Molecular physiological responses and acclimation of the seagrass species *Z. muelleri* to light limitation

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

Submitted January, 2017

Peter A. Davey, BSc. Hons, MSc.

Supervisors: Professor Peter Ralph, Dr Mathieu Pernice and Dr Rudy Dolferus

The Thesis presented meets the standards and requirements set out by the University of Technology Sydney

Certificate of original ownership

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the doctoral degree and/or fully acknowledged within the text. I also certify that this thesis has been written by me (Peter A. Davey). Any help that I have received in my research work and the preparation of this thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in this thesis.

Acknowledgments

I would first like to thank Professor Peter Ralph for giving me the exciting opportunity to work within the Climate Change Cluster (C3) at the University of Technology Sydney, and my co-supervisors, Dr Mathieu Pernice and Dr Rudy Dolferus for the advice and time they dedicated to helping me over the duration of the project. I would also like to thank Dr Martin Schliep, for his support as my co-supervisor in the first year of my PhD. This research was supported by the UTS President's Research Scholarship, UTS International Research Scholarship, and a CSIRO Office of Chief Executive post-graduate scholarship.

I acknowledge Dr Gaurav Sablok for his bioinformatics based teachings, and thank the following colleagues for providing fruitful advice and discussion; Manoj Gupta Kumar; Unnikrishnan Kuzhiumparambil; Justin Ashworth; Chris Evenhuis; Nasim Mohammadi; Kasper Brodersen; Alex Thomson; Stacey Trevathan-Tackett; Milan Szabo; Tony Larkum; Rachel Levin; and John Raven. In relation to fieldwork, the TropWater team associated with James Cook University (Queensland), Katherina Petrou (UTS), Frederic Cadera (visiting scholar) and Vincent Schols (visiting scholar) should be thanked. I thank the lab technicians; Paul Brooks; Stacey Ong; Gemma Armstrong; Peter Jones; Graham Powolski; and Kun Xie for providing first-class support in the laboratories. On many occasions, Mike Lake (IT Department, UTS) Brian Haas (Broad Institute, USA), Don Gilbert (Indiana University, USA), members of the 'biostars.org' and 'seqanswers' online bioinformatics forums should be thanked for providing me with insight on how to solve numerous computational, bioinformatics and software related issues.

Finally, I wish to express my gratitude for the love and support that my family and friends gave me during this journey; My parents; my brother Steven; my late grandmother and my partner Emma; I thank them for encouraging me to undertake a PhD degree on the other side of the world, and for their continuous support and encouragement during my highs and lows.

Preface

The Chapters within this PhD thesis have been written with the intention of submission to scientific journals. The chapters are therefore presented in a typical journal format, ready for submission. Chapter 1 has already been published in a scientific journal (Functional and Integrative Genomics; IF = 2.265) as a critical literature review. Chapters 4 and 5 will be submitted in the near future to scientific journals as original research articles. Scientific work, which I have contributed to, is listed in Appendix 4, one of these pieces of work has been published in another journal (Frontiers in Plant Science; IF =4.495), whilst the other piece of work (a book chapter) is in preparation. Given that this thesis is presented as a series of ready to submit manuscripts, there is an element of repetition in the introduction of some of the chapters.

A foreword

"Look deep into nature, and then you will understand everything better" - Albert Einstein

Personally for me, the above quote sums up my PhD journey over the past three and a half years. Only through bioinformatics analyses, did true meaning come from the observations I made in the laboratory at UTS. To unravel the complexity of one organism over three and a half years has been a huge accomplishment for me, one that I have immenseley enjoyed; however, with the satisfaction came the challenge, one that I found testing at times. By undertaking this PhD, I feel that I have come a long way, learning about myself, and seagrasses in many ways. A journey, which gave me appreciation for how complex nature can be. As the saying goes - "There's more than meets the eye."

General Abbreviations A to Z

2D-IEF	Two Dimensional- Isoelectric Focusing
ABA	Abscisic Acid
AFLP	Amplified Fragment Length Polymorphism
AL-fact	Actinic Light Factor
AMSA	Australian Marine Sciences Association
ANU	Australian National University
ATP	Adenosine 5'-triphosphate
AU	Arbitrary Units
BBH	Bidirectional Best BLAST Hit
BCV	Biological Coefficient of Variation
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BP	Biological Process
BUSCO	Benchmarking Universal Single-Copy Orthologs
BWA	Burrows-Wheeler Aligner
C3	C3 photosynthesis / Climate Change Cluster
C4	C4 photosynthesis
CA	Carbonic Anhydrase
CAP3	Contig Assembly Program 3
CC	Cellular Component
CCM	Carbon Concentrating Mechanism
cDNA	complementary Deoxyribonucleic Acid
CEGMA	Core Eukaryotic Genes Mapping Approach
CHIP-Seq	Chromatin Immunoprecipitation-Sequencing
chl	Chlorophyll
Ci	Inorganic Carbon
CPM	Counts Per Million
CRBB	Conditional Reciprocal Best BLAST
Ct	Cycle threshold
Cyt.b6f	Cytochrome b6f
ddRAD	double digest Restriction Associated DNA
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide triphosphate
DOGMA	Dual Organellar GenoMe Annotator
DVC A	Divinyl chlorophyllide A
EBI	European Bioinformatics Institute
EMBL	European Molecular Biology Laboratory
EST	Expressed Sequence Tag
ETR	Electron Transport Rate
FC	Fold Change
FDM-SSR	Functional Domain Marker - Simple Sequence Repeat
FDR	False Discovery Rate

FPKM	Fragments Per Kilobase of transcript per Million mapped reads
Fv/Fm	Maximum quantum yield of photosystem II
G3P	Glyceraldehyde 3-phosphate
GABA	Gamma-aminobutyric acid
GBS	Genotype By Sequencing
GC-MS	Gas Chromatography Mass Spectrometry
gDNA	genomic Deoxyribonucleic Acid
GGPP	Geranylgeranyl diphosphate
GLM	Generalized Linear Model
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
HPC	High Performance Computing
HPLC	High Performance Liquid Chromatography
Ik	Half-saturation constant
ISBW	International Seagrass Biology Workshop
KAAS	KEGG Automatic Annotation Server
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC-MS/MS	Liquid Chromatography Tandem-Mass Spectrometry
LHC	Light Harvesting Complex
LHCI / LHCA	Photosystem I Light Harvesting Complex
LHCII/ LHCB	Photosystem II Light Harvesting Complex
LL	Light Limited
MDS	Multi-Dimensional Scaling
MF	Molecular Function
Meas. Int	Measuring Intensity
MEME	Multiple EM for Motif Elicitation
MEP	2-C-methyl-D-erythritol 4-phosphate
MIQE	Minimum Information for Publication of Quantitative. Real-Time PCR Experiments
miRNAs	micro Ribonucleic Acid
MISA	MIcroSAtellite identification tool
mRNA-Seq	Messenger RNA-Sequencing
MTDB	Mangrove Transcriptome Database
mya	million years ago
NADP	Nicotinamide adenine dinucleotide phosphate
NADPH / NADP(H)	Reduced nicotinamide adenine dinucleotide phosphate
NCBI	National Center for Biotechnology Information
ncRNAs	non-coding Ribonucleic Acids
NGS	Next Generation Sequencing
No-RT	No Reverse Transcriptase
NPQ	Non-Photochemical Quenching
NRQ	Normalised Relative Quantity
NSW	New South Wales

NZGL	New Zealand Genomics Limited
ORF	Open Reading Frame
P-I	Photosynthesis - Irradiance
PAM	Pulse Amplitude Modulated
PC	Plastocyanin
PCR	Polymerase Chain Reaction
PDA	Photodiode Array Detector
PEPC	Phosphoenolpyruvate carboxylase
PhiX	Enterobacteria Phage PhiX174
polyA	Polyadenylated
P(MC)	Probability (Multiple Correction)
PPOs	Polyphenol Oxidases
PQ	Plastoquinone
PSI	Photosystem I
PSII	Photosystem II
psu	Practical salinity units
PTFE	Polytetrafluoroethylene
qPCR	Quantitative Polymerase Chain Reaction
QC	Quality Check
QTL	Quantitative Trait Loci
RAPD	Random Amplification of Polymorphic DNA
RBH	Reciprocal Best BLAST Hit
rETRmax	relative maximum Electron Transport Rate
RFLP	Random Fragment Length Polymorphism
RIN	RNA Integrity Number
RLC	Rapid Light Curve
RNA	Ribonucleic Acid
RNA-Seq	RNA-Sequencing
ROI	Reactive Oxygen Intermediate
ROS	Reactive Oxygen Species
RPKM	Reads Per Kilobase of transcript per Million mapped reads
RSEM	RNA-Seq by Expectation-Maximization
RT-qPCR	Real Time-quantiative Polymerase Chain Reaction
RuBisCo	Ribulose-1,5-bisphosphate carboxylase/oxygenase
RubP	Ribulose-1,5-bisphosphate
SAM	Sequence Alignment Map
Sat. Width	Saturation Width
Sat.Int	Saturation Intensity
SNP	Single Nucleotide Polymorphism
sRNA	small Ribonucleic Acid
SRP	Signal Recognition Particle
SSL	Super Saturating Light
SSR	Simple Sequenc Repeat

STAR	Spliced Transcripts Alignment to a Reference
TAIR	The Arabidopsis Information Resource
TBAA	tert-Butyl acetoacetate
TCA	Tricarboxylic Acid
TF	Transcription Factor
Tm	Melting temperature
TMM	Trimmed Mean of M-Values
TPM	Transcripts Per Million
UNIX	UNIX programming environment
UPLC	Ultra Performance Liquid Chromatography
UV	Ultra Violet
Yi	Initial quantum yield of photosystem II

List of figures

List of tables **Pg.**

CONTENTS

Thesis summary

Understanding how a keystone marine species responds to its extrinsic environment is important to ensure adequate conservation measures are in place, especially with increasing reports of climate change and anthropogenic disturbance events. For the Southern Hemisphere seagrass, *Zostera muelleri*, this scenario is no different. This keystone species is native to Australia and New Zealand, providing many socio-economic benefits to the coastal zone. Over the past few decades, a reduction in water quality (light limitation) has led to numerous reports of *Z. muelleri* meadow loss in Australia and New Zealand. Although seagrass biologists have a firm understanding of the physiological, morphological and ecological changes within light limited *Z. muelleri* meadows, no current knowledge exists on how *Z. muelleri* responds to light limitation at the transcriptional level. By investigating transcriptional regulation, new knowledge was obtained on how this species responds to light limitation, allowing for more appropriate conservation measures. Encompassing the advances in RNA-Seq, this project has examined how *Z. muelleri* responds to light limitation over a 14-day period, through transcriptional regulation, photobiology and physiology, both at the nuclear and chloroplastic level. Main findings indicate that important regulational shifts occur in genes associated with photosynthesis, photo-pigments, carbon metabolism, reactive oxygen species (ROS) homeostasis and secondary defence metabolism. Both nuclear and chloroplast encoded genes involved in photosynthetic processes have been shown to be correlated with downstream changes in photophysiology, and thus are both crucial for the response as well as the acclimation to light limitation. This research also compared genome-guided transcriptome assembly versus *de novo* assembly, indicating the superiority of genome-guided protocols when a genome is available. Whilst this PhD thesis offers a new level of knowledge to seagrass biologists, it also provides candidate molecular markers, which can be used in future monitoring efforts and population genetic studies.

PhD thesis aims and objectives

The overall aim of this thesis is to investigate how *Zostera muelleri* responds to light limitation using a multi-disciplinary approach. By combining mRNA-seq and RT-qPCR protocols with already established photobiology and pigment profiling techniques, we will not only obtain a new level of understanding on how this seagrass species responds at the transcriptional level, but also how transcriptional regulation is linked to downstream changes in photophysiology. Such work is timely, given that seagrasses are increasingly threatened by light limitation within the coastal environment.

Objectives

- To examine and provide a critical literature review on the current state of molecular profiling and omics techniques in seagrass biology, whilst identifying key knowledge gaps in previous and current research.

- To address the background knowledge associated with light perception in higher plants, seagrasses and *Zostera muelleri*: fundamental knowledge and further direction for research will be discussed.

- To complete *in silico* characterisation and data mining of the *Z. muelleri de novo* transcriptome, based on whole plant tissue.

- To characterise leaf tissue-specific responses of *Z. muelleri* to light limitation; to establish links between transcriptional regulation of nuclear-encoded genes and downstream photophysiology.

- To investigate the expression of chloroplast-encoded photosynthetic genes in *Z. muelleri* in response to light limitation. To designate suitable reference genes and link chloroplastic-encoded gene expression with downstream photobiology.