

# **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.

This research is supported by the Australian Government Research Training Program.

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 Raffaella 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.

I wish to acknowledge and give thanks for the funding I have received from the UTS International Research Scholarship and President's Scholarship, as well as a Scholarship from the non-profit Colfuturo Foundation of Colombia, the Climate Change Cluster (C3) postgraduate funding and the UTS conference funding schemes.

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 L<sup>A</sup>T<sub>E</sub>X, 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

<b>Acknowledgements</b>	<b>iii</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xvi</b>
<b>Abbreviation List</b>	<b>xvii</b>
<b>Thesis Abstract</b>	<b>xix</b>
<b>1 Phytosterol biosynthesis and production by diatoms (Bacillariophyceae)</b>	<b>1</b>
1.1 Introduction . . . . .	2
1.2 Sterol biosynthesis and diversity in diatoms . . . . .	3
1.2.1 Sterols found in diatoms . . . . .	4
1.2.2 Biosynthesis of sterols in diatoms . . . . .	6
1.2.3 Regulation of sterol biosynthesis . . . . .	10
1.3 Diatoms as a source of phytosterols . . . . .	12
1.3.1 Diatoms as a microbial platform . . . . .	12
1.3.2 Biological activities of phytosterols . . . . .	13
1.3.3 Production of phytosterols from plants . . . . .	14
1.3.4 Features of diatoms for phytosterol production . . . . .	15
1.3.5 Challenges for diatom production systems . . . . .	15
1.4 Investigation and optimisation of sterol production in diatoms . . . . .	16
1.4.1 Growth conditions . . . . .	16
1.4.1.1 Light intensity and spectral quality . . . . .	16
1.4.1.2 Temperature . . . . .	17
1.4.1.3 Medium composition . . . . .	17
1.4.2 Genetic experimentation and engineering in diatoms . . . . .	18
1.5 Conclusions . . . . .	19
<b>2 Levels of diatom minor sterols respond to changes in temperature and salinity</b>	<b>21</b>
2.1 Introduction . . . . .	22

2.2	Materials and Methods . . . . .	23
2.2.1	Diatom species . . . . .	23
2.2.2	L1 medium composition . . . . .	24
2.2.3	Growth and harvesting of diatoms . . . . .	24
2.2.4	Reduced temperature experiments . . . . .	24
2.2.5	Assay for salinity tolerances of diatom species . . . . .	25
2.2.6	Cultivation at altered salinities . . . . .	25
2.2.7	Extraction and analysis of sterols by GC–MS . . . . .	25
2.2.8	Statistical analysis . . . . .	27
2.3	Results . . . . .	27
2.3.1	A shift to cold temperature (4 °C) caused species-specific growth effects and changes in minor sterols . . . . .	27
2.3.2	Species-specific tolerance to different salinities . . . . .	28
2.3.3	Salinity affects the relative contents of non-principal sterols in diatoms . . . . .	29
2.4	Discussion . . . . .	29
2.4.1	The composition of sterol types in different diatoms is not simply explained by clade or environment . . . . .	29
2.4.2	<i>P. tricornutum</i> thrives and shifts its sterol content at a reduced temperature . . . . .	31
2.4.3	Minor sterols respond to changes in salinity and temperature . . . . .	33
2.5	Conclusions . . . . .	35
<b>3</b>	<b>The unique sterol biosynthesis pathway of three model diatoms</b>	<b>37</b>
3.1	Introduction . . . . .	38
3.2	Materials and Methods . . . . .	41
3.2.1	Diatoms . . . . .	41
3.2.2	Dose/effect assay for sub-lethal effective concentrations of inhibitors. . . . .	41
3.2.3	Growth and collection of diatoms for analysis under inhibited conditions. . . . .	41
3.2.4	RNA extraction and transcriptome sequencing. . . . .	42
3.2.5	Transcriptome analysis and enzyme annotation. . . . .	42
3.2.6	Extraction and analysis of sterols by GC-MS. . . . .	43
3.3	Results . . . . .	44
3.3.1	Sub-lethal enzyme inhibition reveals shared sterol intermediates among diatoms . . . . .	44
3.3.2	A terbinafine-insensitive alternative squalene epoxidase is a common feature in diatoms . . . . .	46
3.3.3	Sterols are produced via cycloartenol in three model diatom species . . . . .	47

3.3.4	Cholesterol is produced via a cycloartenol-dependent pathway in diatoms . . . . .	47
3.3.5	Transcriptional dynamics of genes involved in sterol biosynthesis . . . . .	48
3.3.6	Reconstruction of sterol precursor biosynthesis pathways . . . . .	50
3.4	Discussion . . . . .	50
3.4.1	A conserved metabolic core for the biosynthesis of sterols in diatoms . . . . .	50
3.4.2	Species-specific sterol repertoires diversify downstream of a common pathway . . . . .	52
3.4.3	Cholesterol biosynthesis in diatoms shares features with both plants and animals . . . . .	53
3.4.4	Divergence and dynamics in diatom sterol precursor pathways . . . . .	54
3.5	Conclusions . . . . .	58
	Supplemental material . . . . .	60
<b>4</b>	<b>Over-expression of key sterol pathway</b>	<b>70</b>
4.1	Introduction . . . . .	71
4.2	Methodology . . . . .	75
4.2.1	Diatom culturing . . . . .	75
4.2.2	Episome construction and transformation . . . . .	75
4.2.3	Diatom transformation and screening . . . . .	76
4.2.4	Experiments with transgenic diatom cultures . . . . .	76
4.2.5	Extraction and analysis of sterols by GC-MS . . . . .	77
4.2.6	Fluorescence imaging . . . . .	78
4.2.7	Multiple sequence alignment and phylogenetic reconstruction . . . . .	78
4.2.8	Statistical analysis . . . . .	78
4.3	Results . . . . .	79
4.3.1	Identification of putative HMGR from <i>T. pseudonana</i> and <i>P. tricornutum</i> . . . . .	79
4.3.2	Phylogenetic analysis of HMGR and conserved protein domains . . . . .	81
4.3.3	Expression and localization of putative HMGR and tHMGR . . . . .	81
4.3.4	Influence of HMGR and tHMGR expression on sterol levels in <i>T. pseudonana</i> and <i>P. tricornutum</i> . . . . .	83
4.3.5	Heterologous expression of a Stramenopile putative SQE . . . . .	87
4.4	Discussion . . . . .	87
4.4.1	HMGR is largely conserved among diatoms and lacks a conventional sterol sensing domain . . . . .	87
4.4.2	HMGR and tHMGR over-expression lead to accumulation of sterol pathway intermediates in <i>P. tricornutum</i> . . . . .	88
4.4.3	Levels of end-point campesterol increased after heterologous expression of SQE in <i>P. tricornutum</i> . . . . .	91

4.5	Conclusions . . . . .	93
	Supplemental material . . . . .	94
<b>5</b>	<b>General discussion, future perspectives and conclusions</b>	<b>119</b>
5.1	General conclusions . . . . .	120
5.1.1	Species-specific differences in sterol composition are related to environmental conditions and diversification of sterol biosynthesis . . . .	120
5.1.2	Metabolic engineering of sterol metabolic pathway of diatoms . . . .	121
5.2	Future perspectives . . . . .	124
5.2.1	Sterol derivatives and specialized reactions of sterol biosynthesis . . .	124
5.2.2	The unexplored sterol homeostasis mechanisms in diatoms . . . . .	126
5.3	Concluding remarks . . . . .	127
	<b>References</b>	<b>128</b>



# List of Figures

Figure 1.1	Steroid skeleton structure according to IUPAC 1989 recommendations (Moss, 1989). Note: previous guidelines (1976) numbered the atoms 24 <sup>1</sup> and 24 <sup>2</sup> as 28 and 29. 24 <sup>1</sup> and 24 <sup>2</sup> can be also found in the literature as 24' and 24'' . . . . .	2
Figure 1.2	Overview of the MVA and MEP pathway and hypothetical early conserved steps of sterol biosynthesis in diatoms. . . . .	9
Figure 1.3	The SREBP pathway in mammals from (Espenshade & Hughes, 2007). .	11
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. . . . .	26
Figure 2.2	Sterol abundance and distribution in the diatoms <i>Chaetoceros muelleri</i> , <i>Phaeodactylum tricornutum</i> and <i>Thalassiosira pseudonana</i> growing under standard conditions: L1 medium at 18 °C under cool white continuous light (100 $\mu\text{mol photons m}^{-2}\text{s}^{-1}$ ; n = 6). . . . .	28
Figure 2.3	Growth curves of diatoms under 4 °C. Time 0 represents the beginning of the temperature shock (n = 3). . . . .	29
Figure 2.4	Sterol levels in diatoms growing at 4 °C. Abundance is given in terms of $\mu\text{g}$ of sterol per mg of biomass dry weight (n = 3). . . . .	30
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 $\mu\text{g}$ of sterol per mg of biomass dry weight (n = 3). A t-test is provided for samples with n = 3. .	32
Figure 2.6	Screening of growth yield and growth rate versus the salinity level in the diatoms <i>C. muelleri</i> (Cm), <i>P. tricornutum</i> (Pt) and <i>T. pseudonana</i> (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. . .	33
Figure 2.7	Sterol levels in diatoms growing at different salinity levels. Abundance is given in terms of $\mu\text{g}$ of sterol per mg of biomass dry weight (n = 3). . .	35

- 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); 14 $\alpha$ -demethylase (E.C. 1.14.14.154);  $\Delta^8$ ,  $\Delta^7$  isomerase (E.C. 5.3.3.5). . . . . 40
- 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: AC-CoA: acetyl-CoA; AACCOA: aceto-acetyl-CoA; HMGGCOA: 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.) . . . . . 49

- 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  $\log$ CPM 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.) . . . . . 51
- 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  $\log$ CPM 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.) . . . . . 56

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 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 <sub>2</sub> fold-changes (Log <sub>2</sub> 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: <i>Chaetoceros muelleri</i> ; Pt: <i>Phaeodactylum tricornutum</i> ; Tp: <i>Thalassiosira pseudonana</i> . 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.) . . . . .	57
Figure 3.S1	Experimental set up in large scale photobioreactors for inhibitors experiment. . . . .	61
Figure 3.S2	Chromatograms from GC-MS analysis of each diatom specie growing without inhibitor addition. (a) Commercial standards (b) <i>P. tricornutum</i> (c) <i>T. pseudonana</i> (d) <i>C. muelleri</i> . . . . .	62
Figure 3.S3	Microscale (200 $\mu$ L) inhibitor dose-effect screen in 96-well plates. . . . .	63
Figure 3.S4	Cell counts during replicate 1.2L aerated batch culture treatment experiments ( $n=3$ ). . . . .	64
Figure 3.S5	Chromatograms from GC-MS analysis of each diatom specie under inhibitors treatment. (a) Terbinafine (b) Fluconazole (c) Ro 48-8071 . . . . .	65
Figure 3.S6	Chromatograms from GC-MS analysis of each diatom species under fenpropimorph treatment (a) Cycloartenol commercial standard, (b) Lanosterol commercial standard (c) <i>P. tricornutum</i> (d) <i>T. pseudonana</i> (e) <i>C. muelleri</i> . . . . .	66
Figure 3.S7	Mass spectra of putatively identified sterols. . . . .	67
Figure 3.S8	GO term enrichments among differentially expressed orthologous transcripts from mRNA-seq transcriptomes. . . . .	68

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). . . . .	69
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 <i>N. oceanica</i> , NoSQE, oxidosqualene cyclase, OSC. . . . .	74
Figure 4.2	Sequence alignment of HMGR protein (catalytic domain) from model organisms. Asterisks (*) indicate conserved catalytic residues. . . . .	80
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 <i>P. tricornutum</i> and <i>T. pseudonana</i> used in this study. . . . .	82
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 <i>N. oceanica</i> , NoSQE. Scale bars correspond to 10 $\mu\text{m}$ . . . . .	84
Figure 4.5	Sterol levels in <i>P. tricornutum</i> transformants. (a) End-point sterol (b) intermediates accumulation. Identical letters denote no statistically significant differences among groups using the Pairwise Wilcoxon Rank Sum tests ( $p < 0.05$ , $n = 9$ ). . . . .	86
Figure 4.S1	Maps of plasmids used for <i>T. pseudonana</i> transformation. CDS are shown in pink, in green promoters and terminators and fluorescence mVenus gene in yellow. . . . .	97

Figure 4.S2	Maps of plasmids used for <i>P. tricornutum</i> transformation. CDS are shown in pink, in green promoters and terminators and fluorescence mVenus gene in yellow. . . . .	97
Figure 4.S3	Screening of transformed <i>T. pseudonana</i> clones. Colonies were transferred from selection agar plates to liquid medium with 100 $\mu\text{g}/\text{mL}$ nourseothricin in 96 well plates, after three days, flow cytometry measurements were taken. . . . .	98
Figure 4.S4	Screening of transformed <i>P. tricornutum</i> clones. Colonies were transferred from selection agar plates to liquid medium with 10 $\mu\text{g}/\text{mL}$ blasticidin in 96 well plates, after three days, flow cytometry measurements were taken. . . . .	99
Figure 4.S5	mVenus fluorescence during full scale experiment in <i>T. pseudonana</i> transformants (n = 3). . . . .	100
Figure 4.S6	Chlorophyll fluorescence during full scale experiment in <i>T. pseudonana</i> transformants (n = 3). . . . .	101
Figure 4.S7	Growth curves during full scale experiment for <i>T. pseudonana</i> transformants (n = 3). . . . .	102
Figure 4.S8	Maximum quantum yield for <i>T. pseudonana</i> transformants during full scale experiment (n = 3). . . . .	103
Figure 4.S9	mVenus fluorescence during full scale experiment in <i>P. tricornutum</i> transformants (n = 3). . . . .	104
Figure 4.S10	Chlorophyll fluorescence during full scale experiment in <i>P. tricornutum</i> transformants (n = 3). . . . .	105
Figure 4.S11	Growth curves during full scale experiment for <i>P. tricornutum</i> transformants (n = 3). . . . .	106
Figure 4.S12	Maximum quantum yield for <i>P. tricornutum</i> transformants during full scale experiment (n = 3). . . . .	107
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. . . . .	122

- Figure 5.2 Expression profile of gene coding for AltSQE (Phatr3-J45494) in *P. tricornutum*. Source: <https://algnaut.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). . . . . 125
- 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) . . . . . 126

# List of Tables

Table 1.1	The occurrence of sterols in diatom species. Reported biological activities are noted. . . . .	5
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. . . . .	10
Table 1.3	Sterol and oil contents cited from different sources on a dry weight (d.w.) basis. . . . .	14
Table 1.4	Culture conditions of diatoms reported for the production of sterol compounds. . . . .	17
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. . . . .	45
Table 3.S1	Read count statistics for mRNA sequencing. . . . .	60
Table 3.S2	Results of Benchmarks for Universal Single-Copy Orthologs (BUSCO) for de novo transcriptome assemblies. . . . .	60
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. . . . .	94
Table 4.S2	Source and number of transmembrane domains of HMGR sequences used for phylogenetic and domain analysis. . . . .	96



# 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-C-methyl-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 cytidyltransferase

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<sub>0</sub>:** Inhibition of enzymes involved in the sterol metabolic pathway of diatoms does not result in the production of phytosterol intermediate compounds

**H<sub>0</sub>:** 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<sub>0</sub>:** It is not possible to transgenically alter the sterol products or amounts of diatoms; the natural levels are strictly balanced and regulated

**H<sub>a</sub>:** 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?