Development of a molecular toolkit to genetically engineer the microalga *Nannochloropsis gaditana* CCMP526 for biotechnology applications

A thesis submitted to the University of Technology Sydney for the degree of DOCTOR OF PHILOSOPHY (PhD)

Submitted October 2020

Margaret Ramarajan, BSc, MSc.

Supervisors: Professor Peter J. Ralph, Dr Michele Fabris,

Climate Change Cluster, School of Life Sciences, University of Technology Sydney

Certificate of original authorship

I, Margaret Ramarajan declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Life Science at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced 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. Any help that I have received in my research work and the preparation of this thesis itself has been acknowledged. This research is supported by the Australian Government Research Training Program.

Signature:

Production Note:
Signature removed prior to publication.

Date: 02/10/2020

Acknowledgements

I thank God for his abundant grace and wisdom to fulfil my PhD research. I would like to thank Professor Peter Ralph for giving me this wonderful opportunity to work in Climate Change Cluster (C3) at the University of Technology Sydney, and my co-supervisor Dr Michele Fabris for their time, strong support and help throughout the entire duration of my PhD project. This research was supported by the UTS International Research Scholarship (UTS IRS) and UTS President's Research Scholarship.

I also would like to thank Dr Mathieu Pernice and Dr. Raffaela M. Abbriano for their help and support with data analysis. I recognise Dr Unnikrishnan Kuzhiumparambil for providing support in GC-MS analysis and Kun Xiao for the much-needed teaching and guidance in flow cytometry. I acknowledge Dr Tim Kahlke for his guidance in using bioinformatics tools. I am grateful for the custom script developed by Dr Chris Evenhuis for microscopic cell-count. I thank Dr Stephen Woodcock for his valuable guidance on statistical analysis. I acknowledge Dr Audrey Commault, Dr Justin Ashworth, Dr Christopher Hall, and Dr Leen Labeeuw for providing useful counsel and discussions.

I thank the technical staff: Paul Brooks, Gemma Armstrong, Scott Allchin, Taya Lapshina, Sue Fenech, and Rachel Keeys, for helping me with the various training and inductions and other lab-based needs. I am grateful for my fellow research scholars for making this PhD journey interesting and fun: Ana Cristina Jaramillo Madrid, Dylan Russell, Jack Adriaans, Jestin George, Lorenzo Barolo, Marco Giardina, Matthias Windhagauer, Navpreet Kaur Walia, Nathan Williams, Nine Le Reun and Shawn Price. Special thanks to Rachelle Tran and Abeeha Khalil for being supportive and all the ice-cream breaks.

Lastly, I extent my heartfelt gratitude for the love and support I received from my family during these three years of my PhD. My parents, my husband and my lovely six-year-old daughter. I thank them for their prayers and encouragement to undertake a PhD degree overseas and being a pillar of strength during hard and stressful times.

Preface

This thesis has been prepared for submission in a thesis by compilation format, whereby the thesis contains a combination of published and publishable work. Considering that this thesis is presented as a series of ready to submit manuscripts, there is a degree of repetition across chapters, particularly within the introductions and materials and methods sections of Chapter 2, 3 and 4. Published work (Chapter 2) has been incorporated into this thesis and appears as it was presented to the journal immediately prior to publication with the following modifications: i) the font and format was changed to maintain consistency across the thesis, ii) figures and tables were re-numbered to reflect the chapter numbering and iii) supplementary information for each chapter appear in the appendix and have been re-numbered accordingly. The referencing format used throughout this thesis conforms to the requirements of the journal Nature.

List of publications:

Chapter 2:

Novel endogenous promoters for genetic engineering of the marine microalga *Nannochloropsis gaditana* CCMP526

Margaret Ramarajan[a], Michele Fabris[a,b], Raffaela M. Abbriano[a] , Mathieu Pernice[a], Peter J. Ralph[a]

[a] University of Technology Sydney, Climate Change Cluster (C3), Faculty of Science, Broadway Campus, Ultimo NSW 2007, Australia

[b] CSIRO Synthetic Biology Future Science Platform, GPO box 2583, Brisbane QLD 4001, Australia

Published in December 2019 within the journal of 'Algal Research- Biomass, Biofuels and Bioproducts' volume 44, as a research article.

Author Contribution: M.R. performed the experiments; M.F. and P.J.R. conceived the research project; M.F, R.A. and M. P. contributed in data analysis; M.R. wrote the

manuscript; all authors edited and reviewed the manuscript.

Chapter 4:

Genetic engineering of the microalga *Nannochloropsis gaditana* for the production of sesquiterpene β-caryophyllene

Margaret Ramarajan[a], Michele Fabris[a,b], Unnikrishnan Kuzhiumparambil[a] , Peter J. Ralph[a]

[a] University of Technology Sydney, Climate Change Cluster (C3), Faculty of Science, Broadway Campus, Ultimo NSW 2007, Australia

[b] CSIRO Synthetic Biology Future Science Platform, GPO box 2583, Brisbane QLD 4001, Australia

Manuscript: In-prep for submission to Metabolic Engineering

Author Contribution: M.R. performed the experiments; M.R, M.F. and P.J.R. conceived the research project; U.K. performed GC-MS; M.F, and U.K. contributed in data analysis; M.R. wrote the manuscript; all authors edited and reviewed the manuscript.

Conferences:

- Poster presentation at the Plant Biology World Summit 2020. Title: Genetic engineering of the marine microalga *Nannochloropsis gaditana* CCMP526 for the production of sesquiterpene β-caryophyllene
- Poster presentation at Australia New Zealand Marine Biotechnology Symposium (ANZMBS 2019). Title: Novel endogenous promoters for genetic engineering of the marine microalga *Nannochloropsis gaditana* CCMP526
- Attended the Synthetic Biology Australasia Conference, Sydney 2017

Table of Contents

List of Figures

Supplementary figures:

List of Tables

Supplementary tables:

General Abbreviations A to Z

Acronym	Definition
AACT	Acetoacetyl-CoA thiolase
ABA	Abscisic acid
ABRE	ACGT-containing abscisic acid response element
AMT	Ammonium transporter
ANOVA	Analysis of variance
ANZMBS	Australia New Zealand Marine Biotechnology Symposium
AOC	Auto Injector/Auto Sampler
AP2	transcription factor APETALA2
ATP	Adenosine triphosphate
BCA	Bovine serum albumin
bHLH	Basic helix-loop-helix transcription factor
BKT	β-carotene ketolase
bZIP	basic leucine-zipper transcription factor
CCAP	The Culture Collection of Algae and Protozoa
CCM	Carbon concentration mechanisms
CDP-ME	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol
CDPMEK	4-diphosphocytidyl-2c-methyl-d-erythritol kinase
CMS	4-Diphosphocytidyl-2C-methyl-D-erythritol synthase
CMV	cytomegalovirus
CRIBI	Centro di Ricerca Interdipartimentale per le Biotecnologie Innovative
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CSM	2-c-methyl-d-erythritol 4-phosphate cytidylyltransferase
CYP	Cytochrome P450 monooxygenase
DCM	Dichloromethane
DHA	Docosahexaenoic acid
DMAPP	Dimethyl allyl pyrophosphate
DMSO	Dimethyl sulfoxide
DNA	Deoxy-ribonucleic acid
DOXP	Deoxyxylulose 5-Phosphate
DXP	1-deoxy-D-xylulose-5-phosphate
DXR	1-deoxy-D-xylulose-5-phosphate reductoisomerase
DXS	1-deoxy-d-xylulose 5-phosphate synthase
ECL	Enhanced chemiluminescence
EDP	Eukaryotic Promoter Database
EEP	Endonuclease/exonuclease/phosphatase
EIC	Extracted-ion chromatogram
EPA	Eicosapentaenoic acid
EPPSII	Extrinsic protein in photosystem II
ESAW	Enriched Seawater, Artificial Water

FACS	Fluorescence assisted cell sorting
FCCP	Fucoxanthin-chlorophyll a-c binding protein
FDA	Food and Drug Administration
FITC	Fluorescein Isothiocyanate
FMDV	Foot-and-mouth disease virus
FPKM	Fragments Per Kilo base of exon per Million fragments mapped
FPP	Farnesyl pyrophosphate
FPPS	Farnesyl pyrophosphate synthase
GAL	Promoter with Gal4p-binding sites
GDPD	Glycerophosphoryl diester phosphodiesterase
GEO	Gene Expression Omnibus
GFP	Green Fluorescent Protein
GGPP	Geranylgeranyl pyrophosphate
GPP	Geranyl pyrophosphate
GUS	β-glucuronidase
HDR	4-hydroxy-3-methylbut-2-enyl diphosphate reductase
HDS	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
HMBDP	4-hydroxy-3- methylbut-2-enyl diphosphate
HMBPP	hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
HMG-CoA	β-Hydroxy β-methylglutaryl-CoA
HMGR	β-Hydroxy β-methylglutaryl-CoA reductase
HMGS	β-Hydroxy β-methylglutaryl-CoA synthase
HSF	Heat shock transcription factor
HSP	Heat shock protein
HTPG	Heat or thermotolerance protein G
HYP	Hypothetical protein
IDI	Isopentenyl diphosphate:dimethylallyl diphosphate isomerase
IDT	Integrated DNA technologies
INR	Initiator element
IPM	Isopropyl myristate
IPP	Isopentenyl pyrophosphate
ISPD	Isoprenoid synthase domain
ISPF	2-c-methyl-d-erythritol–cyclodiphosphate
LDSP	Lipid droplet surface protein
LED	Light Emitting Diode
LIP	Light-inducible protein
MECDP	2-C-methyl-D-erythritol 2,4-cyclodiphosphate
MEP	Methyl-D-erythritol 4-phosphate
MVA	Mevalonate
MYB	Myeloblastosis proto-oncogene, transcription factor
MYC	Myeloblastosis rroto-oncogene, basic helix-loop-helix transcription factor
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NCBI	National Center for Biotechnology Information
NCLDV	Nucleocytoplasmic large DNA viruses

NF YB	Nuclear transcription factor-Y subunit beta
NF YC	Nuclear transcription factor-Y subunit gamma
NIES	Nitrate inducible expression system
NSW	New South Wales
NUDIX	Nucleoside diphosphates linked to some moiety X
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PAM	Pulse-amplitude modulated fluorometry
PBS	Phosphate buffer saline
PCC	The Pasteur culture collection of cyanobacteria
PCR	Polymerase chain reaction
PMD	Mevalonate 5-pyrophosphate decarboxylase
PMK	Mevalonate 5-phosphate kinase
PSII	Photosystem II
PSU	Practical Salinity Unit
PTM	Post-translational modifications
PUFA	Polyunsaturated fatty acids
PVDF	Polyvinylidene difluoride
QLD	Queensland
RIGS	Repeat-induced gene silencing
RNA	Ribonucleic acid
RPKM	Reads Per Kilobase of transcript, per Million mapped reads
SDS	Sodium dodecyl sulfate
SIM	Selective Ion Monitoring
SIT	Silicon transporters
SNP	Single nucleotide polymorphisms
SPS	Sodium Phosphate Symporter
SQDG	Sulfoquinovosyldiacylglycerol
SQS	Squalene synthase
STRE	Stress response element
TAG	Triacylglycerols
TAIR	The Arabidopsis Information Resource
TCA	Tricarboxylic acid
TMHMM	Transmembrane hidden markov model
TSS	Transcription start sites
TSSP	Transcription start sites for plant
TUB	β-tubulin
UEP	Ubiquitin extension protein
USA	United States of America
UTR	Untranslated Region
UTS	University of Technology Sydney
VCP	Violaxanthin/Chlorophyll α-binding Protein
YAC	Yeast Artificial Chromosome
YFP	Yellow Fluorescent Protein

Thesis summary

Model microalgae such as *Chlamydomonas reinhardtii* and *Phaeodactylum tricornutum* are well suited for genetic engineering as they possess an array of well characterised genetic tools that can be used for gene and genome manipulations. *Nannochloropsis gaditana* is a photosynthetic oleaginous microalga that has been studied extensively due to a broad range of industrial applications such as oil, polyunsaturated fatty acids (PUFAs) and pigments. Even though genetic engineering and synthetic biology resources are rapidly advancing and widening this alga's industrial potential, there are still some limitations such as a narrow repertoire of functional promoters with limited diversity in terms of expression range and inducibility; and a lack of knowledge of specific regulatory elements that control transcription in *Nannochloropsis* spp. that would allow for predictable expression and rational design of promoter regions. To date, several studies involving gene over-expression have been conducted in *Nannochloropsis* spp. using endogenous promoters, but in-depth promoter analysis done in other model microalgae like *P. tricornutum* and *C. reinhardtii* is still missing with *Nannochloropsis* spp. By identifying and profiling a suite of promoters for use in *N. gaditana* CCMP526, we have expanded the genetic toolbox available for this industrially relevant microalgal species both for biotechnological applications such as metabolic engineering or recombinant protein production, and to understand the biology of *Nannochloropsis.* In this research we also explored *N. gaditana* as a novel platform for sesquiterpenoid biosynthesis and it represents the first report of terpenoid engineering in this microalga, which shed some light into metabolic pathway engineering for terpenoid production in *N. gaditana*. Whilst this PhD thesis offers new knowledge into terpenoid biosynthesis, it also provides candidate gene promoters, which can be used in a wide variety of biotechnology application in *N. gaditana*.