

**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 immensely 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
rETR _{max}	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-quantitative 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
T _m	Melting temperature
TMM	Trimmed Mean of M-Values
TPM	Transcripts Per Million
UNIX	UNIX programming environment
UPLC	Ultra Performance Liquid Chromatography
UV	Ultra Violet
Y _i	Initial quantum yield of photosystem II

List of figures	Pg.
Figure 1: Illustrative diagram collectively showing the specialised traits of seagrasses, which allow for seagrasses to live a submerged life in the coastal marine environment. Information sources: Marbà et al., 2002; Ackerman, 2006; Larkum et al., 2006; Touchette, 2007; Broderson et al., 2015; Hasler-Sheetal et al., 2015; Olsen et al., 2016. Illustrative model concept is based on the <i>Zostera</i> genus.	8
Figure 2: The advances in seagrass molecular profiling and omics to date. The technologies that have been utilised are shown along the bottom of the illustration, whilst types of study conducted to date are shown at the top of the illustration.	9
Figure 3: The physical light environment of A. terrestrial plants and, B. sub-tidal marine plants. Black arrows represent light scattering and attenuation processes. Spectral light availability is represented by a wavelength spectrum bar.	43
Figure 4: The light dependent reaction of photosynthesis (Z-scheme) in higher plants. PQ = Plastoquinone; Cyt.b6f = Cytochrome b6/f complex; PC = Plastocyanin. P680 refers to photosystem II and P700 refers to photosystem I. Positive (+) characters refer to excited state of the photosystems.	45
Figure 5: Schematic diagram of chlorophyll excitation in response to high light, the various fates of the excited chlorophyll molecule are represented. In association with NPQ, the role of the psbS protein and the xanthophyll cycle are shown.	49
Figure 6: Experimental set-up of the artificial mesocosm study with varying light intensities: $\sim 1,000 \mu\text{mol photons m}^{-2} \text{ s}^{-1}$; $\sim 250 \mu\text{mol photons m}^{-2} \text{ s}^{-1}$; and $\sim 10 \mu\text{mol photons m}^{-2} \text{ s}^{-1}$. Aquarium related equipment is labelled in the legend above. Temperature was kept at 24°C and salinity between 32 and 36 units.	64
Figure 7: A schematic flow diagram representing the methodology used to assemble, characterise and profile the <i>Z. muelleri</i> transcriptome.	66
Figure 8: Venn diagrams representative of the GO terms enriched in each of the light conditions. Labelling units are in $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$.	71

Figure 9: The transcription factors family/types (represented by a minimum of 50 occurrences) identified in the <i>Z. muelleri</i> transcriptome.	74
Figure 10: Pie chart showing the composition (in %) of SSR sequence types in <i>Z. muelleri</i> .	75
Figure 11: Scaled topographical map of Narrabeen Lakes. Sampling site is indicated by a blue star. North arrow and scale represented. Map created using ArcGIS online (esri).	92
Figure 12: Experimental protocol and layout. A. Mesocosm layout detailing acclimation period, experimental treatments and sample collection time points. Light units are $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$. B. mRNA-Seq genome guided pipeline to determine transcriptome and differential gene expression, green reads = $200 \mu\text{mol photons m}^{-2} \text{ s}^{-1}$; red = $20 \mu\text{mol photons m}^{-2} \text{ s}^{-1}$ and orange = genome. C. Physiological and photobiology measurements recorded including the pigments analysed. Plants for RNA-sequencing, pigment analysis and rapid light curve analysis were all independent of one another.	93
Figure 13: A – G. Pigment concentrations ($\mu\text{g/ gram fresh weight}$ in control and light limited seagrass plants; $n = 3$). Significance is represented by an asterix (*). H. Relative electron transport rate maximum of control and light limited plants across Days 2 and 6 ($n = 5$). All error bars are representative of standard deviation of mean. Ctl = Control, LL = Light limited.	99
Figure 14: Differential gene expression represented by hierarchical clustered heatmaps. Normalised LogCPM gene counts are centered and z-scaled. D2 = Day 2, D6 = Day6. Light limited replicates (L1-3; $20 \mu\text{mol photons m}^{-2} \text{ s}^{-1}$) and control plant replicates (C1-3; $200 \mu\text{mol photons m}^{-2} \text{ s}^{-1}$) are indicated.	101
Figure 15: Bootstrapped hierarchical clustering of correlated transcripts. A. A cluster of correlated transcripts that decreased in expression under light limitation. B. Three closely related clusters of transcripts that increased in expression under light limitation on Day 6.	102
Figure 16: Significant photosynthetic gene regulation on Day 2 (FDR threshold = 0.05; fold change of 1). Significant down-regulation of many genes can be observed. <i>CCSI</i> = Cytochrome C biogenesis protein; <i>NDF1</i> = NDH-DEPENDENT CYCLIC ELECTRON FLOW 1; <i>Cyt b6/f</i> =	104

Cytochrome b6/f sub-unit; *PGR5* = Proton Gradient Regulation 5; *LTD* = Light harvesting complex translocation defective protein; *OHP* = One helix proteins; *FtsH* = ATP-dependent Zinc metalloproteases; *YCF* = hypothetical chloroplast open reading frame genes; *psb* = photosystem two sub units; *ELIPs* = Early light induced protein genes.

Figure 17: Significant gene regulation (FDR = 0.05; fold change of 1) associated with the carotenoid, xanthophyll, tocopherol, chlorophyll and ABA synthesis pathways on Day 2 of light limitation. *PSY* = Phytoene synthase; *PDS* = Phytoene desaturase; *ZDS* = Zeta carotene desaturase; *CRTISO* = Carotenoid isomerase / Prolycopene isomerase; *LYCB* = Lycopene beta-cyclase; β -*OHase* = β -carotene hydroxylase; ϵ -*OHase* = ϵ hydroxylase; *ZEP* = Zeaxanthin epoxidase; *VDE* = Violaxanthin de-epoxidase; *CCDs* = Carotenoid cleavage dioxygenases; *NCED* = Nine cis-epoxycarotenoid dioxygenase; *POR* = Protochlorophyllide; *FLU* = Blue fluorescent light; *Chl.red B NOL* = Chlorophyllide B reductase NOL; *ChlH* = Mg chelatase sub-unit H; *DVR* = Divinyl chlorophyllide A 8-vinyl-reductase; *VTE* = Tocopherol cyclase; *gamma-TMT* = *Tocopherol O*-methyltransferase; *ABA 8'-OHase 3* = Absicic acid 8' hydroxylase 3; *NSY* = Neoxanthin synthase.

Figure 18: Relative ETR_{max} of *Z. muelleri* plants during pre-treatment and Day 4 in control (dark grey) and light limited treatments (white). Asterix (*) = statistical significance observed ($p < 0.05$) between treatments.

Figure 19: Gene stability measure of chloroplast-encoded reference genes only. Names and average M-stability values are shown on the plot.

Figure 20: Gene stability measure of chloroplast-encoded and nuclear-encoded reference genes combined. Names and average M-stability values are shown on the plot. *ndhC*, *rpoC2*, *rps11* and *rps2* are chloroplast-encoded genes. *GDPH* and *PP2Aa* are nuclear-encoded genes.

Figure 21: A. Target genes normalised to nuclear-encoded reference genes (Normalised Relative Quantity – NRQ). B. Target genes normalised to chloroplast-encoded reference genes. Asterix (*) and brackets indicate statistically significant difference between control and light limited (low light) *Z. muelleri* plants. $n = 3$.

Figure 22: Schematic diagram highlighting the trends of how <i>Z. muelleri</i> responds to light limitation over a period of 14 days.	156
A.1; Figure 1: Plant GO-Slim characterisation of the <i>Z. muelleri</i> transcriptome ‘Biological Processes’. A threshold of 250 or more unigenes per GO term was applied. Level 3 and below (Blast2GO software) are only shown.	168
A.1; Figure 2: Plant GO-Slim characterisation of the <i>Z. muelleri</i> transcriptome ‘Molecular Functions’. A threshold of 250 or more unigenes per GO term was applied. Level 3 and below (Blast2GO software) are only shown.	169
A.1; Figure 3: Plant GO-Slim characterisation of the <i>Z. muelleri</i> transcriptome ‘Cellular Compartments’. A threshold of 250 or more unigenes per GO term was applied. Level 3 and below (Blast2GO software) are only shown.	170
A.2; Figure 1: RLC analysis of <i>Z. muelleri</i> plants from Narrabeen Lakes (n = 6) to determine control irradiance for the experiment during Australian winter, 2015. Standard deviation represented by error bars.	191
A.2; Figure 2: Odyssey light logger data from the experiment across 6 days: 18/09/15 to the 23/09/15. Mid-day spikes are caused by single time point sampling effort disturbance.	191
A.2; Figure 3: Chromatogram of the pigment peaks identified in HPLC analysis. Retention time (minutes) shown on X axis. Intensity (AU) shown on Y-axis.	192
A.2; Figure 4: Biological coefficient of variation between RNA-Seq samples within the experiment. Common dispersion is represented in green. A BCV value of 0.24 was obtained based on the common dispersion. Trended and tagwise dispersions are also represented; however, common dispersion is used to gauge variation within EdgeR (Mccarthy et al. 2012).	192
A.2; Figure 5: Redundant or duplicated transcripts that are present in the draft genome of <i>Z. muelleri</i> (Lee et al. 2016) and are correlated in expression. The correlation of expression is plotted for 1,051 differentially expressed transcript models in this study versus the percent nucleotide identity against redundant, similar or duplicated transcript models in the	193

draft genome.

A.2; Figure 6: Multi-dimensional scaling (MDS) plot showing similarities between RNA samples (Leading LogFC dimension 1 Vs Leading logFC dimension 2). D2 = Day 2, D6 = Day 6, L1-L3 = LL plants 1 – 3, C1-C3 = Control plants 1-3. **193**

A.2; Figure 7: DNA sequence motifs detected by MEME in the putative upstream promoter regions of A. Genes down-regulated under light limitation by Day 2 and, B. Genes up-regulated under light limitation by Day 6. **194**

List of tables	Pg.
Table 1: Known to us, the molecular profiling and omic studies published since the establishment of the seagrass NGS era in 2006. A comprehensive search was conducted on ‘Web of science’ for the topics ‘seagrass sequencing’, ‘seagrass gene expression’, ‘seagrass proteomics’, ‘seagrass metabolomics’. Additional studies based on personal knowledge are also provided. Last accessed: June, 2016.	11
Table 2: The number of polymorphic SSR loci and SNP loci validated in seagrass species to date. Organised by number of polymorphic loci per species. A search was conducted using ‘Web of Science’ with the search terms ‘microsatellite loci seagrass’, ‘microsatellite markers seagrass’, ‘polymorphic microsatellite seagrass’ and ‘polymorphic loci seagrass’. Last accessed: June, 2016.	25
Table 3: Summary of statistics of the original and final <i>Z. muelleri de novo</i> assemblies.	69
Table 4: Coverage distribution of the <i>Z. muelleri</i> transcriptome across related plant species. Proteome = number of genome predicted proteins, <i>Z. muelleri</i> unigenes =52,616. CRBB = Conditional Reciprocal Best BLAST. Reference proteomes were downloaded from ftp://ftp.ensemblgenomes.org/pub/plants/release-32 . The <i>Spirodela polyrhiza</i> proteome was downloaded from http://spirodelagenome.org , <i>Zostera marina</i> from http://uniprot.org .	69
Table 5: Number of unigenes distributed across GO categories putatively involved in light associated processes and light harvesting.	72
Table 6: Statistical comparisons between control and light limited plants on Days 2 and 6. Median, p-values, U values and Z-test scores are reported for Mann-Whitney-U statistical tests. Significant (P<0.05) results are indicated by an asterix (*). The number of biological replicates was n = 5, except for Day 2 when n = 4 for light limited plants (loss of one replicate sample).	98
Table 7: Chloroplast / nuclear-encoded reference genes, and chloroplast-encoded target genes validated and profiled in this study. Gene names are provided (abbreviated and in full), along with primer melting temperatures (T _m , °C), oligo sequences, amplicon size (AS; bp), amplification efficiencies (E; %) and standard curve co-efficient values (R ²).	131

Table 8: Statistical comparisons between control and light limited plants on Day 4. Median, p-values, U values and Z-test scores are reported for Mann-Whitney-U statistical tests. Significant results are indicated by an asterisk (*), n = 5.	136
Table 9: Significance of gene expression change between control and light limited plants when normalised to chloroplast-encoded (cp) reference genes and nuclear-encoded (nc) reference genes. P(MC) = P-value (multiple correction), T = pairwise t-test value. Statistical significance (p<0.05) is represented by an asterisk (*).	138
A.1; Table 1: Number of paired end reads (bp) in each library used to assemble the <i>Z. muelleri</i> transcriptome.	171
A.1; Table 2: Functional distribution hit analysis showing top 5 most represented plant species in DIAMOND protein alignment searches (E-value of 1.0×10^{-5}); 52,616 = total number of best hits retained.	171
A.1; Table 3: Enriched GO terms in $\sim 10 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ and $\sim 250 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ of light combined. GO = Gene Ontology, NS = Classification, Name = Name of GO term, FDR threshold = 0.05. BP = Biological Process; CC = Cellular Compartment and MF = Molecular Function.	172
A.1; Table 4: Enriched GO terms in $\sim 10 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ of light. GO = Gene Ontology, NS = Classification, Name = Name of GO term, FDR threshold = 0.05. BP = Biological Process; CC = Cellular Compartment and MF = Molecular Function.	174
A.1; Table 5: Enriched GO terms in $\sim 250 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ of light. GO = Gene Ontology, NS = Classification, Name = Name of GO term, FDR threshold = 0.05. BP = Biological Process; CC = Cellular Compartment and MF = Molecular Function.	179
A.1; Table 6: Enriched GO terms in $\sim 1,000 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ of light. GO = Gene Ontology, NS = Classification, Name = Name of GO term, FDR = 0.05. BP = Biological Process; CC = Cellular Compartment and MF = Molecular Function.	182
A.1; Table 7: PlantGSEA metabolic pathway enrichment of terms in $\sim 1,000 \mu\text{mol photons m}^{-2}\text{s}^{-1}$ of light. FDR threshold = 0.05.	184
A.1; Table 8: PlantGSEA metabolic pathway enrichment of terms in ~ 250	186

$\mu\text{mol photons m}^{-2}\text{s}^{-1}$ of light. FDR threshold = 0.05.	
A.1; Table 9: PlantGSEA metabolic pathway enrichment of terms in ~ 10 $\mu\text{mol photons m}^{-2}\text{s}^{-1}$ of light. FDR threshold = 0.05.	187
A.1; Table 10: The top ten most common SSR repeat motifs found within the <i>Z. muelleri</i> transcriptome.	189
A.2; Table 1: Xanthophyll and carotenoid pigments profiles on days 2 and 6 (n = 3). Significant results are highlighted with an asterix (*) character (P<0.05). Independent student t-tests were conducted.	195
A.2; Table 2: Chlorophyll pigments profiles on days 2 and 6 (n = 3). Significant results are highlighted with an asterix (*) character (P<0.05). Independent student t-tests were conducted.	195
A.2; Table 3: Number of paired end reads (bp) in each library used for the assembly of the genome-guided transcriptome.	196
A.2; Table 4: Top 10 species/ genus groups with functional best hits to the <i>Z.</i> <i>muelleri</i> transcriptome.	197
A.2; Table 5: Differentially regulated genes associated with photosynthesis and carbon fixation on Day 2. FC = Fold change, FDR = 0.05.	198
A.2; Table 6: Differentially regulated genes associated with photosynthesis and carbon fixation on Day 6. FC = Fold change, FDR = 0.05.	200
A.2; Table 7: Differentially regulated genes associated with photosynthetic pigments and ABA (abscisic acid) on Day 2. FC = Fold change, FDR = 0.05.	201
A.2; Table 8: Differentially regulated genes associated with photosynthetic pigments and ABA (abscisic acid) on Day 6. FC = Fold change, FDR = 0.05.	203
A.2; Table 9: Differentially regulated genes associated with light reception and signaling on Day 2. FC = Fold change, FDR = 0.05.	205
A.2; Table 10: Differentially regulated genes associated with light reception and signaling on Day 6. FC = Fold change, FDR = 0.05.	207
A.2; Table 11: Enriched GO terms in the up-regulated sub-set of genes in LL plants on Day 2. FDR = 0.05. NS = Classification; MF = Molecular Function; BP = Biological Process; CC = Cellular Compartment.	208
A.2; Table 12: Enriched GO terms the down-regulated sub-set of genes in LL plants on Day 2. FDR = 0.05. NS = Classification; MF = Molecular Function; BP = Biological Process; CC = Cellular Compartment.	209

A.2; Table 13: Enriched GO terms in the up-regulated sub-set of genes in LL plants on Day 6. FDR = 0.05. NS = Classification; MF = Molecular Function; BP = Biological Process; CC = Cellular Compartment.	212
A.2; Table 14: Enriched GO terms in the down-regulated sub-set of genes in LL plants on Day 6. FDR = 0.05. NS = Classification; MF = Molecular Function; BP = Biological Process; CC = Cellular Compartment.	214
A.2; Table 15: Enriched GO terms in the down-regulated sub-set of genes in LL plants on Day 6 compared to LL plants on Day 2. FDR = 0.05. NS = Classification; MF = Molecular Function; BP = Biological Process; CC = Cellular Compartment.	216
A.3; Table 1: Comparison of genome-guided assembly and <i>de novo</i> super assembly statistics.	218

CONTENTS

Thesis summary	XXI
PhD thesis aims and objectives	1
CHAPTER 1	2
Abstract	3
1.1 Introduction	4
1.2 Seagrasses: A unique group of plants	5
1.3 The current status of omics and molecular profiling in seagrass biology	5
1.4 Seagrass light perception and response at the molecular level	10
1.5 Carbon fixation in seagrasses – challenging old beliefs with new technology	13
1.6 Are stress and environmental response signatures between land and marine plants different?	14
1.7 Osmoregulation at the molecular level.....	16
1.8 Tolerance to anoxia and phytotoxic sediment.....	18
1.9 Other areas needing investigation and development.....	19
1.10 Genetic marker utilisation in seagrass biology	22
1.11 Development and advancement of molecular markers	22
1.12 SSR markers: High-resolution popularity	23
1.13 Is there a need for further advances in the field of genetic markers?.....	24
1.14 Concluding remarks	26
1.15 Acknowledgements	27
1.16 References	28
CHAPTER 2	40
2.1 Light and photosynthesis: terrestrial vs. marine environments	42
2.2 Organelles responsible for light absorption in higher plants.....	44
2.3 Light-dependent reaction of photosynthesis	44
2.4 Photosystem-light harvesting complexes	45
2.5 Chlorophyll pigments	46
2.6 Carotenoids and the xanthophyll pigments	47
2.7 The xanthophyll cycle and non-photochemical quenching	48
2.8 Anthocyanins.....	49
2.9 Sensing quality and quantity of environmental light	50
2.10 Common low light responses in plants	51
2.11 Our knowledge of light response in <i>Z. muelleri</i> to date	52
2.12 References	54
CHAPTER 3	60
Abstract	61
3.1 Introduction	62
3.2 Materials and methods	63
3.2.1 Sample collection, experimental layout and sample preparation.....	63
3.2.2 Quality checks and transcriptome assembly	65
3.2.3 Annotation and filtering of unigenes	66
3.2.4 Evaluation of the transcriptome assembly	67
3.2.5 Functional classification and detection of genetic elements.....	67
3.2.6 Enrichment analysis and in silico profiling of light responsive orthologs.....	68
3.3 Results	68
3.3.1 Transcriptome annotation and assessment.....	68
3.3.2 Functional classification and enrichment analysis.....	70
3.3.3 Metabolic pathway enrichment across light irradiances	72

3.3.4	Antennae complex genes involved in light capture and response	73
3.3.5	Identification of transcriptional regulators and microsatellites	73
3.4.	Discussion	75
3.5	Conclusion.....	80
3.6	Acknowledgements	81
3.7	References	82
CHAPTER 4	88
Abstract	89
4.1	Introduction	90
4.2	Materials and methods	91
4.2.1	Field analyses and sample transplantation.....	91
4.2.2	Aquarium set-up and acclimation period.....	91
4.2.3	Photo-biology measurements and rapid light curve analysis.....	94
4.2.4	Leaf tissue sampling and processing.....	94
4.2.5	Read processing, transcriptome assembly and annotation.....	95
4.2.6	Differential gene expression analysis and gene ontology enrichment.....	96
4.2.7	Statistical testing of photobiology and pigment data.....	97
4.3	Results	97
4.3.1	Photobiology	97
4.3.2	Pigment shifts observed through HPLC profiling	98
4.3.3	Transcriptome assembly	100
4.3.4	Differential expression analysis	100
4.3.5	Regulation of genes involved in photosynthesis and carbon fixation.....	103
4.3.6	Regulation of genes associated with pigment pathways.....	104
4.3.7	Regulation of genes involved in light perception and signaling.....	106
4.3.8	Gene Ontology (GO) enrichment analysis.....	106
4.3.8.1	Photosynthesis and Sugar Enriched GO terms	106
4.3.8.2	Light and pigment enriched GO terms.....	107
4.3.8.3	Secondary metabolism and ROS enriched GO terms	107
4.4	Discussion	108
4.5	Conclusion.....	116
4.6	Acknowledgements	118
4.7	References	119
CHAPTER 5	126
Abstract	127
5.1	Introduction	128
5.2	Materials and methods	129
5.2.1	Sample transplantation and experimental set-up	129
5.2.2	Rapid light curve assessment and photobiology	130
5.2.3	RNA sampling, extraction and quality checks.....	132
5.2.4	Prediction of chloroplast genes and primer design	132
5.2.5	Reverse transcription and qPCR protocol.....	133
5.2.6	RT-qPCR data analysis and normalisation	134
5.3	Results	134
5.3.1	Photobiology	134
5.3.2	Selection of reference genes	136
5.3.3	Target gene expression profiling to nuclear-encoded reference genes	138
5.4	Discussion	139
5.5	Conclusion.....	144

5.6	Acknowledgements	145
5.7	References	146
SYNTHESIS, OUTLOOK AND CONCLUSIONS		152
6.1	Overview	153
6.2	Photo-physiological acclimation occurs early in <i>Z. muelleri</i>	154
6.3	Both nuclear and chloroplast-encoded genes play important roles in <i>Z. muelleri</i> , in response to light limitation	154
6.4	Trends observed in <i>Z. muelleri</i> over 14 days in response to light limitation	155
6.5	Genome-guided assembly versus <i>de novo</i> transcriptome assembly; Research challenges and implications for the future	157
6.6	Future research direction from this thesis	160
6.6.1	The seagrass holobiont paradigm and consequences for <i>Z. muelleri</i> under light limited environments	160
6.6.2	Further investigation of light response and acclimation in <i>Z. muelleri</i>	161
6.6.3	Design and implementation of molecular markers for reactive monitoring of light-related stress through targeted gene expression assay .	161
6.7	Conclusions	162
6.8	References	164
APPENDIX 1		167
APPENDIX 2		190
APPENDIX 3		217
APPENDIX 4		219

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 chloroplast-encoded gene expression with downstream photobiology.