in parallel to conventional sterol biosynthesis (Fig. 3.5, **Chapter 3**). A new transcriptome for *C. muelleri* was also measured and assembled during this research, which contributes novel information about this important diatom species that is widely used in the aquaculture industry(Reis Batista et al., 2015).

Considering results from **Chapter 2 and 3**, we believe that enzymes participating in downstream reactions of sterol biosynthesis, such as C-24 alkyl transferases and sterol C-22 desaturase, respond to changes in environmental conditions (Fig. 5.1). Typical C-22 desaturated sterols such as brassicasterol have been found in *P. tricornutum*, consistent with the presence of sterol 22-desaturase enzymes, whereas the same desaturases have not been detected in the centric diatoms *T. pseudonana* and *C. muelleri* which consequently do not produce C-22 desaturated sterols (**Chapter 2**). Whether the presence or absence of these enzymes provide certain diatom species with advantages to survive in more fluctuating environments is not yet demonstrated.

In general, the current process for sterol analysis considerably limits the amount of experiments that can be conducted simultaneously. Sterol extraction and GC-MS analysis is time consuming, and requires dedicated use of laboratory resources. A high throughput technique that allows screening for sterol content in cultures of less than 100 mL would allow the evaluation of sterol content in diatoms growing in a wider range of environmental conditions, as well as screening of many more genetic targets. Currently, in the Climate Change Cluster (C3) efforts are being invested in the creation of a phenomics facility that allows this type of high throughput research.

5.1.2 Metabolic engineering of sterol metabolic pathway of diatoms

Once we established unknown steps in sterol synthesis for the diatoms *T. pseudonana*, *C. muelleri* and *P. tricornutum*, we attempted to modify final levels of sterol by genetic en-
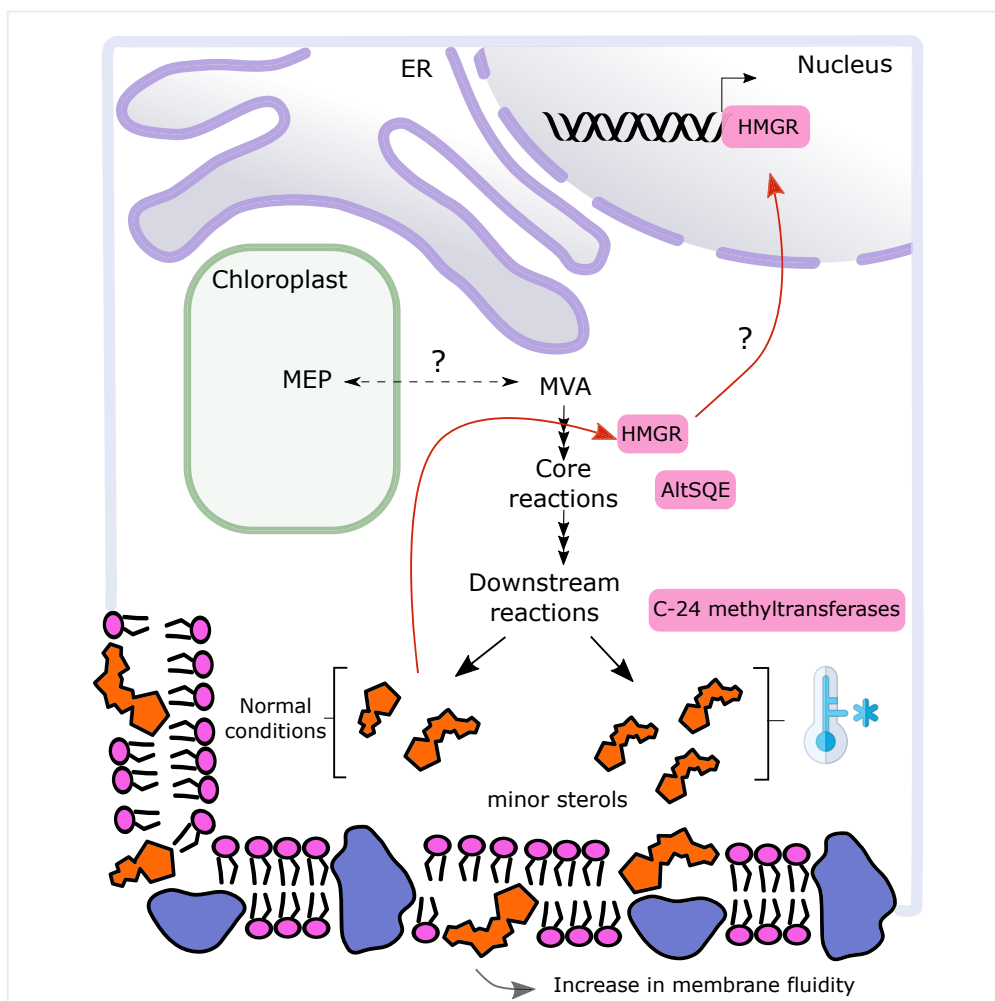
Figure 5.1: Hypothetical model of sterol pathway responses at reduced temperature in diatoms. When temperature decreases an increase of C-24 ethyl branched sterols such as campesterol could contribute to membrane fluidity maintenance (Chapter 2). Enzymes such as HMGR, AltSQE and C-24 methyltransferases are putatively involved in a regulation system (represented by red arrows) that maintains sterol levels and tunes minor sterol contents.

gineering. I consistently obtained transgenic lines of diatoms that were resistant to antibiotics and that expressed fluorescent fusion proteins consisting of target enzymes and mVenus. In **Chapter 4**, I independently over-expressed key enzymes in sterol biosynthesis: a native HMGR, a truncated version of HMGR (tHMGR) and a heterologous SQE from *Nannochloropsis oceanica*. Due to unavailable transformation methods and uncharacterized promoters in the species *C. muelleri*, genetic engineering experiments were carried out only in the model centric *T. pseudonana* and the model pennate diatom *P. tricornutum*.

I found statistically significant accumulation of intermediates from sterol metabolic pathways in *P. tricornutum* transformants. However, these diatoms did not appear to produce different levels of total sterols. It is presumed that several levels of regulation could be affecting the expression, localization, lifetime, and activities of these introduced genes.

The regulatory processes operating in diatoms regarding sterol homeostasis remains completely unknown.

Genetic engineering work done in **Chapter 4** is a demonstration of the utility and experimental accomplishment, with interesting findings that serve as examples of what is possible in diatoms as microbial production platforms. Additionally, in **Chapter 4** biological knowledge regarding localization and phylogeny in diatoms of the rate limiting enzyme HMGR is provided.

Through all the experiments conducted in this thesis, I consistently found that levels of the most abundant sterol in each diatom species tended to remain unchanged: cholesterol in *C. muelleri*, and brassicasterol in *P. tricornutum* and 24-methylcholesta-5,24(24')-dien-3β-ol in *T. pseudonana*. The content of these major sterols did not change after abrupt changes in environmental conditions (**Chapter 2**), under the treatment of some chemical inhibitors (fluconazole and fenpropimorph, **Chapter 3**) and even after over-expression of truncated HMGR and heterologous SQE enzymes (**Chapter 4**). These results suggest that diatoms are able to regulate sterol biosynthesis, through unknown and uncharacterized mechanisms. It is possible that the apparent ability of diatoms to rebalance sterol levels under different conditions, the presence of chemical inhibitors, and heterologously expressed enzymes is a consequence of additional levels of pre- or post-transcriptional or metabolic regulation.

The sequencing of the *T. pseudonana* and *P. tricornutum* genomes in 2004 and 2008, respectively (Armbrust et al., 2004; Bowler et al., 2008) was followed by a rapid increase in genetic engineering techniques in diatoms. However, we are still lacking efficient, reliable and reproducible methods that allow high throughput genetic experiments in diatoms. This includes characterization of new promoters for species without genomes available, such as *C. muelleri*, and the extension of transformation techniques, such as bacterial conjugation, to diatom species beyond the model pennate diatom *P. tricornutum*. Although genetic transformation of several diatom species can be achieved via microparticle bombardment, this technique often results in random and fragmented integration of multiple copies of the plasmid of interest in the genome (George et al. under review, 2019). Therefore, transformation of *P. tricornutum* via bacterial conjugation with an extrachromosomal episome capable of self-maintenance represents more reliable method (Diner et al., 2016; Karas et al., 2015). In this study, we optimised bacterial conjugation for consistent and reproducible transformation of *T. pseudonana*, based on previous studies (Karas et al., 2015). I consistently obtained genetic transformants for *P. tricornutum* and *T. pseudonana*.

Unfortunately, attempts to transform *C. muelleri* via bacterial conjugation were unsuccessful (data not shown). Additionally, I characterized the EF2 promoter in *T. pseudonana*, and implemented the use of blasticidin-resistant for antibiotic selection in *P. tricornutum* instead of the toxin traditionally used, zeocin. These genetic constructs were built using a modular type IIS restriction endonuclease-based system for efficient and versatile DNA fabrication referred to as "loop assembly" (Pollak et al., 2019). In this study, we largely

contribute to the creation of a domesticated library with genetic parts for genetic transformation of diatoms, including promoters, resistant genes and fluorescent tags.

5.2 Future perspectives

The knowledge provided in this thesis brings us one step closer to establishing diatoms as an alternative and sustainable source of phytosterols with nutraceutical applications. The adaptation and engineering of easily culturable diatoms into existing research and systems for industrial and sustainable cultivation can provide an alternative source of interesting phytosterols for natural bioproducts markets. Besides the advantages of diatoms as photosynthetic organisms, I learned that principal sterol levels remains stable under temperature and salinity shocks, while minor, less abundant sterols, appear to shift in order to acclimate to environmental change (**Chapter 2**). This capacity of regulation might represent an advantage when selecting a microalga species for large scale production in open ponds or raceway systems. Temperature, salinity and pH changes across the year are one of the most common reasons for decreases in production yield in large scale bioreactors for microalgae (Lebeau & Robert, 2003*a*).

Additionally, in this study, I establish that diatoms possess a unique sterol biosynthesis core that may represent an advantage for sterol and triterpenoids in general production compared to the conventional route present in most eukaryotic organisms (**Chapter 3**). Furthermore, establishing the sterol metabolic pathway is the first step to optimise sterol production through methods such as metabolic engineering. Flux balance analysis of sterol metabolism in diatoms can now be performed based on the results obtained in this study. Results obtained in **Chapter 4** encourage the use of genetic engineering methods for basic and applied research in diatoms. Now, we can design high throughput experiments that involve the fabrication of many genetic constructs and transformation via bacterial conjugation in diatoms.

5.2.1 Sterol derivatives and specialized reactions of sterol biosynthesis

Although salinity and temperature are some of the most important conditions affecting membrane fluidity and permeability, there are other factors such as growth medium composition that could significantly affect sterol levels in diatoms. Nitrogen and silicate limitation result on lipids accumulation in several diatom species (Chauton et al., 2013; Davis, Abbriano, Smith & Hildebrand, 2017; D'Ippolito et al., 2015; Traller & Hildebrand, 2013). Similarly, transcriptomics analysis revealed that several genes codifying for lipids production, including AltSQE, are highly responsive to phosphorus stress (Cruz de Carvalho et al., 2016; Fig. 5.2). Evaluation of those conditions and use of available transcriptomics data will provide insight on sterol levels and their response to environmental changes. More inter-
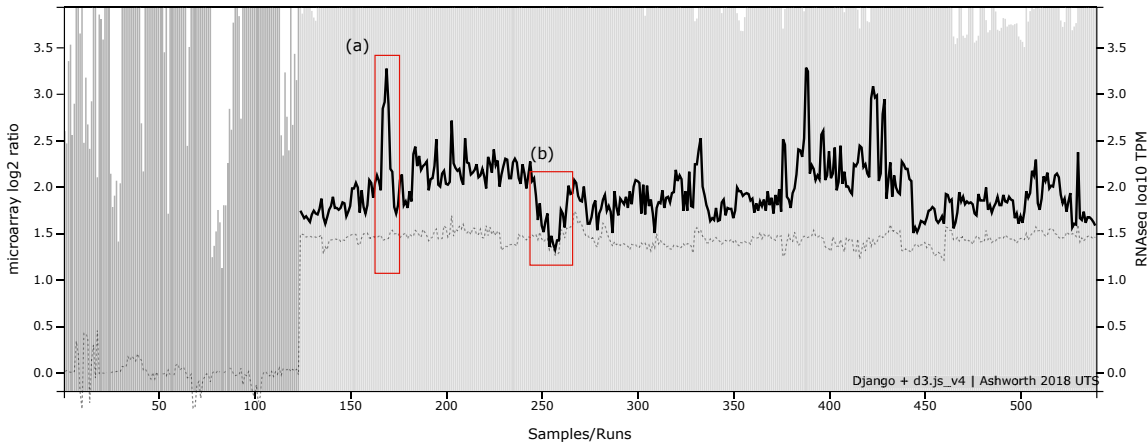
Figure 5.2: Expression profile of gene coding for AltSQE (Phatr3-J45494) in *P. tricornutum*. Source: https://alganaut.uts.edu.au/. Red squares represent some significant changes in transcript abundance (a) global transcriptome changes in response to phosphorus fluctuations over a course of 8 days (Cruz de Carvalho et al., 2016) (b) Responses under light intensity transitions (Heydarizadeh et al., 2017).

estingly would be the development of a method to determine and analyse the presence of conjugated sterols under different growth conditions. In this study, I was only able to identify free sterols due to the method used for sterol extraction. However, it has been reported that sterol sulphates participate in complex signalling processes such as cell death in diatom species that bloom (Gallo et al., 2017). Moreover, the presence in diatoms with rare sterols such as of 23-methyl sterols, cyclopropyl sterols and 27-norsterols is thought to provide a defense mechanism against grazers, potentially interfering with the sterol metabolism of predatory organisms (Giner & Wikfors, 2011). Therefore, exploration into other sterol derivatives and diatom species producing rare sterol compounds could provide insight on biological reasons for the presence of a huge diversity of sterol compounds in diatoms.

Specialized reactions of sterol biosynthesis in diatoms remain to be characterized. Particularly, we hypothesised that isofucosterol is converted into fucosterol in the diatoms *C. muelleri* and *T. pseudonana*. However, no enzyme candidate was proposed since this kind of isomerization has not been reported to be performed in any organism yet. Functional characterization of the final steps in sterol biosynthesis could provide new targets for genetic engineering and potential enzymes involved in sterol regulation. Additionally, the cholesterol pathway presented in **Chapter 3** was hypothetically reconstructed based on identified intermediates; however, further experiments are required to functionally characterize the proposed reactions.
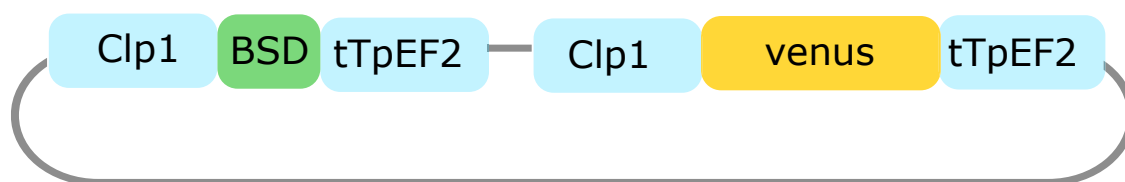
Figure 5.3: Schematic representation of proposed "universal" episome for transformation of different diatom species. Clp1 is a diatom-infecting virus promoter (Kadono et al., 2015) other genetic elements are described in Table 4.S1 (Chapter 4)

5.2.2 The unexplored sterol homeostasis mechanisms in diatoms

One fascinating aspect of this research is that all the results obtained point to the presence of a completely unknown sterol regulation system in diatoms. As mentioned before, I consistently observed the capacity of all diatoms studied to regulate sterol levels even under chemical inhibitors pressure. Transcriptomics data from **Chapter 3** revealed that key enzymes in sterol synthesis were significantly up or down-regulated in the presence of different inhibitors. Furthermore, sterol levels did not significantly change in *T. pseudonana* expressing only the catalytic domain of HMGR which in theory lacks regulatory domains (**Chapter 4**). These results suggest that diatoms have a tight sterol regulation system that may not be related to the conventional regulation model that involves SSD, but some complex system with several regulation points not only in the MVA pathway, but further down the sterol metabolic pathway. Therefore, metabolic engineering experiments targeting several enzymes simultaneously could enhance sterol production as discussed in **Chapter 4**.

Further exploration is required to characterize sterol homeostasis mechanism operating in diatoms. This could be done through targeted mutation of putative candidates involved in sterol regulation such as AltSQE and enzymes downstream the metabolic pathway. Generation of knock-down mutants using CRISPR/Cas9 approaches that have been proven in diatoms should be considered (Daboussi et al., 2014; Hopes et al., 2016; Nymark et al., 2016; Slattery et al., 2018; Weyman et al., 2015). One of the major bottlenecks with genetic engineering of diatoms that involves multi-species analysis (like the present study, **Chapter 4**) is the construction of multiple plasmids with promoters and resistance cassettes specific to each diatom species. As part of genetic engineering efforts done in this thesis, I built a "universal" episome for multiple diatoms transformation using Clp1, a diatom-infecting virus promoter (Kadono et al. 2015), as promoter and blasticidin (bsd) as selection marker (Fig. 5.3). This universal episome was successfully transformed in *T. pseudonana* and *P. tricornutum*, but the number of positive clones was too low to move forward (data not showed). Additional research in versatile promoters and resistance markers that can be used across diatom species will widen genetic engineering studies.

5.3 Concluding remarks

The demand for biologically-sourced sterol products in the pharmaceutical, nutraceutical and food markets is expanding, encouraging the pursuit of new sources of naturally produced sterols. Diatoms are broadly distributed and evolutionarily diversified microalgae that produce a wide range of potentially valuable phytosterol compounds. They possess unique features at the metabolic level that could provide advantages as production platforms of valuable compounds such as triterpenoids. Diatoms capacity to adapt and survive under changing environmental conditions is influenced by their genetic plasticity of metabolism, evident in the diversity of sterols that different species produce. Increasing understanding of diatom metabolism allows the application of genetic technologies for their manipulation to optimize and increase desired production yields.

References

Abumweis, S., Barake, R. & Jones, P. (2008), 'Plant sterols/stanols as cholesterol lowering agents: A meta-analysis of randomized controlled trials', *Food Nutrition Research* **52**(1811).
URL: *https://foodandnutritionresearch.net/index.php/fnr/article/view/1132*

Ahmed, F., Zhou, W. & Schenk, P. M. (2015), 'Pavlova lutheri is a high-level producer of phytosterols', *Algal Research* **10**, 210 – 217.
URL: *http://www.sciencedirect.com/science/article/pii/S2211926415001332*

Alappat, L., Valerio, M. & Awad, A. B. (2010), 'Effect of vitamin D and β-sitosterol on immune function of macrophages', *International Immunopharmacology* **10**(11), 1390 – 1396.
URL: *http://www.sciencedirect.com/science/article/pii/S1567576910002468*

Aldini, R., Micucci, M., Cevenini, M., Fato, R., Bergamini, C., Nanni, C., Cont, M., Camborata, C., Spinozzi, S., Montagnani, M., Roda, G., D'Errico-Grigioni, A., Rosini, F., Roda, A., Mazzella, G., Chiarini, A. & Budriesi, R. (2014), 'Antiinflammatory Effect of Phytosterols in Experimental Murine Colitis Model: Prevention, Induction, Remission Study', *PLOS ONE* **9**(9), 1–16.
URL: *https://doi.org/10.1371/journal.pone.0108112*

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997), 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic acids research* **25**(17), 3389–3402.

Alverson, A. J., Beszteri, B., Julius, M. L. & Theriot, E. C. (2011), 'The model marine diatom Thalassiosira pseudonana likely descended from a freshwater ancestor in the genus Cyclotella', *BMC evolutionary biology* **11**(1), 125.
URL: *https://doi.org/10.1186/1471-2148-11-125*

Andersen, R. A., Morton, S. L. & Sexton, J. P. (1997), 'Provasoli-Guillard National Center for culture of marine phytoplankton 1997 list of strains', *Journal of phycology* **33**, 1–75.
URL: *https://doi.org/10.1111/j.0022-3646.1997.00001.x*

Armbrust, E. V. (2009), 'The life of diatoms in the world's oceans', *Nature* **459**, 185–192.
URL: *https://doi.org/10.1038/nature08057*

Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., Zhou, S., Allen, A. E., Apt, K. E., Bechner, M., Brzezinski, M. A., Chaal, B. K., Chiovitti, A., Davis, A. K., Demarest,

M. S., Detter, J. C., Glavina, T., Goodstein, D., Hadi, M. Z., Hellsten, U., Hildebrand, M., Jenkins, B. D., Jurka, J., Kapitonov, V. V., Kröger, N., Lau, W. W. Y., Lane, T. W., Larimer, F. W., Lippmeier, J. C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M. S., Palenik, B., Pazour, G. J., Richardson, P. M., Rynearson, T. A., Saito, M. A., Schwartz, D. C., Thamatrakoln, K., Valentin, K., Vardi, A., Wilkerson, F. P. & Rokhsar, D. S. (2004), 'The Genome of the Diatom Thalassiosira Pseudonana: Ecology, Evolution, and Metabolism', *Science* **306**(5693), 79–86.
URL: *https://science.sciencemag.org/content/306/5693/79*

Athanasakoglou, A., Grypioti, E., Michailidou, S., Ignea, C., Makris, A. M., Kalantidis, K., Massé, G., Argiriou, A., Verret, F. & Kampranis, S. C. (2019), 'Isoprenoid biosynthesis in the diatom Haslea ostrearia', *New Phytologist* **222**(1), 230–243.
URL: *https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.15586*

Athyros, V., Kakafika, A., Papageorgiou, A., Tziomalos, K., Peletidou, A., Vosikis, C., Karagiannis, A. & Mikhailidis, D. (2011), 'Effect of a plant stanol ester-containing spread, placebo spread, or Mediterranean diet on estimated cardiovascular risk and lipid, inflammatory and haemostatic factors', *Nutrition, Metabolism and Cardiovascular Diseases* **21**, 213–221.

Ayache, N., Hervé, F., Lundholm, N., Amzil, Z. & Caruana, A. M. (2019), 'Acclimation of the marine diatom Pseudo-nitzschia australis to different salinity conditions: effects on growth, photosynthetic activity and domoic acid content', *Journal of phycology* **56**(1), 97–109.
URL: *https://doi.org/10.1111/jpy.12929*

Balzano, S., Sarno, D. & Kooistra, W. H. (2010), 'Effects of salinity on the growth rate and morphology of ten Skeletonema strains', *Journal of Plankton Research* **33**(6), 937–945.
URL: *https://doi.org/10.1093/plankt/fbq150*

Bansal, S., Narnoliya, L. K., Mishra, B., Chandra, M., Yadav, R. K. & Sangwan, N. S. (2018), 'HMG-CoA reductase from Camphor Tulsi (Ocimum kilimandscharicum) regulated MVA dependent biosynthesis of diverse terpenoids in homologous and heterologous plant systems', *Scientific reports* **8**(1), 3547.

Barka, F., Angstenberger, M., Ahrendt, T., Lorenzen, W., Bode, H. B. & Büchel, C. (2016), 'Identification of a triacylglycerol lipase in the diatom Phaeodactylum tricornutum', *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1861**(3), 239– 248.
URL: *http://www.sciencedirect.com/science/article/pii/S1388198115002450*

Barrett, S. M., Volkman, J. K., Dunstan, G. A. & LeRoi, J.-M. (1995), 'Sterols of 14 species of marine diatoms (Bacillariophyta)', *Journal of Phycology* **31**(3), 360–369.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0022-3646.1995.00360.x*

Beck, J. G., Mathieu, D., Loudet, C., Buchoux, S. & Dufourc, E. J. (2007), 'Plant sterols in "rafts": a better way to regulate membrane thermal shocks', *The FASEB Journal* **21**(8), 1714–1723. PMID: 17317727.
URL: *https://doi.org/10.1096/fj.06-7809com*

Benoiston, A.-S., Ibarbalz, F. M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S. & Bowler, C. (2017), 'The evolution of diatoms and their biogeochemical functions', *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**(1728), 20160397.

Beyene, A., Awoke, A. & Triest, L. (2014), 'Estimation of environmental optima and tolerances of diatoms using multifactor multiplicative modeling', *Ecological Informatics* **19**, 53–61.
URL: *http://www.sciencedirect.com/science/article/pii/S1574954113001283*

Bick, J. A. & Lange, B. M. (2003), 'Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane', *Archives of biochemistry and biophysics* **415**(2), 146–154.

Bolger, A. M., Lohse, M. & Usadel, B. (2014), 'Trimmomatic: a flexible trimmer for Illumina sequence data', *Bioinformatics* **30**(15), 2114–2120.

Borowitzka, M. A. (2013), 'High-value products from microalgae – their development and commercialisation', *Journal of Applied Phycology* **25**, 743–756.

Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R. P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J. A., Brownlee, C., Cadoret, J.-P., Chiovitti, A., Choi, C. J., Coesel, S., De Martino, A., Detter, J. C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman, M. J. J., Jenkins, B. D., Jiroutova, K., Jorgensen, R. E., Joubert, Y., Kaplan, A., Kröger, N., Kroth, P. G., La Roche, J., Lindquist, E., Lommer, M., Martin–Jézéquel, V., Lopez, P. J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L. K., Montsant, A., Secq, M.-P. O., Napoli, C., Obornik, M., Parker, M. S., Petit, J.-L. & Porcel, B. M. (2008), 'The Phaeodactylum genome reveals the evolutionary history of diatom genomes', *Nature* **456**, 239–244.

Bozarth, A., Maier, U. & Zauner, S. (2009), 'Diatoms in biotechnology: modern tools and applications', *Applied Microbiology and Biotechnology* **82**, 195–201.

Breteler, W., Schogt, N. & Rampen, S. (2005), 'Effect of diatom nutrient limitation on copepod development: role of essential lipids', *Marine Ecology Progress Series* **291**, 125–133.

Bröker, J. N., Müller, B., van Deenen, N., Prüfer, D. & Gronover, C. S. (2018), 'Upregulating the mevalonate pathway and repressing sterol synthesis in Saccharomyces cerevisiae enhances the production of triterpenes', *Applied microbiology and biotechnology* **102**(16), 6923–6934.

Brown, M. R., Dunstan, G. A., Norwood, S. J. & Miller, K. A. (1996), 'EFFECTS OF HARVEST STAGE AND LIGHT ON THE BIOCHEMICAL COMPOSITION OF THE DIATOM THALASSIOSIRA PSEUDONANA1', *Journal of Phycology* **32**(1), 64–73.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0022-3646.1996.00064.x*

Burg, J. S. & Espenshade, P. J. (2011), 'Regulation of HMG-CoA reductase in mammals and yeast', *Progress in Lipid Research* **50**(4), 403 – 410.
URL: *http://www.sciencedirect.com/science/article/pii/S0163782711000348*

Campagnac, E., Fontaine, J., Sahraoui, A. L.-H., Laruelle, F., Durand, R. & Grandmougin-Ferjani, A. (2009), 'Fenpropimorph slows down the sterol pathway and the development of the arbuscular mycorrhizal fungus Glomus intraradices', *Mycorrhiza* **19**(6), 365–374.

Cerqueira, N. M., Oliveira, E. F., Gesto, D. S., Santos-Martins, D., Moreira, C., Moorthy, H. N., Ramos, M. J. & Fernandes, P. (2016), 'Cholesterol biosynthesis: a mechanistic overview', *Biochemistry* **55**(39), 5483–5506.

Chan, C. X., Bhattacharya, D. & Reyes-Prieto, A. (2012), 'Endosymbiotic and horizontal gene transfer in microbial eukaryotes', *Mobile Genetic Elements* **2**(2), 101–105. PMID: 22934244.
URL: *https://doi.org/10.4161/mge.20110*

Chang, Y. C., Bien, C. M., Lee, H., Espenshade, P. J. & Kwon-Chung, K. J. (2007), 'Sre1p, a regulator of oxygen sensing and sterol homeostasis, is required for virulence in Cryptococcus neoformans', *Molecular Microbiology* **64**(3), 614–629.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2958.2007.05676.x*

Chappell, J., Wolf, F., Proulx, J., Cuellar, R. & Saunders, C. (1995), 'Is the reaction catalyzed by 3-hydroxy-3-methylglutaryl coenzyme A reductase a rate-limiting step for isoprenoid biosynthesis in plants?', *Plant Physiology* **109**(4), 1337–1343.

Chauton, M. S., Olsen, Y. & Vadstein, O. (2013), 'Biomass production from the microalga Phaeodactylum tricornutum: Nutrient stress and chemical composition in exponential fed-batch cultures', *Biomass and Bioenergy* **58**, 87– 94.
URL: *http://www.sciencedirect.com/science/article/pii/S0961953413004261*

Chen, C.-Y., Yeh, K.-L., Aisyah, R., Lee, D.-J. & Chang, J.-S. (2011), 'Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: A critical review', *Bioresource Technology* **102**(1), 71– 81. Special Issue: Biofuels - II: Algal Biofuels and Microbial Fuel Cells.
URL: *http://www.sciencedirect.com/science/article/pii/S0960852410011648*

Cheng, R.-l., Feng, J., Zhang, B.-X., Huang, Y., Cheng, J. & Zhang, C.-X. (2014), 'Transcriptome and gene expression analysis of an oleaginous diatom under different salinity conditions', *BioEnergy Research* **7**(1), 192–205.
URL: *https://doi.org/10.1007/s12155-013-9360-1*

Chisti, Y. (2007), 'Biodiesel from microalgae', *Biotechnology Advances* **25**(3), 294–306.
URL: *http://www.sciencedirect.com/science/article/pii/S0734975007000262*

Conte, M., Lupette, J., Seddiki, K., Meï, C., Dolch, L.-J., Gros, V., Barette, C., Rébeillé, F., Jouhet, J. & Marechal, E. (2018), 'Screening for biologically annotated drugs that trigger triacylglycerol accumulation in the diatom Phaeodactylum', *Plant physiology* **177**(2), 532–552.

Cruz de Carvalho, M. H., Sun, H.-X., Bowler, C. & Chua, N.-H. (2016), 'Noncoding and coding transcriptome responses of a marine diatom to phosphate fluctuations', *New Phytologist* **210**(2), 497–510.

Cvejić, J. H. & Rohmer, M. (2000), 'CO2 as main carbon source for isoprenoid biosynthesis via the mevalonate-independent methylerythritol 4-phosphate route in the marine diatoms Phaeodactylum tricornutum and Nitzschia ovalis', *Phytochemistry* **53**(1), 21–28.

Daboussi, F., Leduc, S., Marechal, A., Dubois, G., Perez-michaut, C., Amato, A., Falciatore, A., Juillerat, A., Beurdeley, M., Voytas, D., Cavarec, L. & Duchateau, P. (2014), 'Genome engineering empowers the diatom Phaeodactylum tricornutum for biotechnology', *Nature Communications* **5**(3831).

D'Adamo, S., Schiano di Visconte, G., Lowe, G., Szaub-Newton, J., Beacham, T., Landels, A., Allen, M. J., Spicer, A. & Matthijs, M. (2019), 'Engineering the unicellular alga Phaeodactylum tricornutum for high-value plant triterpenoid production', *Plant Biotechnology Journal* **17**(1), 75–87.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/pbi.12948*

Davis, A., Abbriano, R., Smith, S. R. & Hildebrand, M. (2017), 'Clarification of Photorespiratory Processes and the Role of Malic Enzyme in Diatoms', *Protist* **168**(1), 134 – 153.
URL: *http://www.sciencedirect.com/science/article/pii/S1434461016300669*

Davis, A., Crum, L. T., Corbeil, L. B. & Hildebrand, M. (2017), 'Expression of Histophilus somni IbpA DR2 protective antigen in the diatom Thalassiosira pseudonana', *Applied Microbiology and Biotechnology* **101**(13), 5313–5324.
URL: *https://doi.org/10.1007/s00253-017-8267-8*

de Jesus Raposo, M. F., de Morais, R. M. S. C. & de Morais, A. M. M. B. (2013), 'Health applications of bioactive compounds from marine microalgae', *Life Sciences* **93**(15), 479 – 486.
URL: *http://www.sciencedirect.com/science/article/pii/S0024320513004578*

Delgado-Vargas, F., Jiménez, A. & Paredes-López, O. (2000), 'Natural pigments: carotenoids, anthocyanins, and betalains—characteristics, biosynthesis, processing, and stability', *Critical reviews in food science and nutrition* **40**(3), 173–289.

Dellas, N., Thomas, S. T., Manning, G. & Noel, J. P. (2013), 'Discovery of a metabolic alternative to the classical mevalonate pathway', *Elife* **2**, e00672.

Desai, D. K., Nandi, S., Srivastava, P. K. & Lynn, A. M. (2011), 'ModEnzA: accurate identification of metabolic enzymes using function specific profile HMMs with optimised discrimination threshold and modified emission probabilities', *Advances in bioinformatics* **2011**.

Desmond, E. & Gribaldo, S. (2009), 'Phylogenomics of Sterol Synthesis: Insights into the Origin, Evolution, and Diversity of a Key Eukaryotic Feature', *Genome Biology and Evolution* **1**, 364–381.
URL: *https://doi.org/10.1093/gbe/evp036*

Diner, R. E., Bielinski, V. A., Dupont, C. L., Allen, A. E. & Weyman, P. D. (2016), 'Refinement of the diatom episome maintenance sequence and improvement of conjugation-based DNA delivery methods', *Frontiers in bioengineering and biotechnology* **4**, 65.

D'Ippolito, G., Sardo, A., Paris, D., Vella, F., Adelfi, M., Botte, P., Gallo, C. & Fontana, A. (2015), 'Potential of lipid metabolism in marine diatoms for biofuel production', *Biotechnology for Biofuels* **8**(28).

Doblas, V. G., Amorim-Silva, V., Posé, D., Rosado, A., Esteban, A., Arró, M., Azevedo, H., Bombarely, A., Borsani, O., Valpuesta, V., Ferrer, A., Tavares, R. M. & Botella, M. A. (2013), 'The SUD1 Gene Encodes a Putative E3 Ubiquitin Ligase and Is a Positive Regulator of 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase Activity in Arabidopsis', *The Plant Cell* **25**(2), 728–743.
URL: *http://www.plantcell.org/content/25/2/728*

Dodson, V. J., Mouget, J.-L., Dahmen, J. L. & Leblond, J. D. (2014), 'The long and short of it: temperature-dependent modifications of fatty acid chain length and unsaturation in the galactolipid profiles of the diatoms Haslea ostrearia and Phaeodactylum tricornutum', *Hydrobiologia* **727**(1), 95–107.
URL: *https://doi.org/10.1007/s10750-013-1790-4*

Donald, K., Hampton, R. Y. & Fritz, I. B. (1997), 'Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in Saccharomyces cerevisiae.', *Appl. Environ. Microbiol.* **63**(9), 3341–3344.

Dong, L., Pollier, J., Bassard, J.-E., Ntallas, G., Almeida, A., Lazaridi, E., Khakimov, B., Arendt, P., de Oliveira, L. S., Lota, F. et al. (2018), 'Co-expression of squalene epoxidases with triterpene cyclases boosts production of triterpenoids in plants and yeast', *Metabolic engineering* **49**, 1–12.

Dufourc, E. J. (2008), 'Sterols and membrane dynamics', *Journal of Chemical Biology* **1**(1), 63–77.
URL: *https://doi.org/10.1007/s12154-008-0010-6*

Eddy, S. R. (1998), 'Profile hidden Markov models.', *Bioinformatics (Oxford, England)* **14**(9), 755–763.

Eizadora, T., Zendejas, F., Lane, P., Gaucher, S., Simmons, B. & Lane, T. (2009), 'Triacylglycerol accumulation and profiling in the model diatoms Thalassiosira pseudonana and Phaeodactylum tricornutum (Baccilariophyceae) during starvation.', *Journal of Applied Phycology* **21**, 669–681.

Espenshade, P. J. & Hughes, A. L. (2007), 'Regulation of Sterol Synthesis in Eukaryotes', *Annual Review of Genetics* **41**(1), 401–427. PMID: 17666007.
URL: *https://doi.org/10.1146/annurev.genet.41.110306.130315*

Espín, J. C., García-Conesa, M. T. & Tomás-Barberán, F. A. (2007), 'Nutraceuticals: Facts and fiction', *Phytochemistry* **68**(22), 2986 – 3008. Highlights in the Evolution of Phytochemistry: 50 Years of the Phytochemical Society of Europe.
URL: *http://www.sciencedirect.com/science/article/pii/S0031942207005717*

Fabris, M., George, J., Kuzhiumparambil, U., Lawson, C. A., Jaramillo-Madrid, A. C., Abbriano, R. M., Vickers, C. E. & Ralph, P. (2020), 'Extrachromosomal genetic engineering of the marine diatom Phaeodactylum tricornutum enables the heterologous production of monoterpenoids', *ACS Synthetic Biology* **9**(3), 598–612.

Fabris, M., Matthijs, M., Carbonelle, S., Moses, T., Pollier, J., Dasseville, R., Baart, G. J. E., Vyverman, W. & Goossens, A. (2014), 'Tracking the sterol biosynthesis pathway of the diatom Phaeodactylum tricornutum', *New Phytologist* **204**(3), 521–535.
URL: *https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.12917*

Fabris, M., Matthijs, M., Rombauts, S., Vyverman, W., Goossens, A. & Baart, G. J. (2012), 'The metabolic blueprint of Phaeodactylum tricornutum reveals a eukaryotic Entner–Doudoroff glycolytic pathway', *The Plant Journal* **70**(6), 1004–1014.

Fahy, E., Subramaniam, S., Brown, H., Glass, C., Merrill, A., Murphy, R., Raetz, C., Russell, D., Seyama, Y., Shaw, W., Shimizu, T., Spener, F., van Meer, G., VanNieuwenhze, M., White, S., Witztum, J. & Dennis, E. (2005), 'A comprehensive classification system for lipids.', *Journal of Lipid Research* **46**, 839–861.

Fernandes, P. & Cabral, J. (2007), 'Phytosterols: Applications and recovery methods', *Bioresource Technology* **98**(12), 2335 – 2350.
URL: *http://www.sciencedirect.com/science/article/pii/S0960852406005244*

Ferriols, V. M. E. N., Yaginuma-Suzuki, R., Fukunaga, K., Kadono, T., Adachi, M., Matsunaga, S. & Okada, S. (2017), 'An exception among diatoms: unique organization of genes involved in isoprenoid biosynthesis in Rhizosolenia setigera CCMP 1694', *The Plant Journal* **92**(5), 822–833.

Friesen, J. A. & Rodwell, V. W. (2004), 'The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases', *Genome biology* **5**(11), 248.

Fu, W., Chaiboonchoe, A., Khraiwesh, B., Sultana, M., Jaiswal, A., Jijakli, K., Nelson, D. R., Al-Hrout, A., Baig, B., Amin, A. & Salehi-Ashtiani, K. (2017), 'Intracellular spectral recompositioning of light enhances algal photosynthetic efficiency', *Science Advances* **3**(9).
URL: *https://advances.sciencemag.org/content/3/9/e1603096*

Fu, W., Wichuk, K. & Brynjólfsson, S. (2015), 'Developing diatoms for value-added products: challenges and opportunities', *New Biotechnology* **32**(6), 547 – 551. European Congress of Biotechnology - ECB 16.
URL: *http://www.sciencedirect.com/science/article/pii/S1871678415000588*

Gabay, O., Sanchez, C., Salvat, C., Chevy, F., Breton, M., Nourissat, G., Wolf, C., Jacques, C. & Berenbaum, F. (2010), 'Stigmasterol: a phytosterol with potential anti-osteoarthritic properties', *Osteoarthritis and Cartilage* **18**, 106–116.

Gallo, C., d'Ippolito, G., Nuzzo, G., Sardo, A. & Fontana, A. (2017), 'Autoinhibitory sterol sulfates mediate programmed cell death in a bloom-forming marine diatom', *Nature communications* **8**(1), 1292.

Garaiová, M., Zambojová, V., Šimová, Z., Griač, P. & Hapala, I. (2014), 'Squalene epoxidase as a target for manipulation of squalene levels in the yeast Saccharomyces cerevisiae', *FEMS yeast research* **14**(2), 310–323.

George, J., Kahlke, T., Abbriano, R. M., Kuzhiumparambil, U., Ralph, P. J. & Fabris, M. (2020), 'Metabolic engineering strategies in diatoms reveal unique phenotypes and genetic configurations with implications for algal genetics and synthetic biology', *Frontiers in Bioengineering and Biotechnology* **8**, 513.

Gill, S., Stevenson, J., Kristiana, I. & Brown, A. J. (2011), 'Cholesterol-Dependent Degradation of Squalene Monooxygenase, a Control Point in Cholesterol Synthesis beyond HMG-CoA Reductase', *Cell Metabolism* **13**(3), 260–273.

Giner, J.-L. & Wikfors, G. H. (2011), '"Dinoflagellate Sterols" in marine diatoms', *Phytochemistry* **72**(14), 1896–1901.
  URL: *http://www.sciencedirect.com/science/article/pii/S0031942211002500*

Gladu, P. K., Patterson, G. W., Wikfors, G. H., Chitwood, D. J. & Lusby, W. (1991), 'Sterols of some diatoms', *Phytochemistry* **30**(7), 2301–2303.
  URL: *http://www.sciencedirect.com/science/article/pii/003194229183634W*

Gohil, N., Bhattacharjee, G., Khambhati, K., Braddick, D. & Singh, V. (2019), 'Engineering Strategies in Microorganisms for the Enhanced Production of Squalene: Advances, Challenges and Opportunities', *Frontiers in Bioengineering and Biotechnology* **7**.

Gold, D. A., Caron, A., Fournier, G. P. & Summons, R. E. (2017), 'Paleoproterozoic sterol biosynthesis and the rise of oxygen', *Nature* **543**(7645), 420.

Gong, Y., Hu, H., Gao, Y., Xu, X. & Gao, H. (2011), 'Microalgae as platforms for production of recombinant proteins and valuable compounds: progress and prospects', *Journal of Industrial Microbiology & Biotechnology* **38**(12), 1879–1890.

Gong, Y., Zhang, J., Guo, X., Wan, X., Liang, Z., Hu, C. J. & Jiang, M. (2013), 'Identification and characterization of PtDGAT2B, an acyltransferase of the DGAT2 acyl-Coenzyme A: Diacylglycerol acyltransferase family in the diatom Phaeodactylum tricornutum', *FEBS Letters* **587**(5), 481–487.
  URL: *https://febs.onlinelibrary.wiley.com/doi/abs/10.1016/j.febslet.2013.01.015*

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q. et al. (2011), 'Full-length transcriptome assembly from RNA-Seq data without a reference genome', *Nature biotechnology* **29**(7), 644.

Griffiths, M. & Harrison, S. (2009), 'Lipid productivity as a key characteristic for choosing algal species for biodiesel production', *Journal of Applied Phycology* **21**(5), 493–507.

Guillard, R. & Hargraves, P. (1993), 'Stichochrysis immobilis is a diatom, not a chrysophyte', *Phycologia* **32**(3), 234–236.

Hallikainen, M., Lyyra-Laitinen, T., Laitinen, T., Moilanen, L., Miettinen, T. A. & Gylling, H. (2008), 'Effects of plant stanol esters on serum cholesterol concentrations, relative markers of cholesterol metabolism and endothelial function in type 1 diabetes', *Atherosclerosis* **199**(2), 432–439.

Hamilton, M. L., Warwick, J., Terry, A., Allen, M. J., Napier, J. A. & Sayanova, O. (2015), 'Towards the Industrial Production of Omega-3 Long Chain Polyunsaturated Fatty Acids from a Genetically Modified Diatom Phaeodactylum tricornutum', *PLOS ONE* **10**(12), 1–15.
URL: *https://doi.org/10.1371/journal.pone.0144054*

Hannich, J., Umebayashi, K. & Riezman, H. (2011), 'Distribution and functions of sterols and sphingolipids.', *Cold Spring Harbor Perspectives in Biology* **3**(a004762).

Harding, T., Roger, A. J. & Simpson, A. G. B. (2017), 'Adaptations to High Salt in a Halophilic Protist: Differential Expression and Gene Acquisitions through Duplications and Gene Transfers', *Frontiers in Microbiology* **8**, 944.
URL: *https://www.frontiersin.org/article/10.3389/fmicb.2017.00944*

Harker, M., Holmberg, N., Clayton, J. C., Gibbard, C. L., Wallace, A. D., Rawlins, S., Hellyer, S. A., Lanot, A. & Safford, R. (2003), 'Enhancement of seed phytosterol levels by expression of an N-terminal truncated Hevea brasiliensis (rubber tree) 3-hydroxy-3-methylglutaryl-CoA reductase', *Plant biotechnology journal* **1**(2), 113–121.

Hemmerlin, A., Hoeffler, J.-F., Meyer, O., Tritsch, D., Kagan, I. A., Grosdemange-Billiard, C., Rohmer, M. & Bach, T. J. (2003), 'Cross-talk between the cytosolic mevalonate and the plastidial methylerythritol phosphate pathways in tobacco bright yellow-2 cells', *Journal of Biological Chemistry* **278**(29), 26666–26676.

Hempel, F. & Maier, U. G. (2012), 'An engineered diatom acting like a plasma cell secreting human IgG antibodies with high efficiency', *Microbial Cell Factories* **11**(126).

Henley, W. J., Litaker, R. W., Novoveská, L., Duke, C. S., Quemada, H. D. & Sayre, R. T. (2013), 'Initial risk assessment of genetically modified (GM) microalgae for commodity-scale biofuel cultivation', *Algal Research* **2**(1), 66–77.
URL: *http://www.sciencedirect.com/science/article/pii/S2211926412000549*

Heydarizadeh, P., Boureba, W., Zahedi, M., Huang, B., Moreau, B., Lukomska, E., Couzinet-Mossion, A., Wielgosz-Collin, G., Martin-Jézéquel, V., Bougaran, G. et al. (2017), 'Response of CO2-starved diatom Phaeodactylum tricornutum to light intensity transition', *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**(1728), 20160396.

Hildebrand, M., Davis, A. K., Smith, S. R., Traller, J. C. & Abbriano, R. (2012), 'The place of diatoms in the biofuels industry', *Biofuels* **3**(2), 221–240.
URL: *https://doi.org/10.4155/bfs.11.157*

Hopes, A., Nekrasov, V., Kamoun, S. & Mock, T. (2016), 'Editing of the urease gene by CRISPR-Cas in the diatom Thalassiosira pseudonana', *Plant Methods* **12**(49).

Howe, V., Sharpe, L. J., Alexopoulos, S. J., Kunze, S. V., Chua, N. K., Li, D. & Brown, A. J. (2016), 'Cholesterol homeostasis: How do cells sense sterol excess?', *Chemistry and Physics of Lipids*

199, 170–178. Properties and Functions of Cholesterol.
URL: *http://www.sciencedirect.com/science/article/pii/S0009308416300330*

Huang, Q., Jiang, F., Wang, L. & Yang, C. (2017), 'Design of Photobioreactors for Mass Cultivation of Photosynthetic Organisms', *Engineering* **3**(3), 318–329.
URL: *http://www.sciencedirect.com/science/article/pii/S2095809917304241*

Huang, W. & Daboussi, F. (2017), 'Genetic and metabolic engineering in diatoms', *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**(1728), 20160411.
URL: *https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2016.0411*

Hughes, A. L., Todd, B. L. & Espenshade, P. J. (2005), 'SREBP Pathway Responds to Sterols and Functions as an Oxygen Sensor in Fission Yeast', *Cell* **120**(6), 831–842.

Huntley, M. E., Johnson, Z. I., Brown, S. L., Sills, D. L., Gerber, L., Archibald, I., Machesky, S. C., Granados, J., Beal, C. & Greene, C. H. (2015), 'Demonstrated large-scale production of marine microalgae for fuels and feed', *Algal Research* **10**, 249–265.
URL: *http://www.sciencedirect.com/science/article/pii/S2211926415001113*

Izar, M. C., Tegani, D. M., Kasmas, S. H. & Fonseca, F. A. (2011), 'Phytosterols and phytosterolemia: gene–diet interactions', *Genes & Nutrition* **6**(1), 17–26.
URL: *https://doi.org/10.1007/s12263-010-0182-x*

Jaramillo-Madrid, A. C., Ashworth, J., Fabris, M. & Ralph, P. J. (2019), 'Phytosterol biosynthesis and production by diatoms (Bacillariophyceae)', *Phytochemistry* **163**, 46–57.

Jaramillo-Madrid, A. C., Ashworth, J., Fabris, M. & Ralph, P. J. (2020a), 'The unique sterol biosynthesis pathway of three model diatoms consists of a conserved core and diversified endpoints', *Algal Research* **48**, 101902.

Jaramillo-Madrid, A. C., Ashworth, J. & Ralph, P. J. (2020b), 'Levels of Diatom Minor Sterols Respond to Changes in Temperature and Salinity', *Journal of Marine Science and Engineering* **8**(2), 85.

Jiang, H. & Gao, K. (2004), 'Effects of lowering temperature during culture on the production of polyunsaturated fatty acids in the marine diatom Phaeodactylum tricornutum (bacillariophyceae) 1', *Journal of Phycology* **40**(4), 651–654.
URL: *https://doi.org/10.1111/j.1529-8817.2004.03112.x*

Johnson, L. K., Alexander, H. & Brown, C. T. (2019), 'Re-assembly, quality evaluation, and annotation of 678 microbial eukaryotic reference transcriptomes', *GigaScience* **8**(4), giy158.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G. et al. (2014), 'InterProScan 5: genome-scale protein function classification', *Bioinformatics* **30**(9), 1236–1240.

Jung, H. A., Islam, M. N., Lee, C. M., Oh, S. H., Lee, S., Jung, J. H. & Choi, J. S. (2013), 'Kinetics and molecular docking studies of an anti-diabetic complication inhibitor fucosterol from edible brown algae Eisenia bicyclis and Ecklonia stolonifera', *Chemico-Biological Interactions* **206**(1), 55–62.
URL: *http://www.sciencedirect.com/science/article/pii/S0009279713002214*

Kadono, T., Miyagawa-Yamaguchi, A., Kira, N., Tomaru, Y., Okami, T., Yoshimatsu, T., Hou, L., Ohama, T., Fukunaga, K., Okauchi, M. et al. (2015), 'Characterization of marine diatom-infecting virus promoters in the model diatom Phaeodactylum tricornutum', *Scientific reports* **5**, 18708.

Kajikawa, M., Kinohira, S., Ando, A., Shimoyama, M., Kato, M. & Fukuzawa, H. (2015), 'Accumulation of squalene in a microalga Chlamydomonas reinhardtii by genetic modification of squalene synthase and squalene epoxidase genes', *PLoS One* **10**(3), e0120446.

Kanehisa, M. & Goto, S. (2000), 'KEGG: kyoto encyclopedia of genes and genomes', *Nucleic acids research* **28**(1), 27–30.

Karas, B. J., Diner, R. E., Lefebvre, S. C., McQuaid, J., Phillips, A. P., Noddings, C. M., Brunson, J. K., Valas, R. E., Deerinck, T. J., Jablanovic, J. et al. (2015), 'Designer diatom episomes delivered by bacterial conjugation', *Nature communications* **6**, 6925.

Katan, M. B., Grundy, S. M., Jones, P., Law, M., Miettinen, T. & Paoletti, R. (2003), 'Efficacy and Safety of Plant Stanols and Sterols in the Management of Blood Cholesterol Levels', *Mayo Clinic Proceedings* **78**(8), 965–978.

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., Armbrust, E. V., Archibald, J. M., Bharti, A. K., Bell, C. J., Beszteri, B., Bidle, K. D., Cameron, C. T., Campbell, L., Caron, D. A., Cattolico, R. A., Collier, J. L., Coyne, K., Davy, S. K., Deschamps, P., Dyhrman, S. T., Edvardsen, B., Gates, R. D., Gobler, C. J., Greenwood, S. J., Guida, S. M., Jacobi, J. L., Jakobsen, K. S., James, E. R., Jenkins, B., John, U., Johnson, M. D., Juhl, A. R., Kamp, A., Katz, L. A., Kiene, R., Kudryavtsev, A., Leander, B. S., Lin, S., Lovejoy, C., Lynn, D., Marchetti, A., McManus, G., Nedelcu, A. M., Menden-Deuer, S., Miceli, C., Mock, T., Montresor, M., Moran, M. A., Murray, S., Nadathur, G., Nagai, S., Ngam, P. B., Palenik, B., Pawlowski, J., Petroni, G., Piganeau, G., Posewitz, M. C., Rengefors, K., Romano, G., Rumpho, M. E., Rynearson, T., Schilling, K. B., Schroeder, D. C., Simpson, A. G. B., Slamovits, C. H., Smith, D. R., Smith, G. J., Smith, S. R., Sosik, H. M., Stief, P., Theriot, E., Twary, S. N., Umale, P. E., Vaulot, D., Wawrik, B., Wheeler, G. L., Wilson, W. H., Xu, Y., Zingone, A. & Worden, A. Z. (2014), 'The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing', *PLOS Biology* **12**(6), 1–6.
URL: *https://doi.org/10.1371/journal.pbio.1001889*

Kim, D., Langmead, B. & Salzberg, S. L. (2015), 'HISAT: a fast spliced aligner with low memory requirements', *Nature methods* **12**(4), 357.

Kim., J., Yoo., G., Lee., H., Lim., J., Kim., K., Kim., C. W., Park., M. S. & Yang, J.-W. (2013), 'Methods of downstream processing for the production of biodiesel from microalgae', *Biotechnology Advances* **31**(6), 862 – 876. "Bioenergy and Biorefinery from Biomass" through innovative technology development.
URL: *http://www.sciencedirect.com/science/article/pii/S0734975013000773*

Kim, Y.-J., Lee, O. R., Oh, J. Y., Jang, M.-G. & Yang, D.-C. (2014), 'Functional analysis of 3-hydroxy-3-methylglutaryl coenzyme a reductase encoding genes in triterpene saponin-producing ginseng', *Plant physiology* **165**(1), 373–387.

Kirby, J. & Keasling, J. D. (2009), 'Biosynthesis of plant isoprenoids: perspectives for microbial engineering', *Annual review of plant biology* **60**, 335–355.

Klug, L. & Daum, G. (2014), 'Yeast lipid metabolism at a glance', *FEMS yeast research* **14**(3), 369–388.

Kritchevsky, D. & Chen, S. C. (2005), 'Phytosterols—health benefits and potential concerns: a review', *Nutrition Research* **25**(5), 413 – 428.
URL: *http://www.sciencedirect.com/science/article/pii/S0271531705000837*

Krogh, A., Larsson, B., Von Heijne, G. & Sonnhammer, E. L. (2001), 'Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes', *Journal of molecular biology* **305**(3), 567–580.

Kroth, P. G. (2002), Protein transport into secondary plastids and the evolution of primary and secondary plastids, *in* 'International review of cytology', Vol. 221, Elsevier, pp. 191–255.

Kuwabara, P. E. & Labouesse, M. (2002), 'The sterol-sensing domain: multiple families, a unique role?', *TRENDS in Genetics* **18**(4), 193–201.

Lange, B. M., Rujan, T., Martin, W. & Croteau, R. (2000), 'Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes', *Proceedings of the National Academy of Sciences* **97**(24), 13172–13177.

Lange, I., Poirier, B. C., Herron, B. K. & Lange, B. M. (2015), 'Comprehensive assessment of transcriptional regulation facilitates metabolic engineering of isoprenoid accumulation in Arabidopsis', *Plant physiology* **169**(3), 1595–1606.

Laranjeira, S., Amorim-Silva, V., Esteban, A., Arró, M., Ferrer, A., Tavares, R. M., Botella, M. A., Rosado, A. & Azevedo, H. (2015), 'Arabidopsis squalene epoxidase 3 (SQE3) complements SQE1 and is important for embryo development and bulk squalene epoxidase activity', *Molecular plant* **8**(7), 1090–1102.

Lebeau, T. & Robert, J. (2003*a*), 'Diatom cultivation and biotechnologically relevant products. Part I: Cultivation at various scales', *Applied Microbiology and Biotechnology* **60**(6), 612–623.

Lebeau, T. & Robert, J. M. (2003*b*), 'Diatom cultivation and biotechnologically relevant products. Part II: Cultivation at various scales', *Applied Microbiology and Biotechnology* **60**(6), 624–632.

Leber, R., Landl, K., Zinser, E., Ahorn, H., Spok, A., Kohlwein, S. D., Turnowsky, F. & Daum, G. (1998), 'Dual localization of squalene epoxidase, Erg1p, in yeast reflects a relationship between the endoplasmic reticulum and lipid particles', *Molecular biology of the cell* **9**(2), 375–386.

Lee, A.-R., Kwon, M., Kang, M.-K., Kim, J., Kim, S.-U. & Ro, D.-K. (2019), 'Increased sesqui-and triterpene production by co-expression of HMG-CoA reductase and biotin carboxyl carrier protein in tobacco (Nicotiana benthamiana)', *Metabolic engineering* **52**, 20–28.

Lee., Y. S., Shin., K. H., Kim., B.-K. & Lee, S. (2004), 'Anti-Diabetic activities of fucosterol from Pelvetia siliquosa', *Archives of Pharmacal Research* **27**(11), 1120–1122.

Leivar, P., Antolín-Llovera, M., Ferrero, S., Closa, M., Arró, M., Ferrer, A., Boronat, A. & Campos, N. (2011), 'Multilevel Control of Arabidopsis 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase by Protein Phosphatase 2A', *The Plant Cell* **23**(4), 1494–1511.
URL: *http://www.plantcell.org/content/23/4/1494*

Leivar, P., González, V. M., Castel, S., Trelease, R. N., López-Iglesias, C., Arró, M., Boronat, A., Campos, N., Ferrer, A. & Fernandez-Busquets, X. (2005), 'Subcellular localization of Arabidopsis 3-hydroxy-3-methylglutaryl-coenzyme A reductase', *Plant physiology* **137**(1), 57–69.

Levitan, O., Dinamarca, J., Zelzion, E., Gorbunov, M. Y. & Falkowski, P. G. (2015), 'An RNA interference knock-down of nitrate reductase enhances lipid biosynthesis in the diatom Phaeodactylum tricornutum', *The Plant Journal* **84**(5), 963–973.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/tpj.13052*

Li, W., Liu, W., Wei, H., He, Q., Chen, J., Zhang, B. & Zhu, S. (2014), 'Species-specific expansion and molecular evolution of the 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) gene family in plants', *PloS one* **9**(4), e94172.

Liang, Y., Sun, M., Tian, C., Cao, C. & Li, Z. (2014), 'Effects of salinity stress on the growth and chlorophyll fluorescence of Phaeodactylum tricornutum and Chaetoceros gracilis (Bacillariophyceae)', *Botanica Marina* **57**(6), 469–476.
URL: *https://doi.org/10.1515/bot-2014-0037*

Liao, P., Wang, H., Wang, M., Hsiao, A.-S., Bach, T. J. & Chye, M.-L. (2014), 'Transgenic Tobacco Overexpressing Brassica juncea HMG-CoA Synthase 1 Shows Increased Plant Growth, Pod Size and Seed Yield', *PLOS ONE* **9**(5), 1–15.
URL: *https://doi.org/10.1371/journal.pone.0098264*

Lichtenthaler, H. K. (2000), 'Non-mevalonate isoprenoid biosynthesis: enzymes, genes and inhibitors', *Biochemical Society Transactions* **28**(6), 785–789.
URL: *https://doi.org/10.1042/bst0280785*

Lin, Y., Knol, D. & Trautwein, E. A. (2016), 'Phytosterol oxidation products (POP) in foods with added phytosterols and estimation of their daily intake: A literature review', *European Journal of*

Lipid Science and Technology **118**(10), 1423–1438.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/ejlt.201500368*

Liu, X., Hempel, F., Stork, S., Bolte, K., Moog, D., Heimerl, T., Maier, U. G. & Zauner, S. (2016), 'Addressing various compartments of the diatom model organism Phaeodactylum tricornutum via sub-cellular marker proteins', *Algal research* **20**, 249–257.

Lohr, M., Schwender, J. & Polle, J. E. (2012), 'Isoprenoid biosynthesis in eukaryotic phototrophs: A spotlight on algae', *Plant Science* **185-186**, 9 – 22.
URL: *http://www.sciencedirect.com/science/article/pii/S0168945211002202*

Lu, Y., Zhou, W., Wei, L., Li, J., Jia, J., Li, F., Smith, S. M. & Xu, J. (2014), 'Regulation of the cholesterol biosynthetic pathway and its integration with fatty acid biosynthesis in the oleaginous microalga Nannochloropsis oceanica', *Biotechnology for biofuels* **7**(1), 81.

Luo, X., Su, P. & Zhang, W. (2015), 'Advances in Microalgae-Derived Phytosterols for Functional Food and Pharmaceutical Applications', *Marine Drugs* **13**(7), 4231–4254.

López-Elías, J. A., Voltolina, D., Enríquez-Ocaña, F. & Gallegos-Simental, G. (2005), 'Indoor and outdoor mass production of the diatom Chaetoceros muelleri in a mexican commercial hatchery', *Aquacultural Engineering* **33**(3), 181 – 191.
URL: *http://www.sciencedirect.com/science/article/pii/S0144860905000051*

Ma, Y.-H., Wang, X., Niu, Y.-F., Yang, Z.-K., Zhang, M.-H., Wang, Z.-M., Yang, W.-D., Liu, J.-S. & Li, H.-Y. (2014), 'Antisense knockdown of pyruvate dehydrogenase kinase promotes the neutral lipid accumulation in the diatom Phaeodactylum tricornutum', *Microbial Cell Factories* **13**(100).

Malik, A., Lenzen, M., Ralph, P. J. & Tamburic, B. (2015), 'Hybrid life-cycle assessment of algal biofuel production', *Bioresource Technology* **184**, 436 – 443. Advances in biofuels and chemicals from algae.
URL: *http://www.sciencedirect.com/science/article/pii/S0960852414015648*

Mann, D. G. & Vanormelingen, P. (2013), 'An Inordinate Fondness? The Number, Distributions, and Origins of Diatom Species', *Journal of Eukaryotic Microbiology* **60**(4), 414–420.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/jeu.12047*

Massé, G., Belt, S. T., Rowland, S. J. & Rohmer, M. (2004), 'Isoprenoid biosynthesis in the diatoms Rhizosolenia setigera (Brightwell) and Haslea ostrearia (Simonsen)', *Proceedings of the National Academy of Sciences* **101**(13), 4413–4418.
URL: *https://www.pnas.org/content/101/13/4413*

Matsumoto, M., Nojima, D. ans Nonoyama, T., Ikeda, K., Maeda, Y., Yoshino, T. & Tanaka, T. (2017), 'Outdoor Cultivation of Marine Diatoms for Year-Round Production of Biofuels', *Marine Drugs* **15**(94).

Medlin, L. K. & Kaczmarska, I. (2004), 'Evolution of the diatoms: V. Morphological and cytological support for the major clades and a taxonomic revision', *Phycologia* **43**(3), 245–270.

Melis, A. (2009), 'Solar energy conversion efficiencies in photosynthesis: Minimizing the chlorophyll antennae to maximize efficiency', *Plant Science* **177**(4), 272 – 280.
URL: *http://www.sciencedirect.com/science/article/pii/S0168945209001861*

Miras-Moreno, B., Sabater-Jara, A. B., Pedreño, M. A. & Almagro, L. (2016), 'Bioactivity of Phytosterols and Their Production in Plant in Vitro Cultures', *Journal of Agricultural and Food Chemistry* **64**(38), 7049–7058.

Mitchell, J. G., Seuront, L., Doubell, M. J., Losic, D., Voelcker, N. H., Seymour, J. & Lal, R. (2013), 'The Role of Diatom Nanostructures in Biasing Diffusion to Improve Uptake in a Patchy Nutrient Environment', *PLOS ONE* **8**(5), 1–9.
URL: *https://doi.org/10.1371/journal.pone.0059548*

Morand, O., Aebi, J., Dehmlow, H., Ji, Y.-H., Gains, N., Lengsfeld, H. & Himber, J. (1997), 'Ro 48-8.071, a new 2, 3-oxidosqualene: lanosterol cyclase inhibitor lowering plasma cholesterol in hamsters, squirrel monkeys, and minipigs: comparison to simvastatin.', *Journal of lipid research* **38**(2), 373–390.

Moreau, R. A., Whitaker, B. D. & Hicks, K. B. (2002), 'Phytosterols, phytostanols, and their conjugates in foods: structural diversity, quantitative analysis, and health-promoting uses', *Progress in Lipid Research* **41**(6), 457 – 500.
URL: *http://www.sciencedirect.com/science/article/pii/S0163782702000061*

Moss, G. P. (1989), 'The nomenclature of steroids', *European Journal of Biochemistry* **186**(3), 429–458.
URL: *https://febs.onlinelibrary.wiley.com/doi/abs/10.1111/j.1432-1033.1989.tb15228.x*

Moustafa, A., Beszteri, B., Maier, U. G., Bowler, C., Valentin, K. & Bhattacharya, D. (2009), 'Genomic Footprints of a Cryptic Plastid Endosymbiosis in Diatoms', *Science* **324**(5935), 1724–1726.
URL: *https://science.sciencemag.org/content/324/5935/1724*

Nagai, M., Sakakibara, J., Nakamura, Y., Gejyo, F. & Ono, T. (2002b), 'SREBP-2 and NF-Y are involved in the transcriptional regulation of squalene epoxidase', *Biochemical and biophysical research communications* **295**(1), 74–80.

Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K. & Miyawaki, A. (2002a), 'A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications', *Nature biotechnology* **20**(1), 87–90.

Nes, W. D. (2011), 'Biosynthesis of cholesterol and other sterols', *Chemical reviews* **111**(10), 6423–6451.

Niu, Y.-F., Wang, X., Hu, D.-X., Balamurugan, S., Li, D.-W., Yang, W.-D., Liu, J.-S. & Li, H.-Y. (2016), 'Molecular characterization of a glycerol-3-phosphate acyltransferase reveals key features essential for triacylglycerol production in Phaeodactylum tricornutum', *Biotechnology for Biofuels* **9**(60).

Noakes, M., Clifton, P., Ntanios, F., Shrapnel, W., Record, I. & McInerney, J. (2002), 'An increase in dietary carotenoids when consuming plant sterols or stanols is effective in maintaining plasma carotenoid concentrations', *The American Journal of Clinical Nutrition* **75**(1), 79–86.
URL: *https://doi.org/10.1093/ajcn/75.1.79*

Nowosielski, M., Hoffmann, M., Wyrwicz, L. S., Stepniak, P., Plewczynski, D. M., Lazniewski, M., Ginalski, K. & Rychlewski, L. (2011), 'Detailed mechanism of squalene epoxidase inhibition by terbinafine', *Journal of chemical information and modeling* **51**(2), 455–462.

Nymark, M., Sharma, A. K., Sparstad, T., Bones, A. M. & Winge, P. (2016), 'A CRISPR/Cas9 system adapted for gene editing in marine algae', *Scientific Reports* **6**(24951).

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. (2016), 'Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown', *Nature protocols* **11**(9), 1650.

Piepho, M., Martin-Creuzburg, D. & Wacker, A. (2011), 'Simultaneous Effects of Light Intensity and Phosphorus Supply on the Sterol Content of Phytoplankton', *PLOS ONE* **5**(12), 1–7.
URL: *https://doi.org/10.1371/journal.pone.0015828*

Piepho, M., Martin-Creuzburg, D. & Wacker, A. (2012), 'Phytoplankton sterol contents vary with temperature, phosphorus and silicate supply: a study on three freshwater species', *European Journal of Phycology* **47**(2), 138–145.
URL: *https://doi.org/10.1080/09670262.2012.665484*

Piironen, V., Lindsay, D. G., Miettinen, T. A., Toivo, J. & Lampi, A.-M. (2000), 'Plant sterols: biosynthesis, biological function and their importance to human nutrition', *Journal of the Science of Food and Agriculture* **80**(7), 939–966.

Piironen, V., Toivo, J., Puupponen-Pimiä, R. & Lampi, A.-M. (2003), 'Plant sterols in vegetables, fruits and berries', *Journal of the Science of Food and Agriculture* **83**(4), 330–337.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/jsfa.1316*

Polakowski, T., Stahl, U. & Lang, C. (1998), 'Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast', *Applied microbiology and biotechnology* **49**(1), 66–71.

Pollak, B., Cerda, A., Delmans, M., Álamos, S., Moyano, T., West, A., Gutiérrez, R. A., Patron, N. J., Federici, F. & Haseloff, J. (2019), 'Loop Assembly: a simple and open system for recursive fabrication of DNA circuits', *New Phytologist* **222**(1), 628–640.

Pollak, O. J. (1953), 'Reduction of Blood Cholesterol in Man', *Circulation* **7**(5), 702–706.
URL: *https://www.ahajournals.org/doi/abs/10.1161/01.CIR.7.5.702*

Pollier, J., Moses, T., Gonzalez-Guzman, M., De Geyter, N., Lippens, S., Bossche, R., Marhavy, P., Kremer, A., Morreel, K., Guerin, C., Tava, A., Oleszek, W., Thevelein, J., Campos, N., Goormachtig, S. & Goossens, A. (2013), 'The protein quality control system manages plant defence compound synthesis', *Nature* **504**, 148–152.

Pollier, J., Vancaester, E., Kuzhiumparambil, U., Vickers, C. E., Vandepoele, K., Goossens, A. & Fabris, M. (2019), 'A widespread alternative squalene epoxidase participates in eukaryote steroid biosynthesis', *Nature Microbiology* **4**, 226–233.

Ponomarenko, L., Stonik, I., Aizdaicher, N., Orlova, T., Popovskaya, G., Pomazkina, G. & Stonik, V. (2004), 'Sterols of marine microalgae Pyramimonas cf. cordata (Prasinophyta), Attheya ussurensis sp. nov. (Bacillariophyta) and a spring diatom bloom from Lake Baikal', *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **138**(1), 65 – 70.
URL: *http://www.sciencedirect.com/science/article/pii/S1096495904000387*

Párraga-Martínez, I., López-Torres-Hidalgo, J. D., del Campo-del Campo, J. M., Galdón-Blesa, M. P., Precioso-Yáñez, J. C., Rabanales-Sotos, J., García-Reyes-Ramos, M., Andrés-Pretel, F., Navarro-Bravo, B. & Ángeles Lloret-Callejo (2015), 'Long-term Effects of Plant Stanols on the Lipid Profile of Patients With Hypercholesterolemia. A Randomized Clinical Trial', *Revista Española de Cardiología (English Edition)* **68**(8), 665 – 671.
URL: *http://www.sciencedirect.com/science/article/pii/S1885585714004150*

Rampen, S. W., Abbas, B. A., Schouten, S. & Sinninghe Damste, J. S. (2010), 'A comprehensive study of sterols in marine diatoms (Bacillariophyta): Implications for their use as tracers for diatom productivity', *Limnology and Oceanography* **55**(1), 91–105.
URL: *https://aslopubs.onlinelibrary.wiley.com/doi/abs/10.4319/lo.2010.55.1.0091*

Ras, R. T., Geleijnse, J. M. & Trautwein, E. A. (2014), 'LDL-cholesterol-lowering effect of plant sterols and stanols across different dose ranges: a meta-analysis of randomised controlled studies', *British Journal of Nutrition* **112**(2), 214–219.

Rastogi, A., Maheswari, U., Dorrell, R. G., Vieira, F. R. J., Maumus, F., Kustka, A., McCarthy, J., Allen, A. E., Kersey, P., Bowler, C. & Tirichine, L. (2018), 'Integrative analysis of large scale transcriptome data draws a comprehensive landscape of Phaeodactylum tricornutum genome and evolutionary origin of diatoms', *Scientific Reports* **8**(4834).

Re, E. B., Jones, D. & Learned, R. M. (1995), 'Co-expression of native and introduced genes reveals cryptic regulation of HMG CoA reductase expression in Arabidopsis', *The Plant Journal* **7**(5), 771–784.

Reis Batista, I., Garcia, A. B., Van Dalen, P., Kamermans, P., Verdegem, M. & Smaal, A. C. (2015), 'Culturing Chaetoceros muelleri using simplified media with different N sources: effects on production and lipid content', *European journal of phycology* **50**(1), 92–99.

Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R. & Keasling, J. D. (2006), 'Production of the antimalarial drug precursor artemisinic acid in engineered yeast', *Nature* **440**, 940–943.

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. (2010), 'edgeR: a Bioconductor package for differential expression analysis of digital gene expression data', *Bioinformatics* **26**(1), 139–140.

Rodríguez-Concepción, M. (2006), 'Early steps in isoprenoid biosynthesis: multilevel regulation of the supply of common precursors in plant cells', *Phytochemistry Reviews* **5**(1), 1–15.

Rosenfeld, J. & Osborne, T. (1998), 'HLH106, a Drosophila sterol regulatory element-binding protein in a natural cholesterol auxotroph', *Journal of Biological Chemistry* **273**(26), 16112–16121.

Round, F., Crawford, R. & Mann (1990), *The Diatoms: Biology and Morphology of the Genera.*, Cambridge University Press, England.

Russell, N. J. (1989), 'Adaptive modifications in membranes of halotolerant and halophilic microorganisms', *Journal of Bioenergetics and Biomembranes* **21**(1), 93–113.

Sáenz, J. P., Grosser, D., Bradley, A. S., Lagny, T. J., Lavrynenko, O., Broda, M. & Simons, K. (2015), 'Hopanoids as functional analogues of cholesterol in bacterial membranes', *Proceedings of the National Academy of Sciences* **112**(38), 11971–11976.
URL: *https://www.pnas.org/content/112/38/11971*

Sasso, S., Pohnert, G., Lohr, M., Mittag, M. & Hertweck, C. (2012), 'Microalgae in the postgenomic era: a blooming reservoir for new natural products', *FEMS Microbiology Reviews* **36**(4), 761–785.
URL: *https://doi.org/10.1111/j.1574-6976.2011.00304.x*

Sato, R., Maeda, Y., Yoshino, T., Tanaka, T. & Matsumoto, M. (2014), 'Seasonal variation of biomass and oil production of the oleaginous diatom Fistulifera sp. in outdoor vertical bubble column and raceway-type bioreactors', *Journal of Bioscience and Bioengineering* **117**(6), 720 – 724.
URL: *http://www.sciencedirect.com/science/article/pii/S1389172313004404*

Satoh, A., Ichii, K., Matsumoto, M., Kubota, C., Nemoto, M., Tanaka, M., Yoshino, T., Matsunaga, T. & Tanaka, T. (2013), 'A process design and productivity evaluation for oil production by indoor mass cultivation of a marine diatom, Fistulifera sp. JPCC DA0580', *Bioresource Technology* **137**, 132 – 138.
URL: *http://www.sciencedirect.com/science/article/pii/S0960852413004550*

Sauer, S. & Plauth, A. (2017), 'Health-beneficial nutraceuticals—myth or reality?', *Applied Microbiology and Biotechnology* **101**(3), 951–961.

Sawai, S., Ohyama, K., Yasumoto, S., Seki, H., Sakuma, T., Yamamoto, T., Takebayashi, Y., Kojima, M., Sakakibara, H., Aoki, T. et al. (2014), 'Sterol side chain reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato', *The Plant Cell* **26**(9), 3763–3774.

Schaller, H., Grausem, B., Benveniste, P., Chye, M.-L., Tan, C.-T., Song, Y.-H. & Chua, N.-H. (1995), 'Expression of the Hevea brasiliensis (HBK) Mull. Arg. 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1 in tobacco results in sterol overproduction', *Plant physiology* **109**(3), 761–770.

Schreiber, U. (2004), Pulse-amplitude-modulation (PAM) fluorometry and saturation pulse method: an overview, *in* 'Chlorophyll a fluorescence', Springer, pp. 279–319.

Seegmiller, A., Dobrosotskaya, I., Goldstein, J., Ho, Y., Brown, M. & Rawson, R. (2002), 'The SREBP Pathway in Drosophila: Regulation by Palmitate, Not Sterols', *Developmental Cell* **2**(2).

Shahzad, N., Khan, W., MD, S., Ali, A., Saluja, S. S., Sharma, S., Al-Allaf, F. A., Abduljaleel, Z., Ibrahim, I. A. A., Abdel-Wahab, A. F., Afify, M. A. & Al-Ghamdi, S. S. (2017), 'Phytosterols as a natural anticancer agent: Current status and future perspective', *Biomedicine & Pharmacotherapy* **88**, 786 – 794.
URL: *http://www.sciencedirect.com/science/article/pii/S0753332216324052*

Shen, C., Dupont, C. L. & Hopkinson, B. M. (2017), 'The diversity of CO2-concentrating mechanisms in marine diatoms as inferred from their genetic content', *Journal of Experimental Botany* **68**(14), 3937–3948.
URL: *https://doi.org/10.1093/jxb/erx163*

Sheppard, V., Poulsen, N. & Kröger, N. (2010), 'Characterization of an endoplasmic reticulum-associated silaffin kinase from the diatom Thalassiosira pseudonana', *Journal of Biological Chemistry* **285**(2), 1166–1176.

Shin, G.-H., Veen, M., Stahl, U. & Lang, C. (2012), 'Overexpression of genes of the fatty acid biosynthetic pathway leads to accumulation of sterols in Saccharomyces cerevisiae', *Yeast* **29**(9), 371–383.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/yea.2916*

Singh, R., De, S. & Belkheir, A. (2013), 'Avena sativa (Oat), A Potential Neutraceutical and Therapeutic Agent: An Overview', *Critical Reviews in Food Science and Nutrition* **53**(2), 126–144. PMID: 23072529.
URL: *https://doi.org/10.1080/10408398.2010.526725*

Slattery, S. S., Diamond, A., Wang, H., Therrien, J. A., Lant, J. T., Jazey, T., Lee, K., Klassen, Z., Desgagné-Penix, I., Karas, B. J. & Edgell, D. R. (2018), 'An Expanded Plasmid-Based Genetic Toolbox Enables Cas9 Genome Editing and Stable Maintenance of Synthetic Pathways in Phaeodactylum tricornutum', *ACS Synth. Biol.* **7**, 328–338.

Smet, E. D., Mensink, R. P. & Plat, J. (2012), 'Effects of plant sterols and stanols on intestinal cholesterol metabolism: Suggested mechanisms from past to present', *Molecular Nutrition & Food Research* **56**(7), 1058–1072.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/mnfr.201100722*

Sonawane, P. D., Pollier, J., Panda, S., Szymanski, J., Massalha, H., Yona, M., Unger, T., Malitsky, S., Arendt, P., Pauwels, L. et al. (2017), 'Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism', *Nature plants* **3**(1), 16205.

Souza, C. M., Schwabe, T. M., Pichler, H., Ploier, B., Leitner, E., Guan, X. L., Wenk, M. R., Riezman, I. & Riezman, H. (2011), 'A stable yeast strain efficiently producing cholesterol instead of ergosterol is functional for tryptophan uptake, but not weak organic acid resistance', *Metabolic Engineering* **13**(5), 555 – 569.
URL: *http://www.sciencedirect.com/science/article/pii/S1096717611000723*

Spilling, K. & Seppälä, J. (2017), 'Measurement of Fluorescence for Monitoring Algal Growth and Health'.

Srivastava, P. K., Desai, D. K., Nandi, S. & Lynn, A. M. (2007), 'HMM-ModE–Improved classification using profile hidden Markov models by optimising the discrimination threshold and modifying emission probabilities with negative training sequences', *BMC bioinformatics* **8**(1), 104.

Stonik, V. & Stonik, I. (2015), 'Low-Molecular-Weight Metabolites from Diatoms: Structures, Biological Roles and Biosynthesis.', *Marine Drugs* **13**(6), 3672–3709.

Strand, T. A., Lale, R., Degnes, K. F., Lando, M. & Valla, S. (2014), 'A new and improved host-independent plasmid system for RK2-based conjugal transfer', *PloS one* **9**(3), e90372.

Summons, R. E., Bradley, A. S., Jahnke, L. L. & Waldbauer, J. R. (2006), 'Steroids, triterpenoids and molecular oxygen', *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**(1470), 951–968.

Szyjka, S. J., Mandal, S., Schoepp, N. G., Tyler, B. M., Yohn, C. B., Poon, Y. S., Villareal, S., Burkart, M. D., Shurin, J. B. & Mayfield, S. P. (2017), 'Evaluation of phenotype stability and ecological risk of a genetically engineered alga in open pond production', *Algal Research* **24**, 378 – 386.
URL: *http://www.sciencedirect.com/science/article/pii/S2211926417300024*

Tan, X. B., Lam, M. K., Uemura, Y., Lim, J. W., Wong, C. Y. & Lee, K. T. (2018), 'Cultivation of microalgae for biodiesel production: A review on upstream and downstream processing', *Chinese Journal of Chemical Engineering* **26**(1), 17 – 30.
URL: *http://www.sciencedirect.com/science/article/pii/S1004954117304305*

Tanaka, T., Yabuuchi, T., Maeda, Y., Nojima, D., Matsumoto, M. & Yoshino, T. (2017), 'Production of eicosapentaenoic acid by high cell density cultivation of the marine oleaginous diatom Fistulifera solaris', *Bioresource Technology* **245**, 567 – 572.
URL: *http://www.sciencedirect.com/science/article/pii/S0960852417315468*

Team, R. C. et al. (2016), 'R: A language and environment for statistical computing'.

Tenenbaum, D. J. (2008), 'Food vs. Fuel: Diversion of Crops Could Cause More Hunger', *Environmental Health Perspectives* **116**(6), 254–257.
URL: *https://ehp.niehs.nih.gov/doi/abs/10.1289/ehp.116-a254*

Theesfeld, C. L., Pourmand, D., Davis, T., Garza, R. M. & Hampton, R. Y. (2011), 'The sterol-sensing domain (SSD) directly mediates signal-regulated endoplasmic reticulum-associated degradation (ERAD) of 3-hydroxy-3-methylglutaryl (HMG)-CoA reductase isozyme Hmg2', *Journal of Biological Chemistry* **286**(30), 26298–26307.

Tholl, D. (2015), Biosynthesis and biological functions of terpenoids in plants, *in* 'Biotechnology of isoprenoids', Springer, pp. 63–106.

Thompson, D. L., Jellum, M. D. & Young, C. T. (1973), 'Effect of controlled temperature environments on oil content and on fatty acid composition of corn oil', *Journal of the American Oil Chemists Society* **50**(12), 540–542.
URL: *https://doi.org/10.1007/BF02640529*

Todd, B. L., Stewart, E. V., Burg, J. S., Hughes, A. L. & Espenshade, P. J. (2006), 'Sterol Regulatory Element Binding Protein Is a Principal Regulator of Anaerobic Gene Expression in Fission Yeast', *Molecular and Cellular Biology* **26**(7), 2817–2831.
URL: *https://mcb.asm.org/content/26/7/2817*

Tokushima, H., Inoue-Kashino, N., Nakazato, Y., Masuda, A., Ifuku, K. & Kashino, Y. (2016), 'Advantageous characteristics of the diatom Chaetoceros gracilis as a sustainable biofuel producer', *Biotechnology for biofuels* **9**(1), 235.
URL: *https://doi.org/10.1186/s13068-016-0649-0*

Tozzi, S., Schofield, O. & Falkowski, P. (2004), 'Historical climate change and ocean turbulence as selective agents for two key phytoplankton functional groups', *Marine Ecology Progress Series* **274**, 123–132.

Traller, J. C. & Hildebrand, M. (2013), 'High throughput imaging to the diatom Cyclotella cryptica demonstrates substantial cell-to-cell variability in the rate and extent of triacylglycerol accumulation', *Algal Research* **2**(3), 244 – 252.
URL: *http://www.sciencedirect.com/science/article/pii/S2211926413000477*

Trentacoste, E. M., Shrestha, R. P., Smith, S. R., Glé, C., Hartmann, A. C., Hildebrand, M. & Gerwick, W. H. (2013), 'Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth', *Proceedings of the National Academy of Sciences* **110**(49), 19748–19753.
URL: *https://www.pnas.org/content/110/49/19748*

Uddin, M. S., Sarker, M. Z. I., Ferdosh, S., Akanda, M. J. H., Easmin, M. S., Bt Shamsudin, S. H. & Yunus, K. B. (2015), 'Phytosterols and their extraction from various plant matrices using supercritical carbon dioxide: a review', *Journal of the Science of Food and Agriculture* **95**(7), 1385–1394.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/jsfa.6833*

Valitova, J., Sulkarnayeva, A. & Minibayeva, F. (2016), 'Plant sterols: diversity, biosynthesis, and physiological functions', *Biochemistry (Moscow)* **81**(8), 819–834.
URL: *https://doi.org/10.1134/S0006297916080046*

Vavitsas, K., Fabris, M. & Vickers, C. (2018), 'Terpenoid Metabolic Engineering in Photosynthetic Microorganisms.', *Genes* **9**(11), 520.

Vik, Å. & Rine, J. (2001), 'Upc2p and Ecm22p, Dual Regulators of Sterol Biosynthesis in Saccharomyces cerevisiae', *Molecular and Cellular Biology* **21**(19), 6395–6405.
URL: *https://mcb.asm.org/content/21/19/6395*

Villanova, V., Fortunato, A. E., Singh, D., Bo, D. D., Conte, M., Obata, T., Jouhet, J., Fernie, A. R., Marechal, E., Falciatore, A., Pagliardini, J., Le Monnier, A., Poolman, M., Curien, G., Petroutsos, D. & Finazzi, G. (2017), 'Investigating mixotrophic metabolism in the model diatom *Phaeodactylum tricornutum*', *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**(1728), 20160404.
URL: *https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2016.0404*

Volkman, J. (2003), 'Sterols in microorganisms', *Applied Microbiology and Biotechnology* **60**(5), 495–506.
URL: *https://doi.org/10.1007/s00253-002-1172-8*

Volkman, J. K. (2005), 'Sterols and other triterpenoids: source specificity and evolution of biosynthetic pathways', *Organic geochemistry* **36**(2), 139–159.

Volkman, J. K. (2016), Sterols in Microalgae, *in* M. Borowitzka, J. Beardall & R. J., eds, 'The Physiology of Microalgae', Springer International Publishing, Cham, pp. 485–505.

Vranová, E., Coman, D. & Gruissem, W. (2013), 'Network analysis of the MVA and MEP pathways for isoprenoid synthesis', *Annual review of plant biology* **64**, 665–700.

Vrieling, E. G., Sun, Q., Tian, M., Kooyman, P. J., Gieskes, W. W., van Santen, R. A. & Sommerdijk, N. A. (2007), 'Salinity-dependent diatom biosilicification implies an important role of external ionic strength', *Proceedings of the National Academy of Sciences* **104**(25), 10441–10446.
URL: *https://doi.org/10.1073/pnas.0608980104*

Vriet, C., Russinova, E. & Reuzeau, C. (2013), 'From Squalene to Brassinolide: The Steroid Metabolic and Signaling Pathways across the Plant Kingdom', *Molecular Plant* **6**(6), 1738–1757.

Véron, B., Billard, C., Dauguet, J.-C. & Hartmann, M.-A. (1996), 'Sterol composition ofPhaeodactylum tricornutum as influenced by growth temperature and light spectral quality', *Lipids* **31**(9), 989–994.
URL: *https://aocs.onlinelibrary.wiley.com/doi/abs/10.1007/BF02522694*

Véron, B., Dauguet, J.-C. & Billard, C. (1998), 'Sterolic biomarkers in marine phytoplankton. II. free and conjugated sterols of seven species used in mariculture', *Journal of Phycology* **34**(2), 273–279.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1529-8817.1998.340273.x*

Wagner, H., Jakob, T. & Wilhelm, C. (2006), 'Balancing the energy flow from captured light to biomass under fluctuating light conditions', *New Phytologist* **169**(1), 95–108.
URL: *https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-8137.2005.01550.x*

Wang, H., Nagegowda, D. A., Rawat, R., Bouvier-Navé, P., Guo, D., Bach, T. J. & Chye, M.-L. (2012), 'Overexpression of Brassica juncea wild-type and mutant HMG-CoA synthase 1 in Arabidopsis up-regulates genes in sterol biosynthesis and enhances sterol production and stress tolerance', *Plant Biotechnology Journal* **10**(1), 31–42.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-7652.2011.00631.x*

Wang, J., Huang, M., Yang, J., Ma, X., Zheng, S., Deng, S., Huang, Y., Yang, X. & Zhao, P. (2017), 'Anti-diabetic activity of stigmasterol from soybean oil by targeting the GLUT4 glucose transporter', *Food & amp; Nutrition Research* **61**.
URL: *https://foodandnutritionresearch.net/index.php/fnr/article/view/1259*

Wang, J.-K. & Seibert, M. (2017), 'Prospects for commercial production of diatoms', *Biotechnology for Biofuels* **10**(1), 16.

Wei, J. H., Yin, X. & Welander, P. V. (2016), 'Sterol Synthesis in Diverse Bacteria', *Frontiers in Microbiology* **7**, 990.
URL: *https://www.frontiersin.org/article/10.3389/fmicb.2016.00990*

Weyer, K. M., Bush, D. R., Darzins, A. & Willson, B. D. (2010), 'Theoretical Maximum Algal Oil Production', *BioEnergy Research* **3**(2), 204–213.
URL: *https://doi.org/10.1007/s12155-009-9046-x*

Weyman, P. D., Beeri, K., Lefebvre, S. C., Rivera, J., McCarthy, J. K., Heuberger, A. L., Peers, G., Allen, A. E. & Dupont, C. L. (2015), 'Inactivation of Phaeodactylum tricornutum urease gene using transcription activator-like effector nuclease-based targeted mutagenesis', *Plant Biotechnology Journal* **13**(4), 460–470.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/pbi.12254*

Whiffin, F., Santomauro, F. & Chuck, C. J. (2016), 'Toward a microbial palm oil substitute: oleaginous yeasts cultured on lignocellulose', *Biofuels, Bioproducts and Biorefining* **10**(3), 316–334.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/bbb.1641*

Wolf, R. B., Cavins, J. F., Kleiman, R. & Black, L. T. (1982), 'Effect of temperature on soybean seed constituents: Oil, protein, moisture, fatty acids, amino acids and sugars', *Journal of the American Oil Chemists' Society* **59**(5), 230–232.
URL: *https://aocs.onlinelibrary.wiley.com/doi/abs/10.1007/BF02582182*

Wollam, J. & Antebi, A. (2011), 'Sterol Regulation of Metabolism, Homeostasis, and Development', *Annual Review of Biochemistry* **80**(1), 885–916. PMID: 21495846.
URL: *https://doi.org/10.1146/annurev-biochem-081308-165917*

Xiao, Y., Savchenko, T., Baidoo, E. E., Chehab, W. E., Hayden, D. M., Tolstikov, V., Corwin, J. A., Kliebenstein, D. J., Keasling, J. D. & Dehesh, K. (2012), 'Retrograde signaling by the plastidial metabolite MEcPP regulates expression of nuclear stress-response genes', *Cell* **149**(7), 1525–1535.

Xue, J., Niu, Y.-F., Huang, T., Yang, W.-D., Liu, J.-S. & Li, H.-Y. (2015), 'Genetic improvement of the microalga Phaeodactylum tricornutum for boosting neutral lipid accumulation', *Metabolic Engineering* **27**, 1 – 9.
  URL: *http://www.sciencedirect.com/science/article/pii/S1096717614001244*

Yang, Z.-K., Niu, Y.-F., Ma, Y.-H., Xue, J., Zhang, M.-H., Yang, W.-D., Liu, J.-S., Lu, S.-H., Guan, Y. & Li, H.-Y. (2013), 'Molecular and cellular mechanisms of neutral lipid accumulation in diatom following nitrogen deprivation', *Biotechnology for Biofuels* **6**(1), 67.

Yao, Y., Lu, Y., Peng, K.-T., Huang, T., Niu, Y.-F., Xie, W.-H., Yang, W.-D., Liu, J.-S. & Li, H.-Y. (2014), 'Glycerol and neutral lipid production in the oleaginous marine diatom Phaeodactylum tricornutum promoted by overexpression of glycerol-3-phosphate dehydrogenase', *Biotechnology for Biofuels* **7**(1), 110.

Zampelas, A. (2014), 'From the Maastricht meeting to the European Atherosclerosis Society Consensus on phytosterols/phytostanols: What is new of an old story?', *Atherosclerosis* **233**(2), 357–358.

Zelazny, A. M., Shaish, A. & Pick, U. (1995), 'Plasma membrane sterols are essential for sensing osmotic changes in the halotolerant alga Dunaliella', *Plant physiology* **109**(4), 1395–1403.
  URL: *https://doi.org/10.1104/pp.109.4.1395*

Zhang, D.-H., Jiang, L.-X., Li, N., Yu, X., Zhao, P., Li, T. & Xu, J.-W. (2017), 'Overexpression of the squalene epoxidase gene alone and in combination with the 3-hydroxy-3-methylglutaryl coenzyme A gene increases ganoderic acid production in Ganoderma lingzhi', *Journal of agricultural and food chemistry* **65**(23), 4683–4690.

Zhang, Z., Sachs, J. P. & Marchetti, A. (2009), 'Hydrogen isotope fractionation in freshwater and marine algae: II. Temperature and nitrogen limited growth rate effects', *Organic Geochemistry* **40**(3), 428–439.

Zhu, B.-H., Shi, H.-P., Yang, G.-P., Lv, N.-N., Yang, M. & Pan, K.-H. (2016), 'Silencing UDP-glucose pyrophosphorylase gene in Phaeodactylum tricornutum affects carbon allocation', *New Biotechnology* **33**(1), 237 – 244.
  URL: *http://www.sciencedirect.com/science/article/pii/S187167841500120X*