IDENTIFICATION AND CHARACTERIZATION OF THE ACTIVATED DEFENCE RESPONSE IN THE COMMERCIALLY IMPORTANT AGAROPHYTE, *GRACILARIA GRACILIS*, FOLLOWING EXPOSURE TO DISEASE ELICITORS

Christopher Shawn Ealand

January 2011

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Molecular and Cell Biology, Faculty of Science,

University of Cape Town, South Africa

TABLE OF CONTENTS

DECLARATION vi

ACKNOWLEDGEMENTS vii

List of Abbreviations viii

List of Figures x

List of Tables xiii

ABSTRACT xiv

CHAPTER 1 1

Literature Review 1

1.1 Introduction 1
1.2 The genus *Gracilaria* 3
1.3 Commercial cultivation of Gracilaria in South Africa 5
1.4 Important biotic and abiotic factors that impact marine aquaculture of *Gracilaria* 8
1.5 Mechanisms of disease resistance in macroalgae 10
1.5.1 Pathogen-induced defence in macroalgae and higher plants 11
1.6 Functional genomics and its role in the identification of macroalgal defence genes 17
1.6.1 High throughput DNA sequencing 17
1.6.2 Molecular tools for transcriptome analysis 18
1.6.3 Overview of cDNA microarray technology 19
1.7 Significance of this study 23

CHAPTER 2 25

Identification of differentially expressed genes in *Gracilaria gracilis* after exposure to disease elicitors using cDNA microarray technology 25

2.1 INTRODUCTION 25

2.2 MATERIALS AND METHODS 27
2.2.1 Construction of microarray 27
2.2.1.1 PCR amplification of cDNA 27
2.2.1.2 Purification of amplified cDNA products 28
2.2.1.3 Printing of microarray slides 29
2.2.2 Seaweed sample acquisition for exposure to disease elicitors 30
2.2.3 RNA extraction protocol 31
2.2.3.1 Quantitative and qualitative analysis of RNA 32
2.2.3.2 RNA amplification and aminoallyl labelling 33
2.2.4 Microarray hybridization of RNA targets to cDNA probes 34
2.2.4.1 Target hybridization mixture 34

2.2.4.2 Preparation of cover-slips for microarray hybridization 35
2.2.4.3 Pre-hybridization of microarray slides 35
2.2.4.4 Target and probe hybridization reaction 35
2.2.5 Image acquisition 36
2.2.5.1 Extraction of features from spots on the microarray 37
2.2.5.1.1 Creation of the GenePix array list (GAL) file 37
2.2.5.1.2 Extraction of foreground intensity pixels 37
2.5.1.1.3 Background correction and quality assessment of individual spots 38
2.2.6 Data processing, analysis and normalization 39
2.2.7 Identification of differentially expressed genes 39
2.2.8 Plasmid DNA isolation and sequencing of putative defence genes 40
2.2.9 Bioinformatics and functional annotation of cDNA sequences 41

2.3 RESULTS AND DISCUSSION 43
2.3.1 Array construction 43
2.3.2 RNA isolation and preparation of RNA targets 44
2.3.3 RNA amplification and aminoallyl labelling 46
2.3.4 Microarray hybridizations and pre-processing of scanned images 49
2.3.5 Normalization of microarray data 51
2.3.6 Statistical analysis and identification of differentially expressed genes 55
2.3.6.1 Sequencing of cDNA inserts and bioinformatics for functional annotation 58
2.3.7 Possible roles of functionally annotated *G. gracilis* genes in the context of activated defence 68
2.3.7.1 Significantly differentially expressed *G. gracilis* genes 69
2.3.7.2 *G. gracilis* genes that failed to meet the criteria for significant differential expression 80

2.4 CONCLUSION 84

CHAPTER 3 86

Validation of microarray experiment and investigation of transcriptional regulation using real-time PCR 86

3.1 INTRODUCTION 86

3.2 MATERIALS AND METHODS 89
3.2.1 Seaweed sample acquisition for exposure to disease elicitors 89
3.2.2 RNA extraction 90
3.2.2.1 Twenty four time-course experiment 90
3.2.2.2 Thirty minute time-course experiment 91
3.2.2.3 Assessment of DNAse efficacy and presence of contaminating gDNA 92
3.2.3 Quantitative and qualitative analysis of RNA 93
3.2.4 Conversion of total RNA to complementary DNA (cDNA synthesis) 93
3.2.5 Designing of qPCR primers 94
3.2.6 Identification of reference genes for normalization of qPCR data 95

3.2.7 Real-time PCR 96
3.2.8 Normalization and relative quantitation of gene expression 97
3.2.9 Statistical analysis and validation of microarray experiment 99

3.3 RESULTS 100
3.3.1 Extraction of total RNA 100
3.3.2 Assessment of DNAse efficiency and genomic DNA contamination in total RNA preparations 103
3.3.3 Identification of reference gene for normalization of 24 hour qPCR data 104
3.3.4 Identification of reference gene for normalization of 30 min qPCR data 107
3.3.5 Relative quantitation of putative defence genes during the 24 hour experiment 109
3.3.6 Relative quantitation of putative defence genes during the 30 min experiment 118
3.3.7 Validation of microarray experiment 125

3.4 DISCUSSION 126
3.4.1 Validation of microarray experiment 126
3.4.2 Analysis of transcriptional regulation of five genes involved in defence response 127

3.5 CONCLUSION 137

CHAPTER 4 139

Investigation of translational regulation of two putative defence genes using western hybridization analysis 139

4.1 INTRODUCTION 139

4.2 MATERIALS AND METHODS 142
4.2.1 Determination of full-length genes 142
4.2.2 PCR amplification of full-length genes 143
4.2.3 Cloning of open reading frames into protein expression vector 144
4.2.4 Recombinant protein expression 145
4.2.5 Affinity purification of recombinant proteins 147
4.2.6 Antibody production and determination of antibody titre 148
4.2.7 Purification of antibodies by polyethylene glycol precipitation 149
4.2.8 Seaweed sample acquisition for exposure to disease elicitors 149
4.2.9 Isolation of seaweed total protein 150
4.2.10 Determination of antibody specificity and optimal amount of *G. gracilis* protein for western hybridization analysis 150
4.2.11 SDS-PAGE and western hybridization analysis of the 24 hour time-course experiment 151
4.2.12 Densitometry analysis of western hybridization immunoblots 153
4.2.13 Statistical analyses of western hybridization analysis 153

4.3 RESULTS 154
4.3.1 Determination of full-length genes 154
4.3.2 PCR amplification of *G. gracilis* ORFs and cloning into pET29a protein expression vectors 160
4.3.3 Recombinant protein expression 162
4.3.4 Affinity purification of recombinant proteins 166
4.3.5 Polyclonal antibody production and determination of the titre and specificity 172
4.3.6 Western hybridization analysis of 24 hour time-course experiment 176

4.4 DISCUSSION 183

4.5 CONCLUSION 188

CHAPTER 5 189

General discussion and future work 189

5.1 General Discussion 189
5.2 Future Work 194

REFERENCES 197

APPENDIX A (Media, Buffers and Solutions) 235

APPENDIX B (DNA primer sequences and PCR cycle profiles) 247

APPENDIX C 249

C.1 Predefined filter for use in image processing 249
C.2 Script used for microarray data analysis 249
C.3 Multiple sequence alignments with PSP genes 252
C.4 Sequence alignment with *G. gracilis* and *C. crispus* peroxiredoxin-like genes 253
C.5 Transit peptide prediction using *G. gracilis* thioredoxin gene sequence 254

APPENDIX D 256

D.1 pDNR-Lib plasmid map 256
D.2 Creator™ SMART™ cDNA library construction kit protocol 257
D.3 pET29a (+) plasmid map 258

APPENDIX E (*Gracilaria gracilis* DNA sequences) 259

DECLARATION

I, Christopher Shawn Ealand, hereby:

(a) Grant the University free license to reproduce the above thesis in whole or in part, for the purpose of research;

(b) Declare that this thesis entitled:

"Identification and characterization of the activated defence response in the commercially important agarophyte, *Gracilaria gracilis*, following exposure to disease elicitors"

(i) is my own unaided work, both in conception and execution, and that apart from the normal guidance of supervisor, I have received no assistance apart from that stated below;

(ii) except as stated below, neither the substance or any part of the thesis has been submitted in the past, or is being, or is to be submitted for a degree at this University or any other University.

(iii) I am now presenting the thesis for examination for the Degree of PhD.

Signature

Date

ACKNOWLEDGMENTS

Firstly, and most importantly, I would like to express my gratitude to my Lord and Saviour Jesus Christ. Throughout my life and academic career, you have always guided me and given me the ability to achieve something this momentous. For all you have done and continue to do for me, may the knowledge I have gained from this journey be used to your glory.

I would also like to express sincere gratitude to my family – my parents, Clayton and Lesley Ealand, my brother André Ealand and my sister Sian Ealand. Thank you Da and Mum for giving me the opportunity and means to pursue this. Thank you all for the constant support and encouragement. I love you all very much and couldn't ask for a better family. This thesis is dedicated to all of you.

Thank you very much Vernon (*Associate Professor Vernon Coyne*) for offering me a place in your research group. Thank you for the guidance and expertise you always provided over the years and for the critical thinking required which has made this achievement possible. The conference in Israel is one of the highlights of my time in your lab. A sincere thank you is also extended to Dr. Suhail Rafudeen – without your positive advice and assistance, I really don't think I would have been so effective in most of my experiments. Thanks also to Dr. Brandon Weber – your assistance with my proteomics section in the final stages of my project was invaluable! Thank you to Wiesner Vos for helping with the statistical analysis of the microarray data. A general thank you is extended to all those who assisted me with various aspects during my lab work. Thanks to all my 201 & 217 lab-mates throughout the years (Brett, Phil, Kim, Rob, Sue, Tanya, Dave H, Dave R, Rene, Amy, Roslyn, Bridget, Kyle, Rogan, Jarrod, Bronwyn, Bronwen, Taryn, Jerry, Karusha and Ayesha). Pub Fridays, golf "tournaments" and camping trips in weather not dissimilar to Hurricane Katrina really made our lab one of the best. I wish you all success in your future endeavours. To the rest of my fellow colleagues and friends in the department of Molecular and Cell Biology – thanks for the experiences and good memories.

Lastly, I must acknowledge my sources of funding over the years. Thank you to the Council for Scientific and Industrial Research (CSIR), the University of Cape Town and the National Research Foundation (NRF). Without your assistance, I would have literally not been able to do this.

List of Abbreviations

α	alpha
β	beta
°C	degrees Celsius
λ	lambda
µg	microgram(s)
µl	microlitre(s)
µm	micromole(s)
µmol	micromolar
ASW	artificial Seawater
aRNA	antisense RNA
BLAST	basic local alignment search tool
BLASTn	basic local alignment search tool nucleotide
BLASTx	basic local alignment search tool nucleotide translated
BSA	bovine serum albumin
bp	base pair(s)
CAPAR	Cape Array Opportunities
cDNA	complementary DNA
Cm	chloramphenicol
Ct	cycle threshold
CY-3/5	cyanine (fluorescent dye label)
DEPC	diethyl pyrocarbonate
dF/dT	first derivative of fluorescence / first derivative of temperature
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
dNTP	deoxyribonucleoside triphosphatases (dATP, dCTP, dGTP & dTTP)
DPI	diphenylene-iodonium
dUTP	deoxyuridine triphosphate
EBI	European bioinformatics institute
EDTA	ethylenediaminetetraacetic acid (disodium salt)
EST	expressed sequence tag
FAO	Food and Agriculture Organization (of the United Nations)
g	standard gravitational acceleration (9.81 m/s^2)
GOI	gene of interest
h	hour(s)
HR	hypersensitive response
IVT	*in vitro* transcription
Kan	kanamycin
kDa	kilodalton(s)
kb	kilobase(s)
L	litre(s)

LA	Luria agar
LB	Luria broth
MA	marine agar
MB	marine broth
MCS	multiple cloning site
mg	milligram(s)
min	minute(s)
mJ	milli Joules
ml	millilitre(s)
MOPS	(3-[N-morpholino]propane-sulfonic acid)
mRNA	messenger RNA
mV	milliVolt(s)
MW	molecular weight
NCBI	National Centre for Biotechnology Information
ng	nanogram(s)
nm	nanometer(s)
OD	optical density
ORF	open reading frame
PBS	phosphate buffered saline
PCD	programmed cell death
PCR	polymerase chain reaction
pdi	post disease induction
PMT	photo multiplier tube
pmol	picomole(s)
qPCR	quantitative real-time PCR
RACE	rapid amplification of cDNA ends
RG	reference gene
RNA	ribonucleic acid
RNAse	ribonuclease
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
SAR	systemic acquired resistance
SDS	sodium dodecyl sulphate (sodium lauryl sulphate)
SEM	standard error of means
SSC	sodium chloride tri-sodium citrate buffer
TAE	Tris-acetate-EDTA buffer
TE	Tris-EDTA buffer
TIFF	Tagged Image File Format
Tris	tris(hydroxymethyl)aminomethane
U	unit(s)
UV	ultraviolet
V	volts
wt	weight
YE	yeast extract

List of Figures

Figure 1.1	Basic frond morphology in *Gracilaria*.	4
Figure 1.2	Map of Saldanha Bay / Langebaan Lagoon system on the west coast of South Africa	5
Figure 1.3	*Gracilaria gracilis* beach-cast collected for the local industry in Saldanha Bay	6
Figure 1.4	Current model of the activated defence response in macroalgae	12
Figure 1.5	Simplified schematic depiction of the cDNA microarray procedure	21
Figure 2.1	Design of *G. gracilis* cDNA microarray	30
Figure 2.2	Local background correction algorithm used by GenePix Pro software	38
Figure 2.3	Purified PCR products from the disease elicitor EST *G. gracilis* cDNA library	43
Figure 2.4	Total RNA preparations extracted from *G. gracilis* samples under control and experimental conditions	45
Figure 2.5	Representative TIFF image corresponding to the scanned image of a cDNA microarray	50
Figure 2.6	Density plot of raw densities for the individual Cy3 (green) and Cy5 (red) channels for all three microarrays before data normalization	52
Figure 2.7	Density plot after robust-spline normalization to correct within each of the three microarrays	54
Figure 2.8	Density plot after robust-spline and Aquantile normalization to correct within and between the three microarrays	55
Figure 2.9	Pie chart representing the different molecular functions significantly regulated in *G. gracilis* following exposure to disease elicitors	64
Figure 3.1	Equation representing the Pfaffl mathematical model for calculating relative expression ratios in qPCR	98
Figure 3.2	Assessment of RNA quality and integrity used in the 24 hour time-course qPCR experiment	101
Figure 3.3	Assessment of RNA quality and integrity used in the 30 min time-course qPCR experiment	102
Figure 3.4	PCR amplification of 18S rRNA gene product to test the amount of residual gDNA in the total RNA samples	104
Figure 3.5	Average Ct values for *RG 329* over the 24 hour time-course	105
Figure 3.6	Average Ct values for *RG 245* over the 24 hour time-course	106
Figure 3.7	Average Ct values for *RG 801* over the 24 hour time-course	106
Figure 3.8	Average Ct values for *18S rRNA RG* over the 24 hour time-course	107

Figure 3.9	Average Ct values for *RG 245* over the 30 min time-course	108
Figure 3.10	Average Ct values for *RG 329* over the 30 min time-course	108
Figure 3.11	Average Ct values of *RG 245* expression in all three biological repeats over the 24 hour time-course	109
Figure 3.12	Melt curve analyses of GOI and RG primer sets	110
Figure 3.13	Gene expression data for PRX2F in the 24 hour qPCR experiment	113
Figure 3.14	Gene expression data for PSP in the 24 hour qPCR experiment	114
Figure 3.15	Gene expression data for PSD in the 24 hour qPCR experiment	116
Figure 3.16	Gene expression data for serine protease-like gene in the 24 hour qPCR experiment	117
Figure 3.17	Gene expression data for thioredoxin in the 24 hour qPCR experiment	118
Figure 3.18	Average Ct values of *RG 329* expression in all three biological repeats over the 30 min time-course	119
Figure 3.19	Melt curve analyses of RG 329 primer set	120
Figure 3.20	Gene expression data for PRX2F in the 30 min time-course qPCR experiment	122
Figure 3.21	Gene expression data for PSP in the 30 min time-course qPCR experiment	123
Figure 3.22	Gene expression data for PSD in the 30 min time-course qPCR experiment	123
Figure 3.23	Gene expression data for serine protease-like gene in the 30 min time-course qPCR experiment	124
Figure 3.24	Gene expression data for thioredoxin in the 30 min time-course qPCR experiment	124
Figure 3.25	Validation of microarray using qPCR	125
Figure 4.1	Central pathway of molecular biology depicting the biological parameters that influence the mRNA-protein correlation	140
Figure 4.2	Prediction of open reading frames present in the cDNA clone harbouring the *G. gracilis* PRX2F gene sequence	155
Figure 4.3	Multiple sequence alignment of eleven PRX2F protein sequences	156
Figure 4.4	Prediction of the open reading frames present in the cDNA clone harbouring the *G .gracilis* serine protease-like gene	158
Figure 4.5	Multiple sequence alignment of four serine protease-like protein sequences	159
Figure 4.6	PCR amplification of PRX2F and serine protease-like open reading frames using proof-reading DNA polymerase	160
Figure 4.7	Restriction digest analyses of pET29a (+) vector constructs used for protein expression studies	161

Figure 4.8	Coomassie stained SDS-PAGE gel showing expression of the His-tagged PRX2F protein in *E. coli* in the small scale study	163
Figure 4.9	Coomassie stained SDS-PAGE gel showing expression of the His-tagged serine protease-like protein in *E. coli* in the small scale study	164
Figure 4.10	Coomassie stained SDS-PAGE gel showing expression of the His-tagged PRX2F protein in *E. coli* in the large scale study	165
Figure 4.11	Coomassie stained SDS-PAGE gel showing expression of the His-tagged serine protease-like protein in *E. coli* in the small scale study	166
Figure 4.12	Elution profile of recombinant PRX2F following Ni+ affinity chromatography	168
Figure 4.13	Coomassie stained SDS-PAGE gel showing the eluted fractions of purified His-tagged PRX2F recombinant protein	169
Figure 4.14	Elution profile of recombinant serine protease-like protein following Ni+ affinity chromatography	170
Figure 4.15	Coomassie stained SDS-PAGE gel showing the eluted fractions of purified His-tagged serine protease-like recombinant protein	171
Figure 4.16	Western hybridization immunoblot of the purified Histidine-tagged recombinant proteins	172
Figure 4.17	'Dot-blot' western hybridization to establish which bleed contained the highest titre of anti-PRX2F antibodies	173
Figure 4.18	'Dot-blot' western hybridization to establish which bleed contained the highest titre of anti-serine protease-like antibodies	174
Figure 4.19	Western hybridization analyses to determine the specificity of polyclonal antibodies	176
Figure 4.20	Coomassie stained SDS-PAGE gels to establish the quality and integrity of *G. gracilis* total protein extracts from the 24 hour time-course experiment	177
Figure 4.21	Western hybridization analysis of *G. gracilis* PRX2F protein levels over 24 hours following exposure to disease elicitors	179
Figure 4.22	Western hybridization analysis of *G. gracilis* Rubisco protein (loading control) over the 24 hour time-course experiment	180
Figure 4.23	Western hybridization analysis of *G. gracilis* serine protease-like protein levels over the 24 hours following exposure to disease elicitors	181
Figure 4.24	Western hybridization analysis of *G. gracilis* Rubisco protein (loading control) over the 24 hour time-course experiment	182

List of Tables

Table 2.1	Constituents of PCR master mix used for colony PCR	28
Table 2.2	Reaction constituents for the BigDye Terminator v.3.1 Cycle Sequencing Kit	41
Table 2.3	Purity and yield of the RNA used for the RNA amplification and subsequent microarray analysis	46
Table 2.4	aRNA yields synthesized from 1.2 µg total RNA for each sample (experiment and control) using the MessageAmp^{TM} II aRNA kit	48
Table 2.5	Characteristics of the fluorescently labelled, purified aRNA used in microarray hybridizations	49
Table 2.6	*G. gracilis* genes differentially expressed in response to exposure to disease elicitors for 24 hours	60
Table 3.1	Gene specific qPCR primers for five putative defence genes and the expected amplicon sizes	94
Table 3.2	Gene specific qPCR primers for four reference genes with expected amplicon sizes	96
Table 3.3	PCR amplification efficiency and correlation efficiency for each GOI and RG 245 primer set for the 24 hour time-course experiment	111
Table 3.4	PCR amplification and correlation efficiencies for each GOI and RG 329 for the 30 min time-course experiment	121
Table 4.1	PCR primers used for amplification and cloning of the *G. gracilis* genes into the pET-29a (+) expression vector	143

ABSTRACT

Southern Africa has a diverse seaweed flora with a wide variety of marine habitats. Commercial interest in these seaweed resources has prompted various studies aimed at increasing knowledge of the biochemical and molecular mechanisms active during abiotic and biotic environmental stresses. Such knowledge is essential for the development of successful seaweed culture systems which may have to circumvent over-exploited or depleted natural seaweed supplies. The Gracilariaceae (Rhodophyta) has emerged as one of the families possessing economic potential as a source of agar and as a potential feed in abalone aquaculture settings. Until recently, macroalgal defence against microorganisms has been regarded as constitutive. However, evidence supports the existence of pathogen-activated macroalgal defence. Furthermore, functional similarities have been observed between the molecular components of defence between macroalgae and higher eukaryotic organisms. Homologies exist in the primary and secondary defence-activating signals, as well as in the enzymes and cellular responses that are activated. The overall lack of knowledge with respect to characterization of macroalgal defence responses has been attributed to the limited amount of DNA sequence information and/or functional annotation of these sequences in publicly available Genebank databases. Two cDNA libraries derived from a red macroalga, *Gracilaria gracilis*, subjected to biotic and abiotic stress (exposure to disease elicitors and nitrogen limitation, respectively) were used to construct a small-scale cDNA microarray. An expression profiling approach used to identify *G. gracilis* genes transcriptionally regulated following exposure to disease elicitors (agar and agar oligosaccharides) is described here. Gene expression profiles were studied 0 and 24 hours after addition of disease elicitors to the culture media surrounding the macroalga. The changes in gene expression were monitored using cDNA microarrays with 1620 different cDNAs potentially representing 1620 unique genes (since cDNAs were not functionally annotated prior to cDNA microarray construction). A total of 51 *G. gracilis* genes were differentially expressed 24 hours post exposure to disease elicitors. Each corresponding cDNA was sequenced and subjected to BLAST analysis in order to functionally annotate the *G. gracilis* genes. Significant transcriptional regulation of *G. gracilis* genes after 24 hours of exposure to the disease elicitors appeared to involve genes encoding proteins involved in stress responses (9%), protein processing (3%), cell structure (9%),

metabolism (14%), respiration (3%), protein synthesis (6%) and DNA modification (3%). Approximately half of the transcriptionally regulated genes identified encoded proteins with unknown or novel functions. Down-regulated transcripts encoded various proteins including thioredoxin, asparaginase, phosphoglycolate phosphatase, polyubiquitin and a peroxiredoxin-like protein. Up-regulated transcripts encoded proteins such as tRNA-dihydrouridine synthase, a high light inducible protein and a SNF2 family chromodomain-helicase DNA-binding protein. However, the responses of five genes (encoding proteins for thioredoxin, a peroxiredoxin-like protein, phosphoserine phosphatase, phosphatidylserine decarboxylase and a serine protease-like protein, respectively) were selected for verification by real-time PCR. In addition to monitoring expression at 0 and 24 hours post exposure to disease elicitors, gene expression profiles were assessed during the first 30 minutes as well as at 1, 8 and 12 hours. No significant changes in gene expression were observed during the first 30 min of exposure to disease elicitors. In contrast, the most dynamic changes in gene expression were observed between 8 and 12 hours after exposure to disease elicitors. A gene encoding an antioxidant peroxiredoxin-like protein was significantly up-regulated whereas transcription of the genes encoding phosphatidylserine decarboxylase, phosphoserine phosphatase, thioredoxin and a serine protease-like protein was generally repressed. To our knowledge, this study represents the first analysis of gene expression using cDNA microarrays in the red macroalga *G. gracilis*. Western hybridization analysis was used to establish whether the observed changes in gene expression following exposure to disease elicitors positively correlated to changes at the protein level. The open reading frames of two genes (encoding a peroxiredoxin-like protein or a serine protease-like protein) were cloned for expression as recombinant proteins in *Escherichia coli*. Polyclonal antibodies raised to both recombinant *G. gracilis* proteins were raised in rabbits. Western hybridizations revealed a positive correlation between the levels of mRNA and protein for the peroxiredoxin-like gene which suggested that *G. gracilis* actively generated antioxidant proteins to ameliorate the oxidative burst as a result of exposure to disease elicitors. In contrast, western hybridization analysis failed to detect any changes in the levels of serine protease-like protein in *G. gracilis* samples exposed to disease elicitors whereas the corresponding mRNA transcripts appeared to be repressed. In conclusion, this study has proved that *G. gracilis* is able to detect disease elicitors and subsequently generate an oxidative burst, presumably to act as

the first line of defence against any invading pathogen. In addition, reactive oxygen species may function as molecular signals that mediate transcriptional regulation of defence genes. This study was limited by the low density of the cDNA microarray and could not be used as a global index of the *G. gracilis* transcriptome during an activated defence response. Nevertheless, similarities in the molecular mechanisms of activated defence between red macroalgae and higher plants were evident. In contrast, the inability to functionally annotate approximately 50% of the transcriptionally regulated *G. gracilis* genes (as determined by the cDNA microarray) suggests that aspects of the macroalgal molecular defence system may be unique.

CHAPTER 1

Literature Review

1.1 INTRODUCTION

The world's oceans simultaneously offer many challenges as well as undiscovered solutions with respect to the sustainability of Earth's resources and the life systems they support. At the same time, there is an increased realization that the oceans are far more vital to humankind than previously thought. The oceans harbour enormous biodiversity, with more than 80% of all life believed to exist beneath its surfaces (Patrzykat *et al*., 2003). With global awareness regarding climate change on the rise in the 21st century, the oceans have become essential variables in the regulation of the atmosphere, controlling weather patterns and the recycling of primary nutrients (Costanza *et al*., 1998; Peterson and Lubchenco, 1997). The notion of unlimited abundance of marine organisms has been a reckless assumption of our generation. Currently, the supplies from natural stocks have been unable to match the global demands of a steadily increasing population, either as a result of over-exploitation or complete depletion (FAO, 2006). Dwindling marine resources has dictated that the sole reliance on natural stocks is not feasible and as such, alternatives are urgently required. Chopin and Sawhney (2009) suggest that aquaculture has the potential to bridge the gap between supply and demand. Aquaculture is defined as 'the growing of aquatic organisms in fresh, brackish or seawater while marine aquaculture involves farming of marine organisms such as fish, molluscs, crustaceans and plants in controlled marine aquatic environments with some form of human intervention to enhance production' (Swift, 1998). Possible downsides of marine aquaculture lie in the threats it may pose to marine and coastal environments in the form of degradation of natural habitats, accidental release of alien organisms, nutrient and waste discharge and transmission of diseases to wild stocks (Phillips, 2009). Despite intense study of the biological organisms structuring marine ecosystems, the dynamics of intermittent, transient phenomena such as disease outbreaks in marine organisms is not clearly understood (Harvell *et al*., 1999). Furthermore, an increase in the incidence of

diseases affecting marine organisms has been documented in the past two decades (Lafferty *et al*., 2004; Epstein *et al*., 1998; Williams and Bunkley-Williams, 1990). Disease outbreaks are promoted by large fluctuations in environmental conditions that either increase prevalence and virulence of existing disease or facilitate new disease (Anderson, 1998). As expected, changes in climate patterns and increased ocean temperatures have thus played major roles in disease epidemics by undermining host resistance and facilitating pathogen transmission (Lafferty *et al*., 2004; Harvell *et al*., 1999). Human intervention has enhanced the global transport of marine species together with their pathogens (Carlton and Cohen, 2003; Carlton and Geller, 1993). Moreover, disease epidemics appear to be most common in aquaculture settings (Farley, 1992 and Ganzhorn *et al*,. 1992) but degradation of natural habitats and pollution may also facilitate disease outbreaks (Osterhaus *et al*., 1995). A major concern facing seaweed biologists is that the phenomenon of increased incidence in disease outbreaks has been documented in several species of commercially important marine algae (seaweeds) (Littler and Littler, 1998; Cole and Babcock, 1996).

An understanding of the potential impact of diseases on the seaweed industry is essential since several seaweeds are of ecological and commercial importance. Together with phytoplankton, seaweeds are required in nature for the production of oxygen in the ocean. Apart from serving as habitats for several marine macro- and microorganisms, seaweeds act as primary producers in the marine food chain (Chan *et al*., 2006). Currently, extracts of certain seaweeds are used as stabilizers, gelling agents or emulsifiers in everyday products such as dyes, toothpaste, salad dressings, flavoured milks, cosmetics, welding rods, inert carriers for drugs as well as culture media in microbiology. Some seaweeds have the capacity to sequester heavy metals from the water and can potentially be used in biomonitoring or for the bioremediation of such pollutants (Chan *et al*., 2006; Oliveira *et al*., 2000; Lee, 1999). Given the prospects and challenges facing the seaweed industry, namely the over-exploitation of natural stocks, implementation and increased reliance on aquaculture alternatives and the potential of increased occurrence of disease outbreaks, there is a need for interdisciplinary studies focussed on developing molecular tools to fully understand biotic interactions and the mechanisms of disease resistance.

1.2 The genus *Gracilaria*

Gracilaria are taxonomically classified as red macroalgae within the phylum Rhodophyta, class Florideophyceae, order Gracilariales and family Gracilariaceae. In addition, the family is comprised of several other genera including *Gracilariopsis*, *Polycavernosa* and *Hydropuntia*, all collectively classified as gracilarioid algae (de Oliveira and Plastino, 1994). Most of the algae in this phylum are typically red in colour but may manifest black, green or yellow forms due to the presence of the phycobilin pigments phycocyanin, phycoerythrin (most dominant) and allophycocyanin. According to Branch & Branch (1981), seaweeds of the genus *Gracilaria* have thalli that are bushy and rigid with relatively short branches (*Gracilaria beckeri*), or slender with ramifying, stringy branches (*Gracilaria gracilis*) (Figure 1.1). Gracilarioid algae are geographically widely distributed with the majority of species concentrated in the warmer waters of the northern hemisphere. A few species (e.g., *G. gracilis*) have been documented in temperate waters. Consequently, *G. gracilis*, which is of particular importance to this study, has been identified in Europe, Chile, and Argentina as well as along the coasts of southern Africa and Namibia (Oliveira *et al*., 2000). This broad distribution of gracilarioid algae has been attributed to the large number of species, including many sibling species and the broad salinity tolerances of some species (Oliveira *et al*., 2000). Temperature has also been considered a major environmental factor controlling geographical distribution of gracilarioid algae (de Oliveira and Plastino, 1994). Although optimal growth temperatures are approximately 20 – 25°C (Critchely, 1993; de Oliveira and Plastino, 1994), certain *Gracilaria* species can survive temperatures from 35°C (Yokoya and Oliveira, 1992) down to freezing and can even withstand being frozen for a few months (Titlyanov *et al*., 1995). With respect to vertical distribution, most species are found in the lower intertidal region where strong surf is uncommon. Only a few species are found in deeper waters or in areas prone to long periods of exposure to air (Oliveira *et al*., 2000). *Gracilaria* often have to survive freshwater dilutions, high inputs of nutrients and low-water motion, in combination with high temperatures and even burial in sand (Santelices and Doty, 1989).

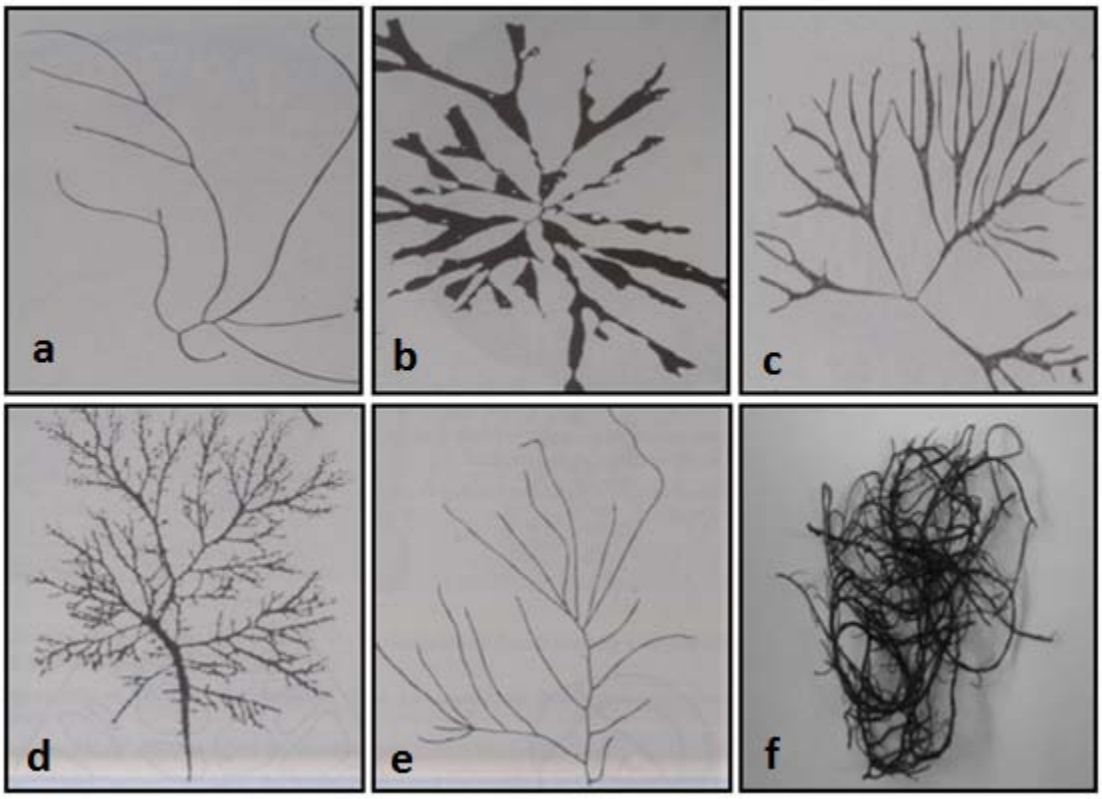


Figure 1.1 Basic frond morphology in *Gracilaria*. (a) *G. verrucosa* (Plastino, 1985); (b) *G. mammilaris* (Plastino, 1985); (c) *G. foliifera* (Plastino, 1985); (d) *G.* cervicornis (Plastino, 1985); (e) *G. tenuifrons* (Plastino, 1985); (f) *G. gracilis*.

Gracilaria species contribute significantly to world-wide agar production. Food grade and sugar reactive agar are of particular commercial importance (Armisen, 1995; Murano, 1995). The main economically important species required for the production of agar are *G. chilensis* and *G. gracilis* (Oliveira *et al*., 2000). Besides being agarophytes, *Gracilaria* species have been used for human consumption (Levring *et al*. 1969), as food for invertebrate cultivation (Chiang, 1981), as fertilizers (Zaneveld, 1959), in the pharmaceutical industry (Chapman, 1950), in the tertiary treatment of sewage (Ryther *et al.*, 1979) and for biogas production (Hanisak, 1981). As with other commercially viable marine organisms, the need for *Gracilaria* cultivation has arisen

as a consequence of the limitations and over-exploitation of natural stocks which has been surpassed by the growing demand for agar.

1.3 Commercial cultivation of *Gracilaria* in South Africa

Saldanha Bay and Langebaan lagoon is a system situated on the southwest coast of South Africa, approximately 18° E 33° S (Figure 1.2). Since World War II the region has been an important *Gracilaria* producer (Rotmann, 1990). Raw material, in the form of beach-cast *G. gracilis* (Stackhouse) Steentoft, Irvine et Farnham has been collected commercially (Figure 1.3) since the 1950s (Anderson *et al*., 1989; Rotmann, 1990). In the 1960s, two agar factories were built to process and produce agar where yields of over 1000 tons dry weight (d wt) of material were typical (Isaac, 1956).

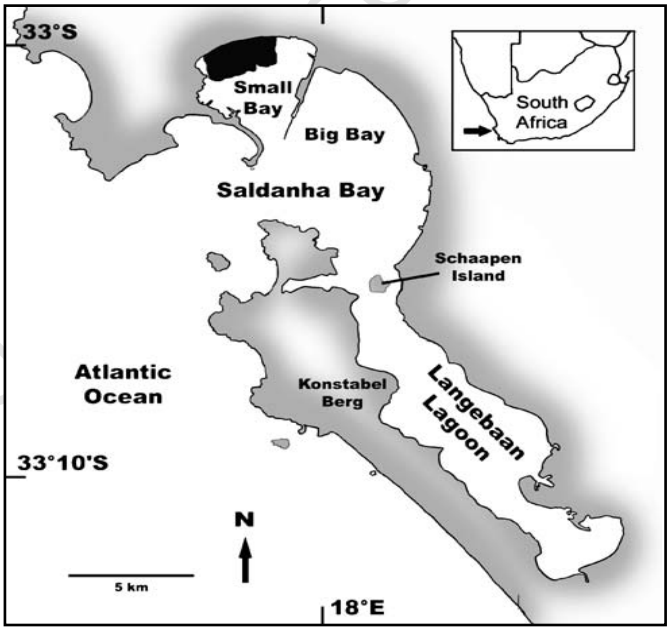


Figure 1.2 Map of Saldanha Bay / Langebaan Lagoon system on the west coast of South Africa. Small Bay is enclosed by breakwater and ore-jetty. The dark patch in Small Bay represents the distribution of gracilarioid algae (Rothman *et al*., 2009).

Figure 1.3 *Gracilaria gracilis* beach-cast collected for the local industry in Saldanha Bay in the 1960s to early 70s (Robert Anderson; http://www.algaebase.org).

In 1974 however, the industry experienced a major decline in natural *G. gracilis* stocks, mainly as a result of the construction of an ore jetty and break-water (Figure 1.2). It was believed that the construction generated changes in the water flow characteristics within the bay which promoted the development of strong thermal stratification of the water column (Anderson *et al*., 1996). Consequently, oligotrophic surfaces characterized by a low accumulation of dissolved nutrients and high oxygen content due to low organic content, were prevalent in summer months. The unavailability of nutrients (particularly nitrogen) and increased temperature led to the collapse of the *Gracilaria* resource. In addition, Schroeder *et al*. (2003) suggested that the effect of pathogenic bacterial infection could not be discounted. The authors observed a positive correlation between the presence of agarolytic epiphytes on the thallus surface and bacterial

pathogenicity. Similarly, Wheeler *et al*. (1979) noted a similar biological phenomenon when seaweeds appeared more susceptible to biotic infections (fungal, bacterial and viral) when temperatures were high and inorganic nutrient levels were low. After a few years of no or very little commercial production in Saldanha Bay, the *Gracilaria* resource slowly started to recover but has never returned to levels comparable to those observed before the construction of the ore jetty. A subsequent collapse of the seaweed resource during the late 1980s, ascribed to herbivorous fish, keyhole limpets and urchins, resulted in the almost complete depletion of *G. gracilis* in the bay (Anderson *et al*., 1993).

Despite the observed collapse of the natural *G. gracilis* resource in Saldanha Bay and extensive research toward establishing a sustainable commercial alternative, there is currently no marine aquaculture of this macroalga locally. However, successful marine aquaculture of *G. gracilis* has been achieved in neighbouring Namibia for several years (Rothman *et al*., 2009) which suggests that Saldanha Bay could be affected by the oligotrophic nature of the water system to a significant degree. In Saldanha Bay, natural stocks of *G. gracilis* have been sporadic as evident by a significantly reduced or even complete lack of beach-cast in the last decade. Furthermore, our research group has observed a major shift in the red macroalgal populations to that of predominantly *Gracilariopsis longissima*. Phenotypically, *G. longissima* is almost indistinguishable from *G. gracilis* but as an agarophyte, *G. longissima* generally provides a much lower quality and yield of agar relative to *Gracilaria* species (Wakibia *et al*., 2001; Rebello *et al*., 1997). Interestingly, wash-ups of *G. longissima* have also been documented in the Langebaan Lagoon (sourced by our research group) as well as in nearby St. Helena Bay, almost 100 km away. This suggests that the natural *G. gracilis* stocks along the west coast of southern Africa have never completely recovered from the collapse in the 1980s. Currently all (if any) of the high quality dried *Gracilaria* from Saldanha Bay is exported to Japan, Korea and Chile for agar processing, while the lower quality product is delivered to Namibia (Rotmann, 1990; Anderson *et al*., 2003). Relying on these unpredictable natural *Gracilaria* stocks to sustain a commercial agar industry will definitely not be feasible. The only long-term solution for the development of a sustainable *Gracilaria* industry in southern Africa is thus open-water marine

aquaculture. Several studies have shown promise in this regard (Wakibia *et al*., 2001 and Anderson *et al*., 1996). As with agriculture of land-based plants, successful cultivation of any given species firstly depends on a thorough understanding of the plant's biology, physiology and biochemistry. Thus, marine aquaculture of *G. gracilis* will only advance through research directed towards understanding its developmental (life history), biochemical, physiological and genetic characteristics as well as the biotic and abiotic interactions between host and pathogen under cultivation conditions.

1.4 Important biotic and abiotic factors that impact marine aquaculture of *Gracilaria*

The methods of *Gracilaria* cultivation currently employed are comprehensively reviewed by Oliveira *et al*. (2000). As highlighted by the authors, several crucial variables dictate the success or failure of such cultivation practices, including:

i) The availability of essential nutrients;
ii) The presence of epiphytes which could adversely affect the seaweed host by competing for light and nutrients, damaging the thallus due to penetration of rhizoids or the production of toxic allelochemicals;
iii) The effect of grazers, such as fish and invertebrates and their potential damage to cultivated beds;
iv) The effect of pathogens and the potential diseases they induce.

Of particular significance to this study is the potential impact of pathogens on the fitness of *Gracilaria* species during cultivation. More specifically, the study aims to characterize some of the molecular interactions that occur between *G. gracilis* and its pathogens when an activated defence response is triggered in the seaweed. The importance of elucidating macroalgal defence responses is reflected by the commencement of similar studies as early as the early 1990s (Correa and Craigie, 1991). In most cases, outward symptoms of disease infection manifest as thallus whitening, tissue softening and ultimate decay (Oliveira *et al*., 2000). *Gracilaria* species

naturally harbour a range of non-pathogenic epiphytes, including several species of bacteria (Correa and Craigie, 1991). Under certain conditions, these microorganisms may be involved in a mutualistic relationship providing the seaweed with dissolved inorganic nitrogen (Bird and Benson, 1987), growth factors, nutrients or protection from other bacteria (Zheng *et al*., 2005; Weinberger *et al*., 1997). Adverse environmental conditions such as low nutrients and increased temperature, as in the case of Saldanha Bay, may alter the intricate balance between host and epiphyte and lead to pathogenesis (Jaffray *et al*., 1997). This has been shown in the ability of certain micro-organisms isolated from healthy seaweed to degrade components of the algal cell wall (Jaffray and Coyne, 1996; Fujita, 1973). Various studies implicate agarolytic bacteria (Jaffray *et al*., 1997; Correa, 1996; Jaffray and Coyne, 1996; Friedlander and Levy, 1995; Weinberger *et al*., 1994; Friedlander and Gunkel, 1992 and Lavilla-Pitogo, 1992), nonagarolytic bacteria (Weinberger *et al*., 1997) and a transmissible, endophytic amoeba in macroalgal diseases (Correa and Flores, 1995).

The presence of various epiphytes has been established as a potential source of disease. When combined with a likely increase in the incidence of disease outbreaks in marine aquaculture settings, failure to incorporate the probable impacts of pathogens and diseases on sustaining a *G. gracilis* industry in southern Africa could be detrimental (Potin, 2008). Several studies emphasise the potentially harmful effects of diseases in aquaculture farms. For example, the development of another red macroalga (*Porphyra yezoensis*) industry led to substantial loss of cultivated crops due to disease outbreaks. The impact of the loss was so significant that it was enough to warrant prioritization of pathological studies (Fujita *et al*., 1972; Suto *et al*., 1972). A similar phenomenon was observed in the farming of *Laminaria japonica* (Ishikawa and Saga, 1989), *Eucheuma* (Ask and Azanza, 2002) and *Kappaphycus* (Hurtado *et al*., 2006). Molecular research focussed on the biotic interactions between pathogens and *G. gracilis* should therefore be prioritised in order to develop a sustainable local *Gracilaria* marine aquaculture industry.

1.5 Mechanisms of disease resistance in macroalgae

Microorganisms grow to higher densities in water than in air (Engel *et al*., 2002). The aquatic environment is thought to promote the formation of biofilms on surfaces and evidence for interactions between macroalgae and microbial epiphytes can be found in literature (Jaffray *et al*., 1996; 1997; Weinberger *et al*., 1997; 1999; Schroeder *et al*., 2003). In striking contrast to the significant amount of literature available on biotic interactions in terrestrial crops, regulation of activated defences is not clearly understood in macroalgae (Potin *et al*., 2002). Many of the specialized structural defences present in vascular plants are absent in macroalgae (Weinberger *et al*., 1999). Given that seaweeds, like terrestrial plants, are mostly sessile organisms attached to surfaces, they cannot retreat as animals do when exposed to a potential threat. However, seaweeds have evolved a variety of defence mechanisms which specifically make use of their chemical repertoire to influence interactions with other organisms and the environment (Paul *et al*., 2006; Pohnert, 2004). A portion of these chemicals may provide constitutive physical barriers against grazers or parasites while constitutive production of secondary metabolites serve as antimicrobial compounds (Kubanek *et al*., 2003) and grazer deterrents (Paul *et al*., 2006). However, since the physiological costs to seaweeds would be metabolically high if they were to constitutively produce these compounds, it is likely that activated and inducible defence mechanisms may have also evolved (Cosse *et al*., 2008). Furthermore, Weinberger (2007) proposed that innate immunity requires more than just the ability of an organism to detect elicitors and activate defences quickly enough to contain pathogens, it also requires that the necessary concentrations of signals and defence compounds be reached under natural growth conditions. Host-pathogen interactions of plants and animals take place in the phylloplane and lymphatic system respectively, where concentrations of compounds can, to an extent, be controlled by the host (Cosse *et al*., 2008). Therefore, in the context of the marine environment, the question of possible dilution of elicitors and signal molecules and whether they are even able to reach appropriate concentrations for induction of the defence response cannot be disregarded. Even though analytical tools that allow real-time monitoring of elicitors directly do not yet exist, there is scientific evidence suggesting that the concentrations of these compounds are sufficient to trigger responses in macroalgae (Cosse *et al*., 2008). Since several similarities between

disease resistance of seaweeds and higher eukaryotes do exist, seaweeds could potentially serve as model organisms to assess the hypothesis that these essential cell functions initially evolved in the oceans.

1.5.1 Pathogen-induced defence in macroalgae and higher plants

Three recent reviews (Cosse *et al*., 2008; Potin, 2008; Weinberger, 2007) comprehensively highlight advances in the field regarding the current model of activated and induced-defence mechanisms in macroalgae.

It has already been established that activated and inducible defence mechanisms in any organism requires that the threat of the pathogen not only be perceived but perceived in the necessary time to mount an effective response. Receptors of vascular plants and vertebrates typically perceive two types of elicitors: exogenous pathogen-associated molecular patterns (PAMPs) and endogenous elicitors or pathogen induced molecular patterns (PIMPs) such as the breakdown products of the host's cell wall caused by enzymatic degradation by the invading pathogen (Nürnberger *et al*., 2004; Mackey and McFall, 2006). Once receptors are activated by elicitors, the first line of defence appears to be an oxidative burst or transient release of reactive oxygen species (ROS). ROS has been implicated in several essential biological processes including signal transduction (Neill *et al*., 2002); the expression of inducible defence genes; synthesis of antioxidant enzymes (catalase and peroxidase) to counteract damaging effects to cellular components; direct toxicity toward the invading pathogen; specific oxidative-burst-associated responses such as the emission of volatile halogenated organic compounds (VHOCs), lipid peroxidation and generation of oxylipins, synthesis of phenolic compounds, cell wall cross-linking, stomatal closure and other wounding responses; programmed cell death (PCD); hypersensitive response (HR) and peroxisome biogenesis (Potin, 2008). Furthermore, transcriptional regulation of defence genes in higher plants has been shown to be regulated by various stresses such as high light intensity, UV-light, ozone, temperature extremes, dehydration,

wounding, pathogen challenge and elicitors which all result in the production of ROS. This model of inducible defence mechanism appears to be conserved in macroalgae.

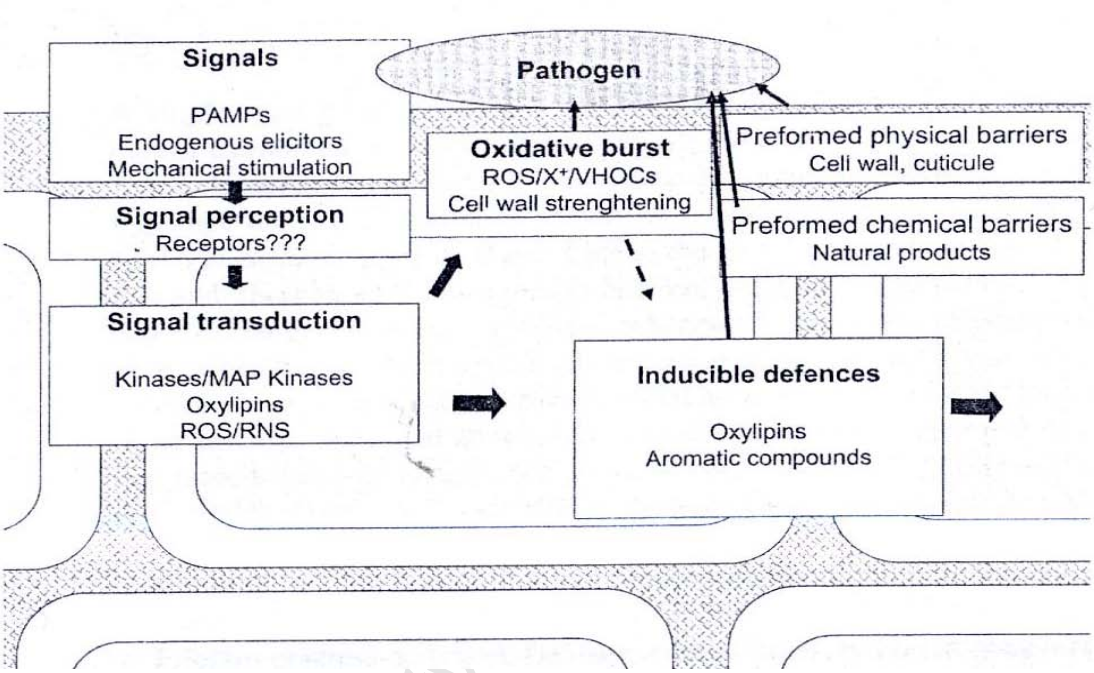


Figure 1.4 Hypothetical, simplified scheme of the complex sensory and signalling mechanisms following the oxidative burst induced by recognition of exogenous PAMPs or endogenous elicitors based on the current knowledge for marine algae. The presence of receptors has not yet been proven in macroalgae (indicated by '???'). Pharmacological evidence and direct measurements of ion influxes suggests that activation of signalling cascades controls the generation of ROS leading to gene-regulated inducible defences and/or for direct toxicity towards the invading pathogen (indicated by thick solid arrows). Cell wall strengthening, preformed physical and chemical barriers as well as inducible responses together constitute a multilayered defence to contain pathogen infection (indicated by thin solid arrows). Abbreviations: PAMPs (pathogen-associated molecular patters), ROS (reactive oxygen species), RNS (reactive nitrogen species), VHOCs (volatile halogenated organic compounds), X^+ (oxidized halides) (Potin, 2008; Nurnberger and Lipka, 2005).

The evolutionary conserved molecular mechanism of receptor-mediated activation in macroalgae and higher plants is highlighted by the high degree of similarity between the functional analogues (defence elicitors) that effect similar defence-related physiological changes (Weinberger *et al*., 2001). As seen in Figure 1.4, the receptor(s) associated with an inducible defence response have not yet been identified in macroalgae (Potin, 2008). Nonetheless, upon detection of an elicitor, macroalgae also respond with an oxidative burst. Such regulation implicitly implies molecular recognition of an elicitor by a receptor and further supports the existence of macroalgal elicitor receptor(s) (Weinberger, 2007). The induced defence-related physiological responses associated with higher plants have been demonstrated in several studies regarding red macroalgae. For example, when species of *Gracilaria* (including *G. gracilis*) were exposed to cell wall breakdown products (agar oligosaccharides), an accumulation of H_2O_2 in the surrounding algal medium was observed within minutes (Weinberger *et al*., 2010; Weinberger *et al*., 1999). However, photosynthetic and non-light-dependent H_2O_2 production has been documented during general plant metabolic processes (Elstner *et al*., 1996). In addition, light-dependent H_2O_2 production has been observed during photosynthetic stress as well as under non-stress conditions in algae (Pedersen *et* al., 1996; Mtolera *et al*., 1995). To prove that the observed oxidative burst was in fact due to exposure to disease elicitors and not light stress or general metabolism, Weinberger *et al.* (1999) measured H_2O_2 production in the absence of light. Significantly increased levels of H_2O_2 were detected which suggested that exposure to elicitors was in fact the cause of ROS release. Transmission electron microscopy (TEM) showed that the site of ROS production was in the plasma membranes of epidermal and sub-epidermal cells (Weinberger *et al*., 2005). The authors showed that the oxidative burst was sensitive to inhibitors of NADPH-dependent enzymes and flavoenzymes suggesting that a NADPH oxidase might be the source of ROS production. As with vascular plants, Ca^{2+} channel inhibitors prevented the oxidative burst while Ca^{2+} ionophores enhanced its intensity. Furthermore, protein kinase inhibitors inhibited the transient production of ROS, while phosphatase inhibitors enhanced production after elicitor perception, thus implicating phosphorylation events in NADPH oxidase activation. Evidence to the contrary suggested that in some instances, elicitor perception may not be required to produce an oxidative burst. In one study, agar oligosaccharides did not activate NADPH oxidase in *G. chilensis*, but were instead oxidized by

an agar oligosaccharide oxidase located in the cell wall (Weinberger *et al*., 2005). However, in a very recent study, Weinberger *et al*. (2010) showed that in fact, the oxidation of agar oligosaccharides is nearly universal in gracilarioids and concluded that a common agar oligosaccharide receptor may be present in Gracilariaceae. Following exposure to elicitors, almost all of the genera tested in the study displayed an increased expression of proteins which exhibited agar oligosaccharide oxidoreductase activity. However, the activation of NADPH-oxidase and subsequent generation of ROS (via elicitor and membrane-bound receptor interactions) was restricted to certain genera, including *G. gracilis*. In the case of NADPH-dependent ROS generation, protein phosphorylation events were indeed required for the activation of NADPH oxidase since it was sensitive to kinase inhibitors. On the contrary, application of the kinase inhibitor had no effect on the increased expression of agar oligosaccharide oxidoreductase which could be inhibited with diphenylene iodonium (DPI), a specific inhibitor of NADPH-dependent enzymes. Thus, up-regulation of agar oligosaccharide oxidoreductase involves a defined cellular signalling pathway but whether activation of NADPH oxidase and induced expression of agar oligosaccharide oxidoreductases share common signalling pathways is yet to be determined. NADPH-derived ROS therefore appears to be required for eliminating pathogens but may also play a role as second messengers in defence-related signalling cascades (Van Breusegem and Dat, 2006; Laloi *et al*., 2004,). Oxidative-burst-associated responses in macroalgae may occur rapidly in order to mount an effective chemical defence against pathogens and grazers. As evident with the oxidative burst of mammalian phagocytes, rapid ROS production following pathogen or elicitor recognition in macroalgae was associated with the production of hypohalous acids which were capable of halogenating various organic substrates (Weinberger *et al*., 1999). An increased production of iodinated, chlorinated or brominated organic compounds was associated with oxidative stress induced by excess light (Mtolera *et al*., 1996), ultra-violet exposure (Laturnus *et al*., 2004), and temperature fluctuations (Abrahamsson *et al*., 2003). In addition, elevated levels of VHOCs were observed in response to grazing pressure (Nightingale *et al*., 1995). Biosynthesis of these halogenated compounds required vanadium haloperoxidases (vHPOs) which catalyze the oxidation of halides (X^-) to generate X^+ needed to produce hypohalous acid (XIO). Marine organisms, especially seaweeds, have long been known to concentrate halides from the environment (Leblanc *et al*., 2006)

highlighting their importance for inducible defence response. Red macroalgae have been shown to produce oxylipins derived from C18 and C20 fatty acids which suggest regulatory roles for these compounds in algal defence (Pohnert, 2004; Potin *et al*., 2002). Signalling compounds such as jasmonic acid (JA) have been shown to play a role in secondary signalling in red macroalgae. For example, Collén *et al*. (2006) exposed the red macroalga *Chondrus crispus* to methyl jasmonate over a 24 hour period. An increase in the transcription of genes annotated to roles in stress responses was observed while genes involved in general metabolism and energy conservation were repressed. In a separate study, up-regulation of an agar oligosaccharide oxidoreductase was observed in different *Gracilaria* species after exposure to agar oligosaccharides for 24 hours, resulting in enhanced resistance to surface attachment by epiphytes (Weinberger, 2007). Up-regulation of the agar oligosaccharide oxidoreductase was prevented with specific inhibitors of nitric oxide (NO) synthase. ROS and NO radicals generated in the course of an oxidative burst are known to be integral components of the regulatory signal networks that modify gene transcription and protein expression patterns in plants (Pitzschke *et al*., 2006; Zeidler *et al*., 2004; Neill *et al*., 2002). Up-regulation of defence proteins in macroalgae therefore suggests that intracellular signalling mediates their transcriptional activation. An interesting similarity between macroalgae and higher plants is the hypersensitive response (HR) and plant cell death (PCD) phenomena. At the site of pathogen invasion, infected and adjacent macroalgal cells undergo hypersensitive cell death in order to minimize pathogen proliferation (Schroeder *et al*., 2003; Jaffray and Coyne, 1996; Weinberger *et al*., 1997, 1994). The HR is a consequence of PCD which is different from accidental death caused by wounding or accumulation of toxic compounds (Van Breusegem and Dat, 2006). In higher plants, regulation of the HR involves sensing changes in intracellular homeostasis of ROS during the oxidative burst (Delledonne *et al*., 2001). An uncoupling of respiration and phosphorylation is required for PCD (Van Breusegem and Dat, 2006; Tiwari *et al*., 2002) which results in increased respiration, amplified ROS generation and depletion of ATP. Evidence suggests that similar mechanisms for PCD exist in red macroalgae. For example, uncoupling of respiration was observed in *Gracilaria* species exposed to agar oligosaccharides. This uncoupling displayed sensitivity to respiration inhibitors and resulted in decreased PCD (Weinberger *et al*., 1999). In addition to the mitochondria, chloroplasts were also involved in the regulation of PCD.

Exposure to small doses of light expedited PCD. In higher plants and other eukaryotes, key enzymes activated during PCD include nucleases (involved in DNA cleavage) and caspases (a family of highly specific proteases) (Lam and del Pozo, 2004; Bucker *et al*., 2000; Heath, 2000; Gou *et al*., 1993). Investigations with the brown macroalga *Laminaria japonica* revealed that PCD after pathogen infection was dependent on caspases while cleavage of DNA was also observed (Wang *et al*., 2004). Even though evidence supporting the existence of an inducible defence response in macroalgae is clear, such a response should ultimately be effective in enhancing host resistance and preventing progression to a diseased state. In this regard, current literature has proved that this is indeed the case. In one study, *Gracilaria* species that were exposed to agar oligosaccharides were able to eliminate up to 60% of the resident bacterial epiphytes within 1 hour (Weinberger and Friedlander, 2000). Single strains of agarolytic bacteria previously isolated from the thallus surfaces of healthy or decaying seaweed proved to be particularly sensitive when re-inoculated onto healthy *G. conferta* previously exposed to agar oligosaccharides. Within 15 min after exposure, up to 90% of these agar-degrading bacteria were eliminated from the thallus surface. In addition, increased resistance against epiphytic microalgae was observed.

This evolutionary conserved model of inducible defence in macroalgae (and higher eukaryotes) is a useful starting point for elucidating disease resistance in *G. gracilis*. However, scientific investigation aimed at the identification, functional annotation and characterization of specific defence genes is still required. Furthermore, insight into the precise molecular mechanisms that govern their transcriptional regulation subsequent to elicitor recognition and release of ROS is essential. In order to discover and then characterize novel defence genes in *G. gracilis*, the relevant molecular tools currently available must be considered.

1.6 Functional genomics and its role in the identification of macroalgal defence genes

A comprehensive understanding of the genetic networks (genomics), proteins (proteomics) and small molecules (metabolomics) that underlie any physiological response in an organism requires the characterization of each of the above molecular components. Similarly, elucidation of a macroalgal defence response will only advance by focussing on genes (and proteins) that are regulated as a result of receptor-elicitor interactions as well as the induced physiological responses to pathogen progression (Potin, 2008; Mahalingam *et al*., 2003; Reymond, 2001).

1.6.1 High throughput DNA sequencing

Molecular techniques such as high throughput DNA sequencing have enabled researchers to identify and characterize every gene in any given genome. However, the challenge associated with knowing a gene's DNA sequence is the need to accurately annotate gene function to the DNA sequence (Kent *et al*., 2002; Wu *et al*., 2001). For example, researchers on the human genome project catalogued more than one million expressed tagged sequences (ESTs) which corresponded to 52 907 unique human genes. At that time, the preliminary functional annotations, transcriptional regulation and expression of more than 80% of the genes were yet to be characterized (Duggan *et al*., 1999). The large discrepancy between DNA sequence and functional annotation would obviously be much less today as researchers are constantly characterizing the human genome, but the point being made here is that for organisms such as *G. gracilis*, with relatively little DNA sequence data presently available, the task of identifying novel defence genes and assigning putative roles in defence mechanisms remains a challenge. Comparisons between *G. gracilis* DNA sequences and various annotated gene or protein databases may be useful for genes that exhibit significant sequence homology, but genes unique to this macroalga would be near impossible to functionally annotate without further scientific investigation. A few algal nuclear and/or plastid genome sequences have been completely determined while several others only have partial coverage (Cosse *et al*., 2008). These include the unicellular red algae *Cyanidioschyzon merolae* (Nozaki *et al*., 2007; Matsukazi *et al*., 2004; McFadden *et al*., 2004; Ohta *et al*., 2003) and *Galderia sulphuraria* (Barbier *et al*., 2005); the

green algae *Chlamydomonas reinhardtii* (Merchant *et al*., 2007); *Ulva linza* (Stanley *et al*., 2005) and *Ostreococcus tauri* (Derelle *et al*., 2006); the brown macroalgae *E. siliculosus* (Cosse *et al*., 2008) and *Laminaria digitata* (Roeder *et al*., 2006); and the diatoms *Thalassiosira pseudonana* (Armbrust *et al*., 2004) and *Phaeodactylum tricornutum* (Scala *et al*., 2002). In the absence of complete genome information, the analysis of dedicated EST libraries (NCBI; http://www.ncbi.nlm.nih.gov/dbEST/) is a promising strategy for assigning putative gene function based on DNA sequence homology. ESTs are nucleotide sequences derived from the ends of cDNA clones which are in turn derived from mRNA transcripts expressed in the cell (Borsani *et al*., 1998). For existing EST libraries to be useful in the assignment of putative biological functions, ESTs with significant sequence homology to un-annotated *G. gracilis* DNA sequences have to already possess a functional annotation. Red macroalgal species such as *Griffithsia okiensis* (Lee *et al*., 2007), *Porphyra yezoensis* (Azamizu *et al.*, 2003; *Nikaido et al*., 2000), *Chondrus crispus* (Collen *et al*., 2006), *G. changii* (Teo *et al.*, 2007) and *G. gracilis* (Lluisama and Ragan, 1997) all contribute to current EST databases but the overall lack of knowledge with respect to macroalgal defence genes can be attributed to the limited amount of DNA sequences and/or functional annotations available in these databases.

1.6.2 Molecular tools for transcriptome analysis

An alternative approach for elucidating gene function (via high throughput sequencing followed by comparative sequence homology analyses) is the determination of gene transcription profiles. Analysis of an organism's transcriptome (mRNA transcripts) can be used to deconstruct which and how different genes work together in a given physiological response. By associating transcriptional regulation to biological response, the function of unknown genes can indirectly determined (Holtorf *et al*., 2002). The elucidation of gene function based on transcript profiling techniques can be further enhanced by combining data with proteomics and metabolomics platforms (Fiehn *et al*., 2002; Fiehn *et al*., 2001). There are several large scale transcript profiling techniques presently available which include cDNA-AFLP (Bachem *et al*., 1996); serial analysis of gene expression (SAGE; Velculescu *et al*., 1995); massive parallel signature

sequencing (MPSS; Brenner *et al*., 2000) and microarray technology (Duggan *et al*., 1999; Schena *et al*., 1995). DNA microarrays are able to explore gene transcription patterns on a genome-wide scale and have become the most common platform to classify synergistic (or antagonistic) transcription patterns amongst the genes assayed (Hoheisel, 2006; Brown and Botstein, 1999; Duggan *et al*., 1999; Iyer *et al*., 1999; Spellman *et al*., 1998). DNA microarray technology was invented using a small set of ESTs from *Arabidopsis thaliana* (Schena *et al*., 1995) and has since been applied to several model (sequenced genomes) organisms including *Escherichia coli*, yeast, human, mouse and the fruit-fly (Richmond *et al*., 1999; Chu *et al*., 1998; Shena *et al*., 1996; Tanaka *et al*., 2000; White *et al*., 1999). Two different array-based technologies (cDNA and oligonucleotide) exist. For detailed reviews on either platform, see Donson *et al*. (2002), Aharoni and Vorst (2001), Duggan *et al*., (1999) and Lipshutz *et al*. (1999). Although both technologies are equally capable of analyzing patterns of gene expression effectively, this study has employed the cDNA-based microarray approach to identify putative defence genes of *G. gracilis*.

1.6.3 Overview of cDNA microarray technology

Microarray technology has become a standard molecular tool that is capable of assessing transcription of several (~thousands) genes in a single experiment (Schindler *et al*., 2005; Schulze and Downward, 2001). By quantitating the relative abundance of mRNA transcripts simultaneously, the functional relationships between genes can be established (Wen *et al*., 1998). The most common application of microarray technology is the identification of differential expression of genes between samples subjected to different treatments (Hoheisel, 2006; Hedge *et al*., 2000). One notable feature of microarrays is the ability to generate custom arrays. In the case of cDNA microarrays, DNA segments (cDNA probes) representing the collection of genes to be assayed are amplified by PCR, purified and mechanically spotted (~5 nl) at high density on coated glass microscope slides using simple x-y-z spatial co-ordinates (Figure 1.5). Microarrays are subsequently queried by co-hybridization using two or more fluorescently labelled targets prepared from total mRNA from the test and reference samples. Prior to hybridization, mRNA

transcripts are reverse-transcribed to cDNA incorporating a modified dUTP and then fluorescently labelled with either Cye3- or Cye5-dye via dye-dUTP coupling (Shulze and Downward, 2001; Duggan *et al*., 1999; Shalon *et al*., 1996). The kinetics of hybridization allows the relative expression of every mRNA target present in each treatment (test and reference) to be established based on the ratio with which each target hybridizes to an individual microarray element (Hedge *et al*., 2000). Laser excitation of the hybridized microarray elements at the appropriate wavelengths for each fluorescent Cye-dye yields characteristic emission spectra. Monochrome images from each Cye-dye channel are imported into software which merges and pseudo-colours the images. Data from a single hybridization experiment is then treated as an expression ratio (i.e. $\log_2$(Cye3/Cye5)) to determine the relative level of that mRNA transcript in relation to each of the treatments (Hoheisel, 2006; Duggan *et al*., 1999).

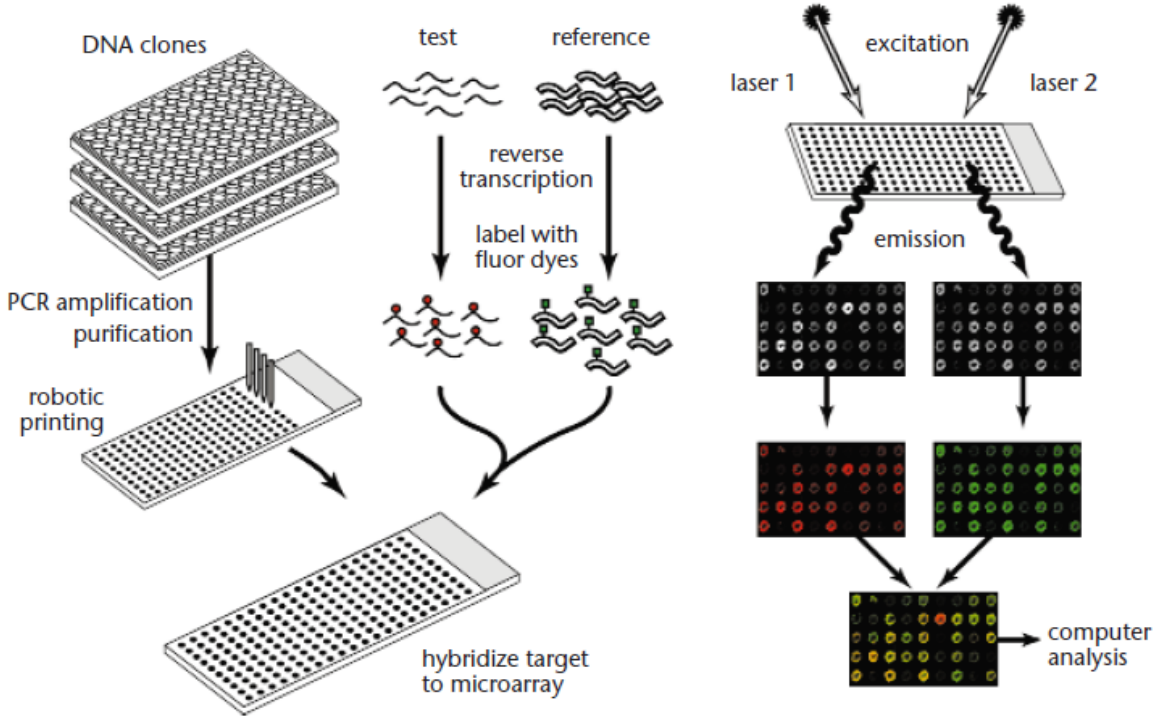


Figure 1.5 Simplified schematic depiction of the cDNA microarray procedure used to measure mRNA transcript changes between test samples (Duggan *et al*., 1999). The cDNA microarray protocol can be divided into three distinct sections, namely (i) construction of the cDNA array, (ii) preparation of RNA targets and hybridization to the array and (iii) Laser scanning, image acquisition and data analysis.

Plant biologists have realized the potential of cDNA microarray technology to identify differentially expressed defence genes (Kuhn, 2001; Richmond and Somerville, 2000). For example, Schenk *et al*. (2000) and Desikan *et al*. (2001) used EST collections of *A. thaliana* to construct cDNA microarray experiments in an initial attempt to identify genes whose expression levels were altered in biotic and oxidative stress, respectively. Using a small cDNA microarray enriched in defence-response genes, Somerville *et al*. (1998) identified *A. thaliana* genes that were differentially regulated in response to powdery mildew infection. Reymond *et al*. (2003) assessed, via cDNA microarrays, the timing and regulation of expression of 150 defence-related *Arabidopsis* genes in response to mechanical damage and insect attack. Transcriptional

regulation of defence genes in species of the pine family was determined using cDNA microarrays (Ralph *et al*., 2006). Fujiwara *et al.* (2004) identified defence genes regulated upon exposure to pathogen flagellin proteins in cultured rice cells. Similarly, Li *et al*. (2006) assessed the expression profile of rice genes in early defence responses to blast and bacterial blight pathogens using cDNA microarrays. With the advancement in high throughput DNA sequencing and potential of microarray-based transcriptome analyses, it is not surprising that functional genomics based approaches has found widespread acceptance in algal research (Cosse *et al*., 2008). The pioneering work of Collén *et al*. (2006) highlighted the feasibility of using EST libraries from the red macroalga *C. crispus* to construct cDNA microarrays. Transcriptome analyses after exposure to methyl jasmonate were subsequently used to identify differentially regulated genes. Moreover, the authors were able to utilise the observed transcription patterns to cluster genes into groups which appeared to respond similarly (synergistic or antagonistic) after exposure to methyl jasmonate. For instance, expression of *C. crispus* genes annotated to roles in stress responses were induced while genes annotated to general metabolism and energy conservation were repressed. Although their work was the first report of a transcriptomic study to elucidate the transcriptional component of activated defence in a red marine macroalga, to our knowledge this is the first study to establish the transcription profile of *G. gracilis* exposed to disease elicitors in order to identify novel defence genes. In very recent studies, Ho *et al*. (2009) and Teo *et al*. (2009) used the same approach to construct cDNA microarrays for the red macroalga *G. changii*. The authors successfully identified novel transcriptionally regulated genes in response to abiotic factors such as light depravation and osmotic stress. As macroalgal EST libraries expand and DNA microarray technology advances, whole genome chips for various macroalgae will eventually become available. Such resources may serve as predictive tools for global transcriptional changes of defence genes in response to biotic stresses. Furthermore, it may be possible to identify gene expression signatures unique for each pathogen, thus providing novel tools for diagnosis and management of infectious diseases (Cummings and Relman, 2006) in *G. gracilis* farming.

1.7 Significance and aims of this study

It is now clear that one of the major obstacles hindering progress in the field of macroalgal biotic (and abiotic) stress responses has been the lack of genomic information (Collen *et al*., 2007). According to Weinberger (2007), systematic comparisons of gene transcripts in macroalgae that have either been or not been exposed to disease elicitors should provide a clearer indication of whether, and which, defence mechanisms are induced after elicitor-receptor interactions. Therefore, to better understand the complex biotic interactions that occur between *G. gracilis* and its pathogens, research in this field must continue to deconstruct the molecular components of defence as well as the possible avoidance or suppression tactics of pathogens to bypass algal defences (Cosse et al., 2008; Chan *et al*., 2006). Several of the innate immunity traits of marine plants, such as pathogen recognition and the oxidative burst machinery, are conserved in higher eukaryotic lineages which suggests that the underlying biochemical machinery arose early in evolution (Potin *et al*., 2002). Thus activated defence mechanisms in macroalgae may serve as simple models for similar physiological and molecular responses in higher eukaryotes. According to current literature, genetic transformation studies have already commenced for several red seaweeds, including *Gracilaria* species (Qin *et al*., 2005). In regard to the development of a sustainable *G. gracilis* marine aquaculture industry in southern Africa, the defence genes identified in this study could ultimately serve as targets for genetic manipulation to create seaweed strains which exhibit enhanced disease resistance. The broader objectives of this study were three-fold:

i. Due to the general lack of genetic information and characterization of the biotic interactions between pathogens and *G. gracilis*, this study aimed to identify transcriptionally regulated genes by means of a systematic comparison between seaweed that had either been exposed or not exposed to disease elicitors using cDNA microarray technology. Furthermore, putative defence gene ESTs would be sequenced and compared to functionally annotated genes of other organisms to establish their possible functions. In so doing, it would be possible to assess the level of evolutionary conservation between the various molecular components of activated defence mechanisms between macroalgae and higher organisms.

Based on these objectives, the significance of this study clearly emerges because not only will novel genome data for *G. gracilis* be generated, putative defence genes will also be functionally annotated. DNA sequences and their corresponding biological functions will be deposited in EST databases thereby expanding macroalgal genomic data available to other macroalgal researchers.

ii. The use of cDNA microarray technology has already been shown to be feasible for the identification of differentially regulated genes in other macroalgae. However, such studies also emphasize the importance of independently validating microarray data using a different molecular technique (Rockett *et al*., 2004; Chuaqui *et al*., 2002; Rajeevan *et al*., 2001). Therefore, the current study will validate transcriptional regulation for several putative defence genes using real-time PCR.

iii. In general, a physiological response by an organism to a stimulus is comprised of a genetic component (DNA and RNA), protein products and the small effector molecules that result in the appropriate response which allows the organism to adapt effectively. However, transcriptional regulation of a gene and its biological function can only be effected in a cell via its translation (or lack of translation in the case of gene repression) into a protein product. Exposure of macroalgae to disease elicitors has been shown in literature to result in altered gene transcription as well as various physiological changes that lead to disease resistance. Transcriptional regulation of any gene however, cannot be assumed to coincide with a corresponding change in its protein concentration within the cell. Therefore, this current study aimed to assess whether a correlation existed between several transcriptionally regulated *G. gracilis* mRNA transcripts and the levels of their protein products following exposure to disease elicitors. The significance of this would be that if transcriptional changes are positively correlated to altered protein concentrations that lead to disease resistance, then seaweed biologists could ultimately utilise proteomics platforms (which are cheaper and more efficient than transcriptomics platforms) as predictive or diagnostic tools in the context of *G. gracilis* farming.

CHAPTER 2

Identification of differentially expressed genes in *Gracilaria gracilis* after exposure to disease elicitors using cDNA microarray technology

2.1 INTRODUCTION

The nuclear genome of *G. gracilis* is at present poorly characterized. As such, genes pertinent to activated defence as well as their transcriptional regulation have remained elusive. In addition to the need for expanding the available genetic information, a clear understanding of the molecular mechanisms governing macroalgal activated defence is ultimately required in order to develop a sustainable *G. gracilis* aquaculture programme in Southern Africa. Published scientific literature has previously demonstrated the feasibility of utilizing high throughput DNA sequencing and/or microarray-based transcriptome analyses to identify and characterize novel algal genes under various abiotic and biotic stresses (Cosse *et al*., 2008; Collén *et al*., 2007; Collén *et al*., 2006). More recently, cDNA microarrays have been successfully used to identify differentially expressed genes in the red macroalga *Gracilaria changii* (Ho *et al*., 2009; Teo *et al*., 2009). Therefore, this study utilised cDNA microarrays to identify differentially expressed genes in the red macroalga *G. gracilis* following exposure to disease elicitors for 24 hours.

Several challenges were associated with conducting the cDNA microarray experiment. For example, fluorescent labelling procedures require large amounts of RNA per cDNA experiment (Zhoa *et al*., 2002). Extraction of RNA from *G. gracilis* was particularly difficult due to the presence of polysaccharides, polyphenolic compounds and ribonucleases (all released upon cell disruption) which are known to decrease RNA yields (Rodriguez *et al*., 2009; Marrion *et al*., 2005; Chan *et al*., 2004; Bellanger *et al*., 1990; Loomis, 1974). Thus, a suitable and reproducible RNA extraction method was developed for this study. In addition, gene expression levels as measured by microarrays have been shown to be influenced by technical variables such as the efficiency of reverse transcription, labelling bias of fluorescent dyes, standardization of data, image acquisition and data analysis (Beneš and Muckentaler,

2003; Schuchhardt *et al*., 2000; Claverie, 1999; Eisen and Brown, 1999; Winzeler *et al*., 1999; Chen *et al*., 1997). Consequently, appropriate data normalization techniques were applied to minimize any potential bias before data analysis (Leung and Cavalieri, 2003; Yang *et al*., 2002).

The cDNAs corresponding to differentially expressed genes were sequenced. Multiple sequence alignments and comparisons to various genome and protein databases were performed to determine the level of homology as well as to functionally annotate novel *G. gracilis* genes. Based on these functional annotations, roles for these genes (and their protein products) were proposed in the context of an activated defence response in *G. gracilis*.

2.2 MATERIALS AND METHODS

All media and solutions used in this study are listed in Appendix A.

All DNA oligonucleotide primers were synthesized in the Department of Molecular and Cell Biology, University of Cape Town.

2.2.1 Construction of microarray

2.2.1.1 PCR amplification of cDNA

A *G. gracilis* cDNA microarray was constructed using cDNA clones from two previously constructed EST libraries, namely, biotic stress (exposure to disease elicitors (Iyer, unpublished data)) and abiotic stress (nitrogen limitation (Gebrekiros, 2003)). *Escherichia coli* XL1 Blue cells harbouring cDNA clones were initially inoculated and maintained on Luria agar (LA) (Appendix A.1) supplemented with 30 $\mu g.ml^{-1}$ chloramphenicol (Cm) (Appendix A.3.1). Bacterial clones were subsequently inoculated into sterile 96-well round-bottomed plates (NUNC^{TM}) containing 100 µl Luria broth (LB) (Appendix A.1) supplemented with 30 $\mu g.ml^{-1}$ Cm and incubated overnight at 37°C. A total of 95 wells were inoculated, while the remaining well served as a no template PCR control. A small volume of each overnight bacterial culture (~ 5 µl) was transferred into a sterile 96-well PCR plate (COSTAR) while the remaining portions were supplemented with 100 µl of 50% glycerol, gently vortexed and stored at –70°C. The PCR plate was sealed with a plastic cover and heated at 96°C for 15 min in order to denature the bacterial cells. A PCR master mix (final volume 100 µl) (Table 2.1) was concurrently prepared, pre-heated at 96°C for 15 min before 98 µl of the master mix was added to each of the 5 µl overnight bacterial cultures. The plasmid cDNA inserts were amplified in a TECHE GENIUS thermocycler (RHYS international) using universal forward and reverse M13 primers (Appendix B.1) according to predefined cycling conditions: an initial 5 minute denaturation was followed by 25 cycles of 30 s denaturation at 94°C, 30 s of annealing at 61°C and 3 min extension at 72°C. The PCR reactions were terminated by an additional incubation of 7 min at 72°C and a cooling step at 4°C for 10 min. The cDNA amplification products were electrophoresed on a 1% TAE

(Appendix A.2) agarose gel. Each product was visually examined and subjectively classified as follows: 'strong single band', 'weak or absent band' or 'multiple bands'. PCR products that failed to meet the quality control criteria of 'one strong band' were flagged and all classifications were stored in a database for future reference.

Table 2.1 Constituents of PCR master mix used for colony PCR.

PCR reagent	Volume	Final Concentration
M13 universal primers	3 µl	0.3 µM
dNTPs (10 mM)	0.4 µl	40 µM
Taq polymerase buffer (10 X)	10 µl	1 X
$MgCl_2$ (25 mM)	12 µl	3.0 mM
Taq polymerase (5 U. μl^{-1})	0.4 µl	0.02 U. μl^{-1}
Nuclease-free dH_2O	74.2 µl	-
TOTAL	*100 µl*	

2.2.1.2 Purification of amplified cDNA products

In order to remove unincorporated nucleotides, primers and excess salts from the PCR reactions, amplified cDNA fragments were transferred to 96 well Millipore plates (Millipore, MultiScreen^{TM} PCR). The Millipore plates were placed on a vacuum apparatus (Millipore, Vacuum/Pressure Pump) and a vacuum was applied. Amplified cDNA products were retained on the membrane via size exclusion while the unnecessary reagents were washed away. The vacuum was released and the membrane at the bottom of each well was washed with 100 µl of sterile Mill Q water. Water was removed by vacuum and the wash procedure repeated. Amplified cDNA fragments bound to the membrane were subsequently re-suspended in 50 µl of 50% DMSO (Sigma) and placed on an orbital shaker at 300 rpm for 2 hours to allow for complete re-suspension. Purified cDNA amplification products were once

again electrophoresed though a 1% TAE agarose gel to visually establish whether the quality criteria for each product had been maintained. Following this, the 50 µl of purified PCR product was divided such that 20 µl was transferred into a 384 well printing plate (GENETIX) while the remainder was stored at –20°C. Seventeen 96-well plates in total were re-racked into five 384-well plates for high throughput robotic deposition onto glass slides.

2.2.1.3 Printing of microarray slides

Purified cDNAs were printed onto aminosaline coated glass slides (GAPII, Corning, NY, USA) using a Micro Grid Biorobotics printer by CAPAR (University of Cape Town). Pins were set up in a 4 x 4 arrangement, allowing 16 different inventory wells (each containing amplified cDNA product) to be visited, loaded and printed in a single source visit. Each array was duplicated on a slide and each PCR product was printed in duplicate (one adjacent to the other) within the array (Figure 2.1). In total, 1620 amplified cDNAs were printed on an array, i.e. 1044 cDNAs from the EST library that represented 18 days of nitrogen limitation and 576 cDNAs from the library that represented exposure to disease elicitors. In addition to the *G. gracilis* cDNA fragments printed on the slide, 23 artificial control genes (Lucidea Universal Scorecard controls, Amersham) were replicated on each array to serve as internal normalization controls. Following printing, each slide was cross-linked under UV-light (120 mJ) to prevent mixing and cross-contamination of cDNAs. Each cDNA on the microarray was assigned a unique ID. For example, Gg_IB9_01C07 was assigned to a single cDNA where ‘Gg’ represented *G. gracilis*, ‘IB9’ represented the EST library from which the cDNA originated and ‘01C07’ represented the plate number and position of the cDNA clone in the original set of 96-well plates. The printed slides were stored in the dark at room temperature until further use.

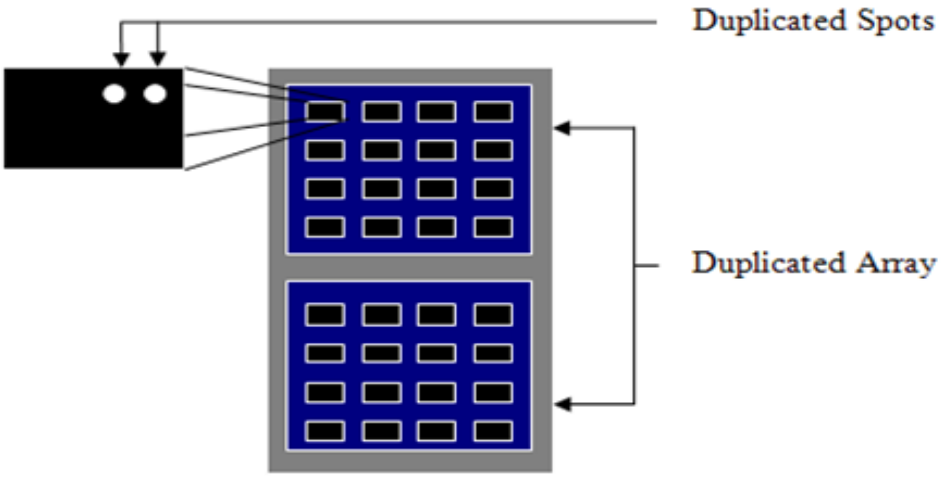


Figure 2.1 *G. gracilis* microarray slide layout showing duplication of the entire array on the slide as well as duplication of the cDNA target within the array (adapted from Lebi, 2006).

2.2.2 Seaweed sample acquisition for exposure to disease elicitors

Samples of *G. gracilis* were sourced from Jacobsbaai Sea Products abalone farm, Cape Town, South Africa. Deionized water was used to remove any sandy or slimy sediment from the thallus surface while any visible epiphytes were detached by hand. Seaweed was maintained in a 20 L plastic tank supplied with seawater. Water movement was achieved by pumping compressed air to the tank through a plastic airline. The water temperature was maintained at 18 – 20°C, and illumination was provided by cool white fluorescent tubes at 7400 Lux with a 16:8 light-dark cycle.

G. gracilis thalli weighing 6.0 ± 0.05 g fresh weight were transferred to eight 0.5 L Erlenmeyer flasks containing 250 ml artificial seawater (ASW) enriched with 1/3 strength PES (0.6% v/v) (Appendix A.1). All seaweed samples were acclimatized in a temperature controlled growth room for three days under a 16:8 light-dark cycle at 18°C. The growth media was discarded after acclimatization. Half the flasks were assigned ‘experiment’ and the other half ‘control’. Disease elicitors used in this experiment were based on the protocol

established by Weinberger *et al*. (1999; personal communication) in which red macroalga *Gracilaria* sp. were exposed to microbial degradation products of the agar cell wall matrix. Agar (Biolab) (0.3% w/v) was autoclaved in ASW for 20 min followed by a 30 min centrifugation at 7000 rpm. The supernatant, which contained agar oligosaccharides, was supplemented with 1/3 strength PES (0.6% v/v) and 250 ml was added to the seaweed in the flask assigned 'experiment'. Two hundred and fifty millilitres of ASW enriched with 1/3 strength PES (0.6% v/v) was added to the seaweed in each of flasks assigned 'control'. Seaweed samples were immediately removed from the control flasks to serve as a time 0 (healthy) control. The control samples were flash frozen in liquid nitrogen and stored at –70°C until RNA extraction. Seaweed samples in the experimental flasks were incubated for 24 hours under a 16:8 light-dark cycle at 18°C. Samples were then removed and flash frozen in liquid nitrogen for storage at –70°C until RNA extraction. In the context of this study, a biological repeat was defined as one pair of seaweed samples (experiment and control). In total, four biological repeats were prepared for the microarray experiment.

2.2.3 RNA extraction protocol

In order to prepare fluorescently-labelled RNA targets for hybridization to the cDNA probes on the microarray slides, RNA had to be extracted from the experimental and control *G. gracilis* samples. All solutions were treated with 0.1% Diethylpyrocarbonate (DEPC) water (Appendix A.3.2.1) and autoclaved for 20 min at 121°C. Glassware was treated with chloroform, followed by 100% ethanol and finally DEPC-water to remove RNAses. Glass- and plasticware were additionally autoclaved at 121°C for 40 min prior to use.

Total RNA was extracted according to the protocol outlined by Azvedo *et al*., 2003. Briefly, *G. gracilis* samples (1.3 g wet weight) were finely ground using liquid nitrogen in a cooled, sterile pestle and mortar. Ground tissue was divided across two 25 ml centrifuge tubes (Beckman) containing 15 ml extraction buffer (Appendix A3.2.2) which had been heated in a 42°C water-bath for 10 min. The ground tissue mixture was vortexed vigorously for 5 min and then incubated for a further 90 min at 42°C to promote cell lysis.

Following incubation, 15 ml chloroform-isoamyl alcohol (24:1 [v/v]) was added to each centrifuge tube to promote extraction of RNA. The mixture was vortexed thoroughly and subsequently centrifuged at 15000 g for 15 min at 4°C. The top aqueous phase, which contained the RNA, was carefully removed and transferred to sterile 25 ml centrifuge tubes. The extraction procedure was repeated by addition of 1 volume of chloroform-isoamyl alcohol to the RNA solution. Centrifugation was repeated as before (15000 rpm for 15 min at 4 °C). The top aqueous phase was recovered and transferred to sterile 25 ml centrifuge tubes. RNA was precipitated out of solution by adding ¼ volume of 10 M LiCl (Appendix A.3.2.3), followed by incubation at 4°C overnight.

Precipitated RNA was subsequently collected by centrifugation at 15000 rpm for 30 min at 4°C. RNA pellets were washed in 2 M LiCl and centrifugation was repeated as before. Following this, RNA pellets were washed in 2 ml of cooled 75% ethanol and centrifuged at 10000 rpm for 10 min at 4°C. Ethanol was decanted and the RNA pellets were allowed to air-dry for 15 min. Finally, the dried pellets were re-suspended in 200 µl DEPC-water. In order to remove contaminating polysaccharides, the re-suspended RNA was incubated at 65°C for 10 min followed by immediate transfer to 4°C for 10 min. Centrifugation at 15000 rpm was performed and the RNA suspension recovered was transferred to sterile 1.5 ml microfuge tubes. Where required, any remaining insoluble material was removed by repeating the centrifugation step (as above).

2.2.3.1 Quantitative and qualitative analysis of RNA

Quantitative analysis of RNA was performed using a Nanodrop spectrophotometer (Thermo Scientific). Absorbance readings were measured at 260 and 280 nm, with one unit of absorption at 260 nm representing 40 µg ml^{-1} of RNA. Ratio measurements at 230, 260 and 280 nm were used to assess the purity of the RNA samples. RNA integrity was determined by assessing a 1 µg sample on a 1.2 % formaldehyde-agarose gel (Appendix A.3.2.5). Formaldehyde-agarose gel electrophoresis was performed in 1 X MOPS (Appendix A.3.2.6) at 70V as described by Sambrook *et al*. (1989). RNA was subsequently visualized on a 254 nm UV-transilluminator.

2.2.3.2 RNA amplification and aminoallyl labelling

Amplification of *G. gracilis* total RNA and Lucidea Universal Scorecard controls was performed using the MessageAmpTM II aRNA kit (Ambion) according to the manufacturer's instructions. Briefly, 1.2 µg of total RNA (spiked with the Lucidea synthetic DNA controls) was mixed with an oligo(dT) primer bearing a T7 promoter to generate first strand cDNA. Second strand cDNA was subsequently synthesized to generate a template for a T7 RNA polymerase which was then purified with a Qiagen RNEasy mini kit (SIGMA) according to the manufacturer's instructions. A 16 hour *in vitro* transcription (IVT) reaction, configured to incorporate the modified nucleotide 5-(3-aminoallyl)-UTP (aaUTP) generated antisense RNA (aRNA). The amplification reaction was terminated by raising the final volume to 100 µl using nuclease-free water. To purify aRNA, unincorporated aaUTPs and free amines were removed using the Qiagen RNEasy mini kit (SIGMA) according to the manufacturer's instructions.

The concentration of aRNA was determined using a Nanodrop spectrophotometer. The aRNA was subsequently fluorescently labelled by chemically coupling either Cy3 (green) or Cy5 (red) NHS ester dyes (AmershamPharmacia) to the aaUTPs. Ambion's online master mix volume calculator (http://www.ambion.com/techlib/append/mm_calcs/msgamp2_96_mm_calc.php) was used to determine the volume of reagents required for preparation of master mixes for the various reactions. Briefly, aRNA (20 µg) was vacuum-dried and then re-suspended in coupling buffer (Appendix A.3.3.1). The Cy-dyes (Appendix A.3.3.2) were then added and the reaction mixture was incubated at room temperature for 30 min in the dark. Finally, hydroxylamine (4 M) (Appendix A.3.3.3) was used to quench the labelling reaction. Labelled aRNA was purified using the Qiagen RNEasy Mini Kit (SIGMA) as per the manufacturer's instructions. A dye-swap was performed in which the assignment of Cy-dye to its respective RNA sample was reversed, i.e. aRNA targets from biological repeats 1 and 2 were labelled as experimental-Cy5 and control-Cy3 whereas this assignment was swapped in biological repeats 3 and 4.

Purified, labelled aRNA was diluted 1:10 with nuclease-free water and quantitated using a Nanodrop spectrophotometer set to read microarray samples. Absorbance readings at 260 nm

were used in Promega's online labelled aRNA calculator (http://www.prontosystems.com/technical_support/calculator/index.asp) to determine the frequency of dye incorporation or the number of Cy-labelled nucleotides incorporated per 1000 nucleotides of cDNA. These values were checked manually by first calculating the total picomoles (pmol) of aRNA synthesized using the following equation:

Pmol nucleotides = [OD_{260} x volume (µl) x 37 ng.µl^{-1} x 1000 pg.ng^{-1}] / 324.5pg.mol^{-1}

where 1 OD_{260} = 37 ng/µl for cDNA and the average molecular weight of a dNTP = 324.5 pg/pmol.

The frequency of dye incorporation was then calculated using:

pmol Cy3 = $_{550}$ x volume (µl) / 0.15

pmol Cy5 = OD_{650} x volume (µl) / 0.25

nucleotides / dye ratio = pmol cDNA / pmol Cy dye

2.2.4 Microarray hybridization of RNA targets to cDNA probes

2.2.4.1 Target hybridization mixture

A target hybridization mixture (Appendix A.3.3.4) for each biological repeat was prepared by pooling 10 µg of purified (5 µg of each treatment) labelled cDNA synthesized from experimental and control *G. gracilis* RNA samples together with blocking reagents, mouse COT1-DNA (Life Technologies) and poly(A)-DNA (Sigma). The final volume of the target hybridization mixture was adjusted to 30 µl. An equal volume (30 µl) of 2X hybridization buffer (20X SSC; 50% formamide and 0.2% SDS) (Appendix A.3.3.5) was stored at room temperature until it could be hybridized to the microarray slide.

2.2.4.2 Preparation of cover-slips for microarray hybridization

Cover-slips (22 mm X 60 mm) (Erie Scientific Company) used in the hybridizations were submerged in 100% acetone for 1 hour at room temperature on a LASEC bench rotational shaker. The cover-slips were then washed in 0.2% SDS for 10 min followed by a further two washes for 10 min in MilliQ water. The cover-slips were then placed in a pre-heated 42ºC oven and left to dry.

2.2.4.3 Pre-hybridization of microarray slides

Microarray slides were pre-hybridized in 60 µl of pre-hybridization buffer (5X SSC; 0.1% SDS; 1% BSA) (Appendix A.3.3.9) before use. A diamond marker was used to outline the edges of the array in order to clearly demarcate the dimensions of the printed microarray. The slides were placed into a hybridization chamber (Arraylt™) and a prepared cover-slip was gently placed onto each slide. Pre-hybridization buffer was injected underneath the cover-slip for dispersion over the entire array. The hybridization chamber was sealed and incubated for 2 hours in a pre-heated container of water in a 42ºC oven. Slides were washed and the cover-slip removed by gentle submersion in a series of five separate containers of MilliQ-Plus (Millipore) water followed by immersion in isopropanol for 1 second and promptly dried by centrifugation at 1000 rpm for 5 min. The slides were used immediately to ensure optimal hybridization efficiency.

2.2.4.4 Target and probe hybridization reaction

The microarray hybridization area was prepared by gently lowering a pre-cleaned glass cover-slip (section 2.2.4.2) over the array area of a pre-hybridized slide (section 2.2.4.3) and placing the slide into a hybridization chamber. The target cDNA hybridization mixture (10 µg) (section 2.2.4.1) was denatured at 95ºC for 3 min, snap-cooled on ice for 30 s and briefly centrifuged. The cDNA mixture (60 µl) was immediately injected under one corner of the cover-slip, allowing the solution to wick along the length of the array. To maintain humidity inside the chamber, 10 µl of MilliQ water was added to each of the reservoir wells at either

end of the hybridization chamber. The chamber was then tightly sealed and placed into a container of water pre-heated to 42ºC. The water container housing the hybridization chamber was covered in foil and placed in a 42ºC oven for 16 – 20 hours.

Following hybridization, the slides were removed from the chamber and washed sequentially with increasing stringency wash buffers as follows. First the cover-slips were removed by submerging the slide in a staining dish containing 42ºC pre-heated low stringency wash buffer (Appendix A.3.3.10). The staining dish was then covered in foil and placed on a rotational shaker (LASEC Lab Rotar) for 5 min at room temperature. The slides were then removed and placed in a staining dish containing a medium stringency wash buffer (Appendix A.3.3.11) and incubated for 5 min on the shaker as before. This process was repeated by placing the slides in a high stringency wash buffer (Appendix A.3.3.12) after which the slides were briefly dipped several times in a staining dish containing 100% ethanol. Slides were immediately dried through centrifugation at 1000 rpm for 5 min and placed in a light-tight slide box until scanned.

2.2.5 Image acquisition

Scanning was performed within one hour of hybridization to obtain maximal fluorescence signal readings for each slide using a GenePix 4000B dual-colour laser scanner (Axon) operated by GenePix 6.0.27 Pro software (Axon Instruments, Inc. Molecular Devices Corporation, CA, USA). Photomultiplier tube (PMT) settings of both channels, i.e. 532 nm (Cy3) and 635 nm (Cy5), were adjusted to levels such that the ratio of signal intensities from the two channels was as close to 1 as possible, with similar range distributions and minimal pixel saturation. A paired 16 bit tagged image file (TIFF) was then generated in order to store the fluorescence intensities for each channel.

2.2.5.1 Extraction of features from spots on the microarray

A spot on the microarray was defined as the representation of the fluorescent signal generated as a result of hybridization between RNA target and cDNA probe. Image segmentation was achieved through four main steps using GenePix Pro software:

i. Generation of a GenePix Array List (GAL) file;
ii. Extraction of foreground intensity pixels;
iii. Correction for background artefacts;
iv. Flagging and exclusion of spots from further analysis.

2.2.5.1.1 Creation of the GenePix array list (GAL) file

During the printing process a data file (section 2.2.1.3) containing the original positions and unique identities of each cDNA probe was created. From this file, a GAL file required for automated array analysis was generated. This GAL file was superimposed onto the pixels generated for each microarray image in a manner that isolated and identified each probe.

2.2.5.1.2 Extraction of foreground intensity pixels

During printing of the array, variation occurred regarding the shape and size of the spots within and across microarrays. As a result, GenePix Pro's proprietary spot-finding algorithm was used to align the features within the GAL file to their appropriate areas on the array. The foreground intensities of each spot were then determined using an adaptive circle segmentation method that fits the best diameter for each spot.

2.2.5.1.3 Background correction and quality assessment of individual spots

Background intensities due to artefacts such as non-specific binding of labelled target, dust, comets as well as auto-fluorescence from probe cDNA, were corrected for using the GenePix Pro local background correction algorithm (Figure 2.2). This algorithm works by creating a circle three times the diameter of the circle identifying the foreground region of a target (Yang *et al*., 2001b). Pixels occurring within the larger outer circle, excluding a two pixel wide region around neighbouring target circles, are considered to be background artefacts. The foreground and background intensities measured for each Cy-dye were determined and background correction was implemented through subtraction of the background intensity from the respective foreground intensity values.

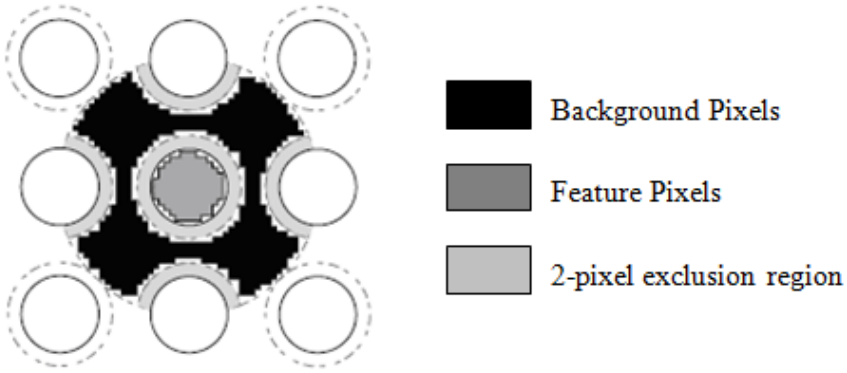


Figure 2.2 Diagram of the local background correction algorithm used by GenePix Pro 6.0.27 software (GenePix Pro 6 user manual).

Automatic morphological spot alignment and background estimation was applied by the software and was manually adjusted where necessary. A predefined filter (Appendix C.1) was used to flag spots that failed to meet minimum quality criteria such as the absence of fluorescent signal, effect of dust on the slide or any other microarray artefact. Flagged spots were not removed from the data set, merely excluded from normalization of the other spots.

2.2.6 Data processing, analysis and normalization

Microarray data normalization and exploratory data analysis was performed using the open-source software package R version 2.3.0 (R Development Team, http://www.r-project.org) (Ihaka and Gentleman, 1996). Specifically, the Bioconductor (http://www.bioconductor.org) (Gentleman *et al*., 2004) package LIMMA (Linear Models for Microarray Data) (Smyth *et al*., 2005) was used for data normalization (see Appendix C.2 for the complete 'R' command script). Normalization of data was performed on two levels. Firstly, data within a microarray slide (one biological repeat) was normalized using a 'Robust-spline' algorithm. Secondly, data across the three microarray slides (between biological repeats) were normalized using an 'Aquantile' algorithm. Selection of the appropriate normalization algorithms within LIMMA was ultimately dictated by the characteristics of the entire microarray data set.

2.2.7 Identification of differentially expressed genes

In the context of this study, genes were deemed to be differentially expressed based on two different approaches:

i. Replicate feature values for each spot were merged using the LIMMA software. Normalized gene expression data was summarized by a design matrix of intensity log-ratios M = $\log_2$(R(intensity of experiment)/G(intensity of control)), with m rows corresponding to the genes under analysis and $n = n_1 + n_2$ columns corresponding to n_1 control hybridizations and n_2 treatment hybridizations. A linear model was fitted to the data and an empirical Bayes statistical function was applied to compute moderated t, *P* and B-statistics. Log_2-fold changes for each gene were generated and ranked according to *P*-value to identify the most statistically significant differentially expressed genes. In this approach, differential gene expression was defined as a statistical outlier relative to the control or when gene expression was associated with a *P*-value less than 0.05.

ii. An alternative approach used to identify differentially expressed genes was based on that published by Causton *et al*. (2003) and Quackenbush (2002). A mean (μ) and standard

deviation (σ) was calculated from the global distribution of the normalized $\log_2$(ratios) for all the spots on the microarray. Genes were deemed to be differentially expressed (95% confidence level) if their M-values deviated more than 1.96(σ) from the mean (μ) of the global distribution. In other words, if the $\log_2$(R/G) ratio (M-value) of any gene was greater than (μ + 1.96(σ)), it was deemed differentially expressed or transcriptionally up-regulated. Conversely, if a gene's M-value was less than (μ – 1.96(σ)), it was deemed transcriptionally down-regulated.

Thus, any gene with a *P*-value < 0.05 and/or (μ - 1.96(σ)) > M-value > (μ + 1.96(σ)) would be considered differentially expressed and selected for DNA sequencing.

2.2.8 Plasmid DNA isolation and sequencing of putative defence genes

Glycerol stocks of *E. coli* XL1 Blue transformants (Section 2.2.1.1) with plasmids harbouring cDNA fragments corresponding to differentially expressed genes were grown at 37°C on Luria agar supplemented with 30 μg/ml chloramphenicol. Single colonies were used to inoculate 5 ml Luria broth (Appendix A.1.2) supplemented with 30 μg/ml Cm. Bacterial cultures were incubated overnight at 37^{o}C on a shaker at 100 rpm. Plasmid DNA was extracted according to Sambrook *et al*. (1989) (Appendix C.1). Plasmid DNA was re-suspended in 50 μl sterile distilled water and quantitated with a Nanodrop spectrophotometer. Additionally, DNA samples were electrophoresed through a 1.2% TAE agarose gel to verify plasmid quality and integrity. All cDNAs were PCR amplified as before (Section 2.2.1.1) with one modification: instead of the M13-F universal forward primer, a synthesized forward primer LIB-F was used (Appendix B.1.2). PCR products were subsequently electrophoresed through a 1.2% TAE agarose gel to verify amplification of a single insert prior to DNA sequencing. Amplified cDNA inserts that met this quality criteria were purified with the E.Z.N.A® Cycle-Pure Kit (peQLab Biotechnologie GmbH) according to the manufacturer's instructions. Cycle sequencing was performed using the BigDye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems). A 10 μl sequencing reaction was set-up as per Table 2.2. Samples were placed in a 2720 Thermal Cycler (Applied Biosystems) set to pre-defined cycling conditions (Appendix B.2.2). Samples for sequencing were purified using

the E.Z.N.A® Cycle-Pure Kit (peQLab Biotechnologie GmbH) according to the manufacturer's instructions. Nucleotide sequences of the purified PCR products were determined using a 3130 Genetic Analyzer (Applied Biosystems) and 3130 Genetic Analyzer Data Collection Software (Version 3.0). Data was subsequently captured using DNA Sequencing Analysis Software (Version 5.2). DNA sequences were edited to remove plasmid vector sequence using CHROMAS (Version 2.01; Technelysium) and analyzed using DNAMAN (Version 4.13; Lynnon Biosoft).

Table 2.2 Reaction constituents for the BigDye Terminator v.3.1 Cycle Sequencing Kit

Reaction constituents	Stock	Volume	Final Concentration
Primer (LibF or M13R)	1.6 pmol/µl	1 µl	3.2 pMol
DNA (cDNA insert)	20 ng/µl	1 µl	20 ng/ µl
Sequencing Buffer	5 X	2 µl	1 X
Ready reaction mix	2.5 X	2 µl	0.5 X
Nuclease-free dH_2O	-	4 µl	-

2.2.9 Bioinformatics and functional annotation of cDNA sequences

BLASTn (Zhang *et al*., 2000) and BLASTx (version 2.2.22) (Altschul *et al*., 1997) sequence homology searches were conducted against the National Center for Biotechnology Information database (NCBI) http://www.ncbi.nlm.nih.gov/Blast.cgi using the non-redundant nucleotide sequence (all GenBank + RefSeq Nucleotides + EMBL + DDBI + PDB sequences) and protein sequence databases (all non-redundant GenBank CDS translations + PDB + SwissProt + PIR + PRF excluding environmental samples from WGS projects),

respectively. Any two or more sequences that exhibited homologous BLAST matches were assessed for sequence redundancy using CLUSTALx (Version 1.81).

In addition, sequences were compared by BLAST analyses using the online resource provided by EMBL-EBI (http://www.ebi.ac.uk/Tools/fasta33/index.html) and Uniprot (http://www.uniprot.org). Homology searches were conducted using the FASTA and FASTX programs for nucleic acid and translated sequences respectively. Sequences that returned BLAST matches with an E value of $1e^{-06}$ or less were considered significant. In this way, sequences were functionally annotated and assigned putative biological roles. Sequences that did not return any significant BLAST results were assessed for the presence of an open reading frame (ORF) and/or conserved domains using the European Bioinformatics Institute InterProScan facility (EMBL-EBI) (http://www.ebi.ac.uk/Tools/InterProScan). Only ORFs longer than 50 codons (150 nucleotides) were accepted as potentially transcribed genes (Lee *et al*., 1999). Where applicable, sequences were analyzed for putative transit peptides for targeting to plastids or mitochondria using Predotar (Version 1.03), ChloroP (Version 1.1) and TargetP (Version 1.1) available at http://www.expasy.ch.

2.3 RESULTS AND DISCUSSION

2.3.1 Array construction

A total of 1620 cDNA probes derived from two EST libraries were successfully PCR-amplified and purified (1044 provided courtesy of Tanya Lebi). Amplified cDNAs were visually assessed and classified into three categories, namely 'strong single band', 'multiple bands' and 'weak or absent band'. As seen in Figure 2.3 (representative of all cDNA 96-well amplifications), some cDNAs were not amplified at all (e.g. row 2, lanes 7 and 22), others contained multiple products (e.g. row 1, lane 24) but on the whole, a high PCR amplification rate was observed as evident by single products in the majority of the lanes. These PCR products were purified and of sufficient quality to be printed on the microarray slides.

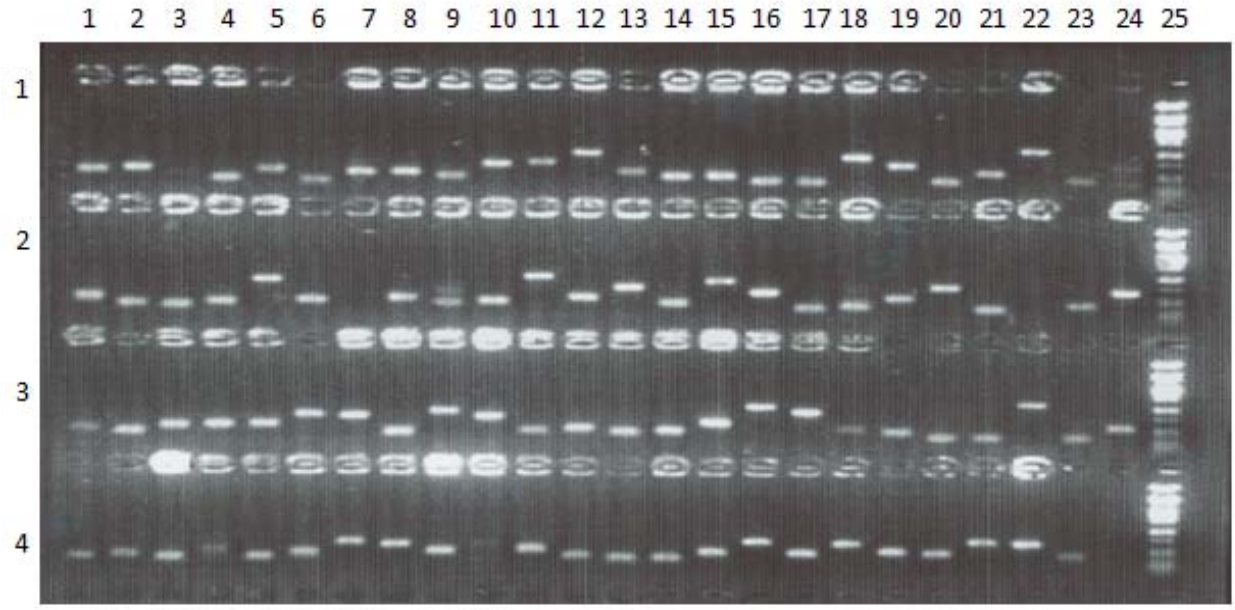


Figure 2.3 Purified PCR products from the disease elicitor EST *G. gracilis* cDNA library. The purified PCR products from the nitrogen stress library were assessed similarly (Lebi, 2006) (data not shown). Rows 1-4, lanes 1-24, contain the purified PCR products from a single 96 well plate. Row 4, lane 24 served as a no template (NTC) control, while row 1-4, lane 25, contained a λPst DNA molecular weight marker.

2.3.2 RNA isolation and preparation of RNA targets

RNA was isolated from *G. gracilis* samples under two experimental conditions, namely, 'experimental' (exposure to disease elicitors) and 'control' (no exposure to disease elicitors). One microgram of RNA was used to assess integrity following electrophoresis through a 1.2% denaturing formaldehyde agarose gel (Figure 2.4). RNA concentration and quality ratios were determined using a Nanodrop spectrophotometer (Table 2.3). Isolation of intact, high-quality RNA from *G. gracilis* was achieved despite the fact that red macroalgae are known to release high levels of polysaccharides, polyphenols and ribonucleases upon cell disruption (Rodriguez *et al*., 2009; Marrion *et al*., 2005; Bellanger *et al*., 1990) which hinder RNA isolation and yield (Sharma *et al*., 2003; Azevedo *et al*., 2003) or inhibit electrophoretic migration and down-stream applications (Wilkins and Smart, 1996). In addition to high integrity, the purity of the RNA isolated was high, as evident by the optimal A_{260}/A_{230} and A_{260}/A_{280} ratios. Ratios above 1.8 indicated low polysaccharide levels and reduced protein contamination, respectively (Chan *et al*., 2004; Azvedo *et al*., 2003). The method of RNA isolation was efficient, reliable and reproducible. By using a high concentration of proteinase K combined with a prolonged 42°C incubation step, enzymatic digestion of ribonucleases resulted in total RNA yields ranging between 9.4 – 74 µg (for cDNA synthesis). No signs of degradation were evident and RNA was thus deemed suitable for the downstream applications of cDNA synthesis and microarray analysis. However, typical fluorescent labelling procedures require large amounts of RNA (2 – 4 µg poly(A)+ RNA or 25 – 50 µg total RNA) per cDNA microarray experiment (Zhoa *et al*., 2002). Since insufficient yields of isolated RNA were observed in for two of the total RNA isolations, i.e. control samples for biological repeats 1 and 2 (< 10µg), a RNA amplification technique was used to circumvent this limitation (Diboun *et al*., 2006). The most common RNA amplification technique is the T7 based linear amplification first developed in the 1990s (Phillips and Eberwine, 1996; Van Gelder *et al*., 1990). It uses a synthetic oligo(dT) primer containing the phage T7 RNA polymerase promoter to prime synthesis of first strand cDNA by reverse transcription of poly(A)+ RNA. The poly(A)+ RNA strand is then degraded with RNAse H, followed by second strand cDNA synthesis with *E. coli* DNA polymerase. Finally, amplified antisense RNA (aRNA) is obtained from *in vitro* transcription of the double-stranded cDNA template using T7 RNA polymerase (Zhoa *et al*., 2002). Several protocols based on this amplification

mechanism have been developed specifically for microarray experiments (Hu *et al*.. 2002; Wang *et al*., 2000; Lou *et al*., 1999; Wodicka *et al*., 1997).

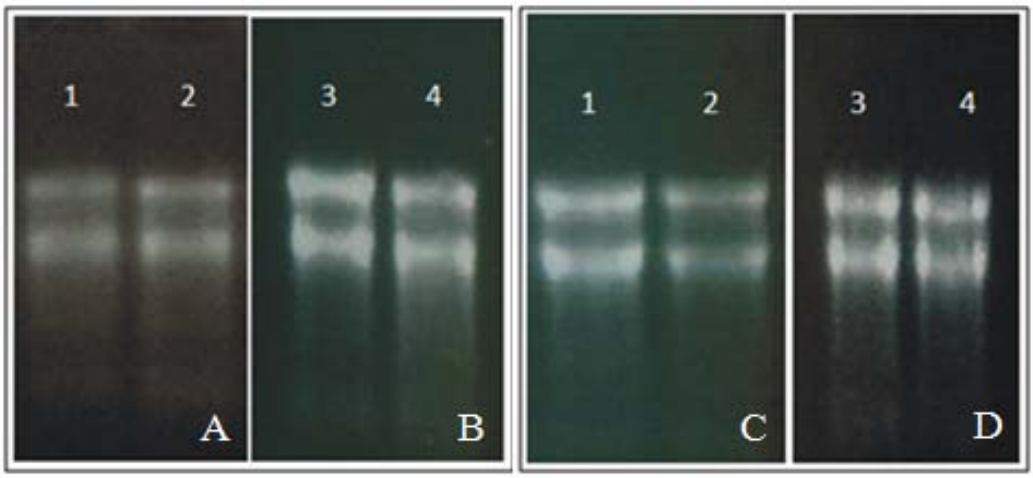


Figure 2.4 Denaturing formaldehyde agarose gel electrophoresis of total RNA preparations (1 µg) from *G. gracilis* samples under control and experimental conditions. (A) RNA samples isolated from *G. gracilis* under control conditions; biological repeats 1 and 2. (B) RNA samples isolated from *G. gracilis* under control conditions; biological repeats 3 and 4. (C) RNA samples isolated from *G. gracilis* under experimental conditions; biological repeats 1 and 2. (D) RNA samples isolated from *G. gracilis* under experimental conditions; biological repeats 3 and 4. Numbers represent biological repeat number.

Table 2.3 Purity and yield of the RNA used for the RNA amplification and subsequent microarray analysis for each biological repeat based on $A_{260/280}$ and $A_{260/230}$ ratios. Control represents *G. gracilis* samples not exposed to disease elicitors. Experiment represents *G. gracilis* samples exposed to disease elicitors for 24 hours. RNA was re-suspended in 200 µl DEPC-water.

Biological Repeat	Origin of *G. gracilis* sample	Yield ($ng.\mu l^{-1}$)	Total Yield (µg)	$A_{260/280}$	$A_{260/230}$
1	Control	47.0	9.4	1.89	2.06
	Experiment	368.2	73.6	2.12	2.40
2	Control	46.0	9.2	1.95	2.12
	Experiment	175.7	35.1	2.08	2.42
3	Control	131.4	26.3	2.20	2.14
	Experiment	295.9	59.2	2.17	2.18
4	Control	138.0	27.6	2.21	2.12
	Experiment	181.0	36.2	2.19	1.86

2.3.3 RNA amplification and aminoallyl labelling

All RNA samples (experiment and control) were amplified using the MessageAmp™ II aRNA kit (Ambion) as per the manufacturer's instructions. The concentration of aRNA was determined by measuring its absorbance at 260 nm on the Nanodrop. As recommended by Ambion, an amount of 5 – 20 µg total aRNA (after amplification) was required for Cy-dye labelling and microarray hybridizations. As seen in Table 2.4, sufficient amounts of aRNA were synthesized to use 20 µg of aRNA for indirect labelling with Cy-dyes (with the exception of Biological Repeat 3 'control' sample which had an aRNA yield of only 8.3 µg). Amplification and fluorescent-labelling of aRNA for biological repeats 1 and 2 were

performed before 3 and 4 (on separate days) and an amount of 20 µg of each labelled aRNA was selected for use. As a consequence of the insufficient yield (8.3 µg) of synthesized aRNA obtained from the control sample of biological repeat 3, it was decided to remove biological repeat 3 from subsequent microarray hybridizations since it would not be biologically meaningful when compared to the three other microarray hybridizations which employed 20 µg of fluorescently labelled aRNA. Although the original experimental design for the microarray hybridizations dictated that a balanced dye-swap be performed using 4 biological repeats, elimination of biological repeat 3 resulted in an unbalanced dye swap experimental design being implemented. There was no obvious reason as to why the control sample for biological repeat 3 yielded such a low amount of aRNA relative to the other samples. The integrity and quality of the RNA sample isolated from biological repeat 3 was comparable to the other biological repeats. However, one possibility could the purity of the RNA sample. According to the MessageAmp™ II aRNA Amplification Kit user manual (Ambion), significant amounts of contaminating DNA, ethanol or salts may adversely affect aRNA synthesis. Alternatively, partially degraded RNA generates shorter cDNA molecules which may affect the average size of the aRNA population and subsequently reduce the yield of aRNA. The latter explanation is unlikely because upon inspection of the RNA obtained from the control sample for biological repeat 3, no degradation was evident (Figure 2.4, B, number 3).

Table 2.4 aRNA yields synthesized from 1.2 µg total RNA for each sample (experiment and control) using the MessageAmp^{TM} II aRNA kit.

Biological Repeat	Sample	aRNA synthesized (µg)
1	Control	30.6
	Experiment	42.0
2	Control	20.0
	Experiment	28.0
3	Control	8.3
	Experiment	45.0
4	Control	23.9
	Experiment	23.0

Following purification and quantitation of the aRNA, Cy-dyes were indirectly coupled to the aaUTPs molecules. Fluorescently labelled aRNA was purified again and the frequency of dye incorporation was calculated using the A260 nm values as determined using the Nanodrop spectrophotometer. Overall, dye incorporation of greater than 200 pmol per sample and a ratio of less than 50 nucleotides/dye molecule was considered optimal for microarray hybridizations (Hedge *et al*., 2000). Fluorescent labelling of all aRNA samples met these criteria (Table 2.5) and was thus deemed suitable for the subsequent microarray hybridizations.

Table 2.5 Characteristics of the fluorescently labelled, purified aRNA used in microarray hybridizations. Fluorescent dye assignment indicated which Cy-dye was used to label each aRNA sample and reflects the dye-swap configuration between biological repeats. The concentration of each aRNA sample was determined spectrophotometrically while the Frequency of Incorporation (FOI) for each aRNA sample was determined empirically using A260nm values.

Biological Repeat	Sample	Fluorescent Label	OD_{550}	OD_{650}	Concentration ($ng.ul^{-1}$)	FOI
1	Control	Green	0.135	-	706.7	41.33
	Experiment	Red	-	0.181	773.3	30.38
2	Control	Green	0.11	-	617.9	38.51
	Experiment	Red	-	0.136	599.4	29.45
4	Control	Red	-	0.253	725.0	46.71
	Experiment	Green	0.120	-	590.0	46.77

2.3.4 Microarray hybridizations and pre-processing of scanned images

The microarrays were queried in a co-hybridization assay using the two fluorescently labelled targets prepared from messenger RNA from the cellular phenotypes of interest (exposure or no exposure to disease elicitors). Pre-hybridization and hybridization conditions as well as wash steps were optimized for high specificity and to minimize cross-hybridization and background signal. Hybridized microarrays were scanned using a dual-laser system capable of analyzing the fluorescent signals from both the Cy3 (532 nm) and Cy5 (635 nm) channels producing separate TIFF images for each (Figure 2.5).

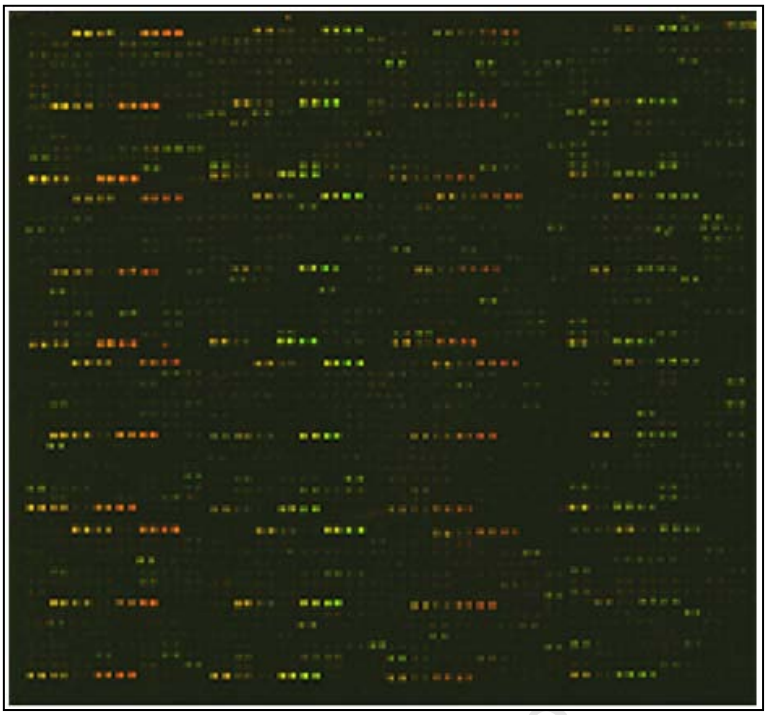


Figure 2.5 Representative TIFF image corresponding to the scanned image of biological repeat 1 cDNA microarray. TIFF images for biological repeats 2 and 4 were similar and are therefore not shown. The microarray contained 1620 *G. gracilis* cDNA probes and 576 Lucidea univeral scorecard controls which were hybridized to 10 µg fluorescently labelled aRNA. Both channels were laser scanned simultaneously (green Cy3 dye at 532 nm and red Cy5 dye at 635 nm), images were overlayerd to generate a false colour image which represented the relative expression levels of each gene on the microarray. Yellow spots represented equal expression of the gene in both treatments whereas predominantly red or green represented up or down-regulation of that gene depending on Cy-dye assignment of the sample (exposure to elicitor or no exposure to elicitor). The contrast on the image was adjusted to allow the majority of the spots in the array to be easily visualized.

As seen in Figure 2.5, there was successful hybrdization between the fluorescntly labelled aRNA and the cDNA probes printed on the slides with minimum background noise. Images were then assessed to clearly identify and delineate each individual microarray spot, to determine the local background intensities for subtraction and to establish the relative fluorescence intensities of each spot. At this stage of the analysis, spots that exhibited fluorescent intensities that failed to meet a set of quality criteria (Appendix D.1) were

automatically flagged within the GenePix software program. The contribution of the flagged spots to normalization of the remaining microarray data was set to zero. An average of 18% of the spots on each of the three microarrays was flagged. A result file was generated which consisted of the total foreground and background fluorescence intensities for both channels (Cy3 and Cy5). The Cy3/Cy5 ratios of each feature were $\log_2$ transformed to generate M-values which were a more informative display of gene expression, i.e. a two-fold induction of gene expression would equal 1 while a half-fold repression would equal –1.

2.3.5 Normalization of microarray data

Before the M-values for each of the spots could be analysed to identify transcriptionally regulated *G. gracilis* genes, the raw microarray data (results file) of each biological repeat generated after image processing had to be normalized. Data normalization was applied to minimize any technical bias such as potential differences in the efficiency of reverse transcription of each RNA sample as well as labelling and detection biases for each fluorescent label (Cy3 and Cy5). Scientific evidence suggests that fluorescent dyes exhibit different quantum yields and/or extinction coefficients and are differentially sensitive to photo-bleaching which may affect observed changes in gene expression (Cox *et al*., 2004; Benes *et al*., 2003; Tseng *et al*., 2002; Yang and Speed, 2002; Wildsmith *et al*., 2001; Worley *et al*., 2000). This phenomenon was circumvented by incorporating the dye swap into the microarray experimental design. However, the dye swap could not be carried out as planned due to elimination of biological repeat 3 from the microarray experiments. Analysis of the fluorescent intensities within both channels without the application of normalization algorithms highlighted the variation between both channels within individual microarrays as well as across the three biological repeats (Figure 2.6). There was a clear labelling bias with the Cy3 (green) dye as reflected by the higher densities relative to the Cy5 (red) dye. In addition, the intensities of the fluorescent signals for both dyes differed across the three biological repeats. Based on these raw data, normalization was obviously required.

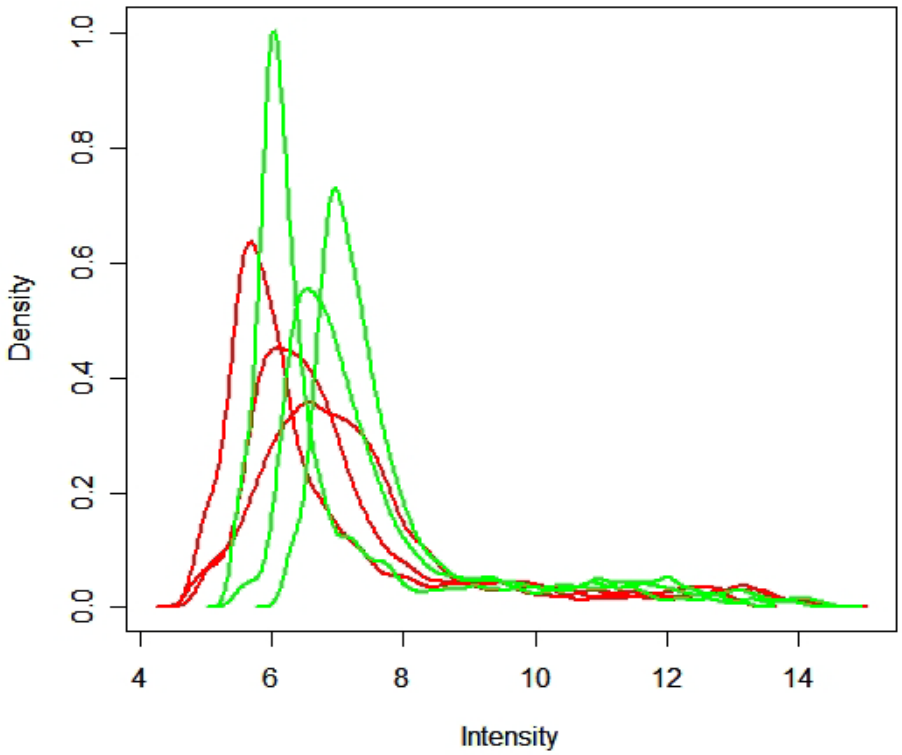


Figure 2.6 Density plot of smoothed empirical densities for the individual Cy3 (green) and Cy5 (red) channels for all three microarrays before data normalization. The plot was generated using the LIMMA package in the 'R' software.

An appropriate normalization strategy had to essentially achieve three aims simultaneously: (i) it had to enable the extraction of biologically significant data from each individual microarray; (ii) it had to adjust the data such that the measured gene expression values across the three biological repeats were comparable and (iii) it had to suit the relatively small density of the microarray used (~1600 cDNAs) since all the assumptions for higher density microarrays would not be met (see Yang *et al*., 2002 and *Hedge et al*., 2000 for reviews). Reference genes whose transcription was not altered in either of the treatments for this experiment could not be used to normalize the microarray data since the amplified cDNA probes printed on the microarray were of unknown identity. Consequently, the Lucidea Universal scorecard controls were included in the construction of the microarray and spiked into each aRNA sample for normalization purposes. Synthetic Lucidea DNA sequences display no sequence homology to *G. gracilis* which made cross-hybridization impossible.

Upon complementary hybridization of the targets to the cDNA probes, the synthetic control sequences were designed to exhibit equal green and red fluorescent intensities in order to generate a normalization factor for each respective microarray slide. A prerequisite to using the Lucidea normalization controls was that the fluorescent signal intensities for each of the *G. gracilis* cDNA probes had to be within the lowest and highest fluorescent intensity values of the normalization controls (Yang *et al*., 2002). However, investigation of the fluorescent intensity data for the Lucidea controls revealed that they exhibited much higher fluorescent intensity values relative to the *G. gracilis* cDNA probes. As a result, they could not be used to normalize the Cy3 and Cy5 channels for this microarray experiment.

Since the original approach (synthetic Lucidea DNA controls) could not be used to generate normalization factors for each microarray slide, an alternative normalization strategy was sought. The Bioconductor package LIMMA was consequently used in conjunction with R software (version 2.3.0) to normalize the microarray data. The LIMMA software package was selected because it was designed specifically for two-colour spotted arrays to generate more stable statistical inferences and improved power for experiments with a relatively small number of microarrays (3 microarray hybridizations in this study) (Smyth, 2004). Moreover, the LIMMA software package could normalize fluorescence intensity data within and between the three separate microarrays. Within microarray normalization employed the robust-spline algorithm which was an empirical Bayes compromise between print-tip loess and global loess normalization. Loess normalization assumes that the majority of the probes on the microarray are not differentially expressed (Smyth, 2004). It does not assume that there are equal numbers of up and down regulated genes or that differential expression is symmetric about zero. Robust-spline normalization of the data resulted in more balanced Cy3 and Cy5 fluorescent intensities within each individual microarray slide (Figure 2.7) relative to un-normalized data (Figure 2.6). The distinct bias detected in the Cy3 (green) channel was significantly reduced by robust-spline normalization but a slight bias was still evident. Even though normalization within each of the three microarrays was successfully applied, the fluorescent intensities between microarrays were not equal. Consequently, gene

expression data between the three microarrays could not be compared and therefore normalization to correct between microarrays was required.

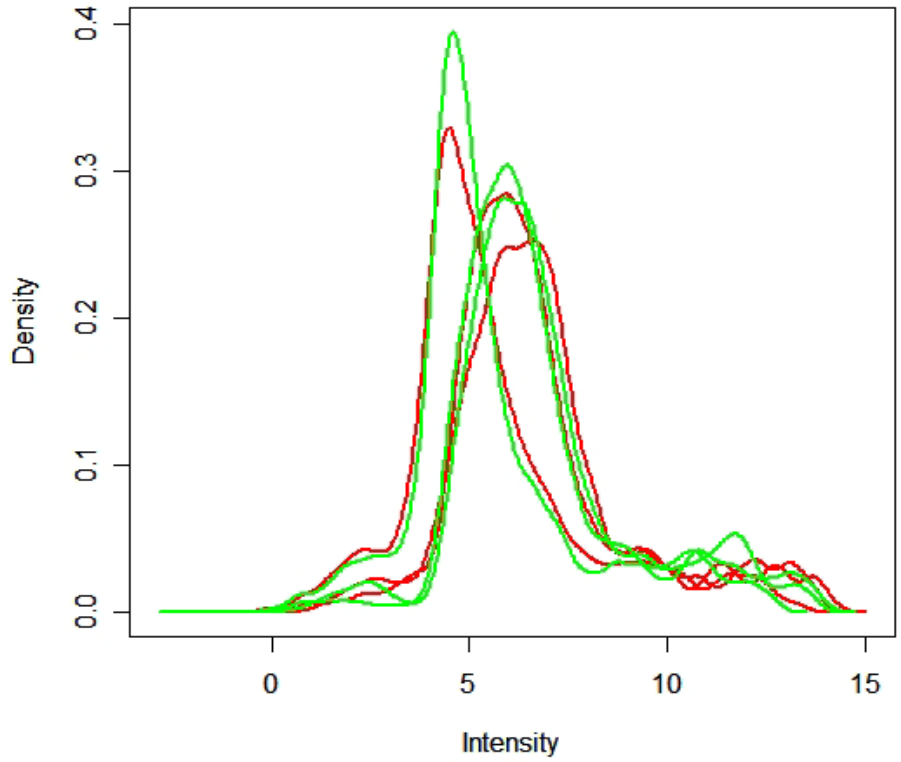


Figure 2.7 Density plot of smoothed empirical densities for the individual Cy3 (green) and Cy5 (red) channels after data robust-spline normalization within the three microarrays using the robust-spline algorithm. The plot was generated using the LIMMA package in the 'R' software.

In order to correct for the differences in fluorescent intensities across the three microarrays, the Aquantile algorithm included in the LIMMA package was applied to the data. The fluorescent intensities within each of the microarray data sets were consequently scaled such that the A-values ($A = ((\log_2(R) + \log_2(G)) / 2$ or the average expression level for that gene across all the microarrays and fluorescence channels) had the same distribution across all microarrays (Smyth, 2004). As evident in Figure 2.8, the fluorescent data intensities in both channels were balanced between the three microarrays after application of the Aquantile

normalization algorithm. Successful normalization of the data within and between the three microarrays served to prevent any one microarray dominating subsequent statistical analyses which would adversely affect accurate identification of differentially expressed genes.

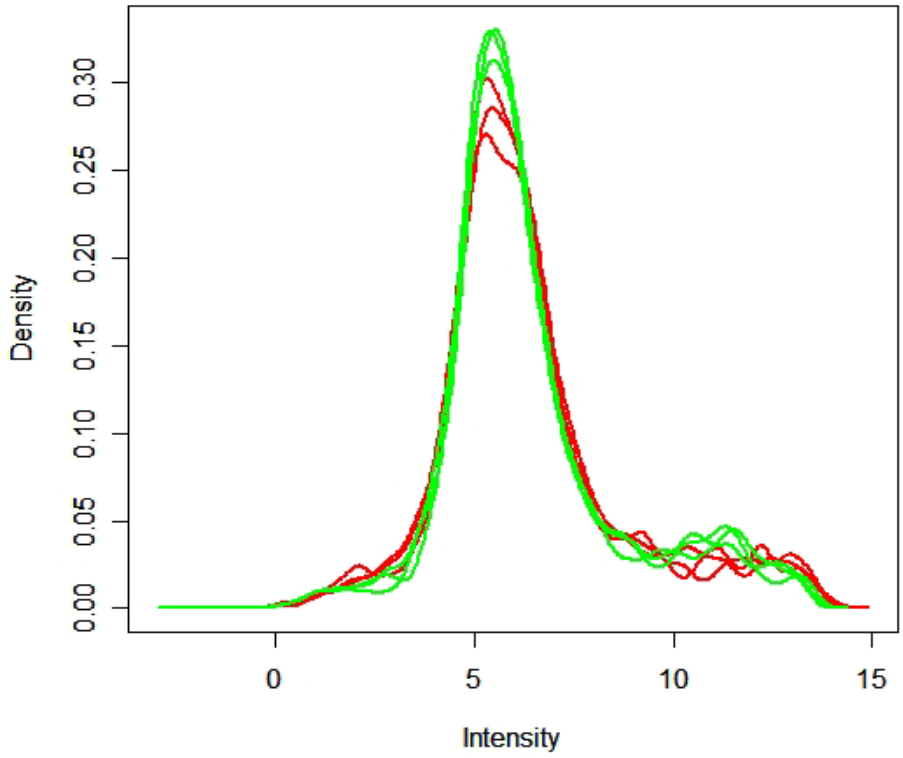


Figure 2.8 Density plot of smoothed empirical densities for the individual Cy3 (green) and Cy5 (red) channels after data normalization within arrays using the robust-spline algorithm and between arrays using Aquantile normalization algorithm. The plot was generated using the LIMMA package in the 'R' software.

2.3.6 Statistical analyses and identification of differentially expressed genes

Following normalization, the typical systemic variations associated with microarray experiments were removed from the data making identification of biologically significant differentially expressed genes possible (Yang *et al.*, 2002; Hedge *et al.*, 2000). Normalized replicate fluorescence data for each gene were merged and average M-values ($\log_2$ fold

change) for all genes were estimated by fitting a linear model using the LIMMA package in the R software. As two independent sample types, namely exposure (experiment) or no exposure to disease elicitors (control), were analyzed in the microarray experiment, it was appropriate to employ a standard *t*-test to compare the expression levels between the samples (Cui and Churchill, 2003; Dopazo *et al*., 2001). When many hypotheses are tested, as is the case in a microarray experiment, the probability that at least one type I error (false positives) is committed can increase sharply with the number of hypotheses (Dudoit *et al*., 2000; Shaffer, 1995; Holm, 1979). A statistical *P* value was produced after *t*-test analysis. This *P*-value was created from the distribution of the moderated t-statistic. It was not adjusted for multiple testing because the moderated t-statistic used to generate it already was (Smyth, 2004). No further adjustment of *P*-values was supported by Dudoit *et al*. (2000) who highlight a common criticism of procedures that control for the type I error too conservatively and state that strong control is not always needed. Within the LIMMA package, empirical Bayes statistics (t and *P* statistics) were computed where the moderated t-statistic was the ratio of the M-value to its standard error. In the current investigation, the null hypothesis set for gene expression analysis was that no genes were differentially expressed when *G. gracilis* was exposed to disease elicitors versus the control. Alternatively stated, any difference in gene expression observed between the two treatments was due to random chance and not differential expression. Thus, a *P*-value of less than 0.05 meant the gene tested was likely to be differentially expressed thereby rejecting the null hypothesis. However, because the *G. gracilis* cDNA targets printed on the microarray were not functionally annotated, the identities and biological roles of these putative defence genes remained unknown. The cDNAs which corresponded to differentially expressed genes had to be sequenced to assist in functional annotation. Differentially expressed genes were ranked by *P*-value to pinpoint expression levels with the highest statistical significance. Thirty genes were observed to have *P*-values less than 0.05.

An alternative approach to identifying differentially expressed genes based on post-normalization fold change was also assessed (Causton *et al*., 2003; Quackenbush, 2002). Essentially, this entailed analyzing changes in gene expression independent of statistical tests.

The global mean (μ) and standard deviation (σ) of the entire microarray data set was calculated as –0.04 and 0.370, respectively. Thus, any gene with an expression ratio (μ - 1.96(σ)) < M-value > (μ + 1.96(σ)) was considered differentially expressed, i.e. M > 0.68 or M < – 0.765. In total, twenty four genes met this condition. Analysis of fold change in transcriptional regulation revealed that the majority of differentially expressed genes were down-regulated at 24 hours post exposure to disease elicitors. Expression of all the genes on the microarray (including differentially regulated genes) ranged between 2-fold induction (M = 1) and 3-fold repression (M = – 1.5). Interestingly, most published studies have used a post-normalization cut-off of two-fold up- or down-regulation to identify genes exhibiting the most significant variation (Quackenbush, 2002; Hedge *et al*., 2000). There is increasing evidence that such arbitrary cut-off selection criteria may be out of context relative to what is biologically relevant in the cell (Baldi and Long, 2001; Quackenbush, 2001; Thomas *et al*., 2001; Woolf and Wang, 2000; DeRisi *et al*., 1996; Schena *et al*., 1996, Schena *et al*., 1995). Mutch *et al*. (2002) and Claverie (1999) suggest that there are inherent problems associated with selecting differential expression based on arbitrary cut-offs. For example, genes with low absolute expression have a greater error in their measured levels and will be selected for by the microarray technique because they tend to numerically meet any given fold change cut-off even if not differentially expressed. The inverse would also be true since highly expressed genes will have less error in their measured levels and may therefore not meet the cut-off, even when they are actually differentially expressed. Therefore, in the context of the current microarray study, the relatively small changes (less than 2-fold changes) in differential expression observed in this study were considered biologically significant because McCarthy and Smyth (2009) suggest that in biological terms, a gene may be considered differentially expressed if a change (even if relatively small) in its expression level leads to a biological effect in the cell. Consequently, the cDNAs which corresponded to differentially expressed genes had to be sequenced in order to establish the putative functions of these genes in the context of a *G. gracilis* defence response.

Selection of differentially expressed genes based solely on fold change is not the best approach since it does not accommodate for background noise or non-specific hybridization

which are typical characteristics of microarray experiments (Baldi and Long, 2001). A drawback of using post-normalization fold-changes is that it is not a statistical test and has no associated level of confidence for selection of differential expression (Cui and Churchill, 2003). On the other hand, using statistical analyses alone to identify differential expression in microarray experiments may be prone to the statistical type I error associated with multiple t-testing. Furthermore, the high cost of repeating microarray experiments limited the number of separate biological repeat hybridizations to a maximum of four which was relatively low. A drawback with a small number of repeats was that inaccurate estimates of variance and a low power in the statistical tests to differentiate differentially expressed genes was more likely (Baldi and Long, 2001). Therefore, in the current microarray study, a combination of fold change and statistical testing was employed to firstly account for the limitations in using either one separately and secondly, to increase the confidence in the genes identified as differentially expressed. It is important to mention that in both approaches, a large number of false positives were observed (~40%). This was not surprising though, because a large number of spots were initially flagged in the pre-processing of the scanned images (Section 2.3.4). Furthermore, flagged spots were not removed from the microarray data set, but merely excluded from normalization of the remaining spots of higher quality (Smyth *et al*., 2005). All false positives were eliminated for further analysis and the findings of both approaches to identify differential expression were merged. Ultimately, a total of 51 genes appeared to be differentially expressed in *G. gracilis* after 24 hours of exposure to disease elicitors (Table 2.6).

2.3.6.1 Sequencing of cDNA inserts and bioinformatics for functional annotation

In order to elucidate the functions and roles of the 51 differentially regulated genes identified in the microarray experiment, their DNA sequences were determined (Appendix F.1) and compared to previously annotated sequences in gene databases. Prior to sequencing, plasmid integrity was assessed (data not shown) and cDNA inserts were PCR amplified (Section 2.2.8). PCR products were gel electrophoresed to assess their quality and success of the PCR reactions (data not shown). Any smearing, absence of PCR product, multiple PCR products or PCR products sized at less than 300 bp (data not shown) resulted in exclusion of the cDNA

fragment for sequencing. Only one cDNA insert (Table 2.6, Clone ID 17) was excluded. Consequently, a total of 50 cDNA inserts which corresponded to significantly differentially expressed genes were sequenced. In addition, it was decided to sequence several cDNA inserts that displayed differential expression which failed to meet the criteria used to detect significant transcriptional changes ([*P*-value < 0.05] or [μ - 1.96(σ)) < M-value > (μ + 1.96(σ)]) after 24 hours post exposure to disease elicitors. Sequencing these cDNA inserts may seem counter-intuitive but it served two purposes in the context of this study: (i) it expanded the genomic data presently available for *G. gracilis* and (ii) it enabled designing DNA primers to assess whether these genes were significantly transcriptionally regulated at an earlier time point (i.e. sooner than 24 hours post-exposure to disease elicitors). The literature suggested that the most significant transcriptional changes could have occurred within the first few hours of exposure to the disease elicitors. Therefore, the experiments planned to validate the microarray results (see Chapter 3) were specifically designed to assay gene expression at several time-points before and including the 24 hour sampling point following exposure to disease elicitors. After cDNA sequencing, all the novel DNA sequences were subjected to homology searches of the NCBI databases using the BLASTx and BLASTn algorithms, as well the InterProScan and KEGG online resources to aid in the functional annotation of each gene (Table 2.6).

Table 2.6 *G. gracilis* genes differentially expressed in response to exposure to disease elicitors for 24 hours. Significant differential expression was defined as any gene with a P-value < 0.05 or ($\mu + 1.96(\sigma)$) < M-value < ($\mu - 1.96(\sigma)$) where $\sigma = 0.370$ and $\mu = -0.04$. Functional annotation of each gene was established by comparing sequence homology to characterized genes and proteins in publicly available databases. The molecular function of each gene was determined according to the guidelines stipulated by Lee *et al*. (1999). M-values represent transcriptional regulation of each gene as determined in the microarray experiment, while positive and negative values correspond to up-regulation and down-regulation, respectively.

Clone ID	M-value	*P*-Value	Putative functional annotation (Organism with highest homology)	E-value	Molecular Function
4	-0.66599	0.014588	Thioredoxin, cytoplasmic (*Griffithsia japonica*)	9.00E-17	Stress response
828	-0.68302	0.015311	Hypothetical trans-membrane trafficking protein (*Xenopus tropicalis*)	1.60E-03	Cell structure
230	-1.18554	0.016235	Conserved hypothetical protein (*Zea mays*)	4.00E-08	Unknown
10	-0.77155	0.016806	Chloroplast hypothetical protein (*Zea mays*)	3.00E-03	Metabolism
34	-0.96098	0.017064	Conserved hypothetical protein (*Vibrio vulnificus* CMCP6)	1.00E-06	Unknown
116	-0.70817	0.017524	pG1 protein (*Lactobacillis jensenii*)	5.00E-09	Cell structure
822	-1.63413	0.019651	No sequence similarity found	-	Unknown
123	-1.56438	0.021275	Chloroplast hypothetical protein (*Zea mays*)	3.00E-09	Metabolism
301	-1.51938	0.022433	Chloroplast hypothetical protein (*Zea mays*)	3.00E-09	Metabolism
450	-1.4145	0.025536	No sequence similarity found	-	Unknown
216	-0.49759	0.027806	Putative 23S ribosomal RNA (*Vigna unguiculata*)	1.00E-03	Protein synthesis
26	-1.30698	0.029448	Chloroplast hypothetical protein (*Zea mays*)	7.00E-07	Metabolism
538	-1.29244	0.030047	Unknown (*Glycine max*)	0.085	Unknown
685	-0.76876	0.03011	Novel asparaginase (*Danio rerio*)	3.00E-13	Metabolism
623	-1.28657	0.030293	Phosphoglycolate phosphatase (*Thalassiosira pseudonana* CCMP1335)	4.00E-61	Respiration
912	-0.95961	0.032714	Predicted protein (*Micromonas* sp. RCC299)	9.00E-03	Unknown
205	0.457117	0.034899	No sequence similarity found	-	Unknown

Table 2.6 *continued*

233	-1.17206	0.035797	Peroxiredoxin-like (Redoxin superfamily) (*Synechococcus* sp. PCC 7335)	4.00E-47	Stress response
19	-1.15722	0.036619	Chloroplast hypothetical protein (*Zea mays*)	6.00E-08	Metabolism
116	-1.12133	0.038731	Conserved hypothetical protein (*Vibrio vulnificus* CMCP6)	1.00E-06	Unknown
314	-0.54625	0.039387	No sequence similarity found	-	Unknown
17	-1.09647	0.0403	Not sequenced (cDNA insert < 300 bp)	-	-
265	-0.6941	0.040684	No sequence similarity found	-	Unknown
542	0.470237	0.041236	No sequence similarity found	-	Unknown
17	-0.91283	0.042007	Putative 23S ribosomal RNA (*Vigna unguiculata*)	2.00E-07	Protein synthesis
422	-0.41742	0.046838	Putative ATP synthase CF0 subunit I (*Vigna unguiculata*)	4.00E-05	Metabolism
1034	1.005998	0.046924	Cytochrome c oxidase subunit I (*Chrysomya megacephala*)	1.3	Metabolism
394	0.986438	0.048571	No sequence similarity found	-	Unknown
924	-0.96696	0.050298	Polyubiquitin (*Griffithsia japonica*)	2.00E-70	Protein processing
671	-0.96578	0.050406	pG1 protein (*Lactobacillis jensenii*)	5.00E-09	Cell structure
488	0.963839	0.050583	Putative 23S ribosomal RNA (*Vigna unguiculata*)	1.00E-08	Protein synthesis
5	0.940597	0.052787	Predicted protein (*Thalassiosira pseudonana* CCMP1335)	0.11	Unknown
134	0.725009	0.053884	tRNA-dihydrouridine synthase (*Micromonas* sp. RCC299)	4.00E-16	Protein synthesis
282	-0.92688	0.054155	Putative cysteine desulphurase (*Ricinus communis*) / glycerol-3-phosphate dehydrogenase	1.00E-06	Metabolism
335	0.897711	0.057249	No sequence similarity found	-	Unknown
264	-0.86058	0.061585	Putative 23S ribosomal RNA (*Vigna unguiculata*)	1.00E-08	Protein synthesis
208	-0.85901	0.061779	Conserved hypothetical protein (*Vibrio vulnificus* CMCP6)	1.00E-06	Unknown
102	-0.84238	0.063893	Putative 23S ribosomal RNA (*Vigna unguiculata*)	6.00E-08	Protein synthesis

Table 2.6 *continued*

1003	0.686956	0.065159	High light inducible protein (*Acaryochloris marina* MBIC11017)	2.00E-06	Stress response
454	-0.82482	0.066243	No sequence similarity found	-	Unknown
252	0.813737	0.067794	SNF2 family chromodomain-helicase DNA-binding protein (*Nectria haematococca* mpVI 77-13-4)	2.00E-14	DNA modification
100	-0.81152	0.068111	pG1 protein (*Lactobacillis jensenii*)	6.00E-09	Cell structure
562	0.811088	0.068173	Hypothetical protein	-	Unknown
221	-0.81076	0.06822	Unknown (*Zea mays*)	3.00E-09	Unknown
470	-0.78763	0.071667	No sequence similarity found	-	Unknown
476	-0.78447	0.072158	GDP-fucose transporter 1 (*Caligus clemensi*)	5.00E-08	Metabolism
1041	0.740829	0.079485	CHCH domain containing protein (*Pyrenophora tritici-repentis* Pt-1C-BFP)	0.004	Protein processing
1010	0.732751	0.080962	Cytidine and Deoxycytidylate deaminase family protein (*Rhodobacterales bacterium* HTCC2654)	0.15	Metabolism
583	-0.72097	0.083189	Hypothetical protein	5.4	Unknown
100	-0.71854	0.083661	pG1 protein (*Lactobacillis jensenii*)	6.00E-08	Cell structure
466	-0.71451	0.084449	Glycolipid transfer protein domain-containing protein 1 (*Homo sapiens*)	6.00E-04	Cell structure
428*	0.350777	0.083533	Putative serine protease-like protein (*Oryza sativa*)	1.00E-24	Protein processing
991*	0.227905	0.085252	Phosphoserine phosphatase	1.00E-16	Metabolism
897*	-0.63652	0.10217	Phosphotidylserine decarboxylase (*Vibrio fischeri* ES114)	2.00E-28	Metabolism
675*	-0.55672	0.12655	Alanine-glyoxylate aminotransferase (*Griffithsia japonica*)	2.00E-95	Respiration

* Represents *G. gracilis* genes that failed to meet the criteria for significant differential expression after 24 hours of exposure to disease elicitors. Functional annotation after cDNA sequencing and literary evidence suggested possible roles in biotic stress responses in other organisms. These genes were therefore earmarked for further investigation in future transcriptional regulation experiments (Chapter 3).

Sequencing and subsequent BLAST analyses (BLASTn and BLASTX) of the cDNA inserts revealed a significant degree of redundancy in the 50 differentially expressed genes (Table 2.6). For example, three separate cDNA inserts displayed sequence homology to a pG1 protein from *Lactobacillus jensenii*; five were homologous to a chloroplast hypothetical protein from *Zea mays* and five more were homologous to a putative 23S ribosomal RNA protein from *Vigna unguiculata*. Redundancy was not a surprising result since the EST libraries used to construct the microarrays were not normalized. This resulted in over-representation of genes that were highly-expressed at the time of EST library creation. Redundancy in the respective cDNA sequences was later confirmed when they were aligned using multiple sequence analysis. In each case, a 100% homology between sequences was observed (data not shown). A large proportion (53%) of *G. gracilis* genes displayed no significant homology to genes in genome databases or alternatively showed significant homology to genes of unknown function (Figure 2.9). This finding was corroborated by Lluisma and Ragan (1997) who showed that out of 200 *G. gracilis* ESTs they constructed, 146 (73%) showed no significant matches to previously characterized genes. In a similar study, Collén *et al*. (2006) observed that even though several *C. crispus* genes were significantly regulated by methyl jasmonate, they could not be functionally annotated. Ho *et al*. (2009) and Teo *et al*. (2008) observed that up to 50% of the genes differentially expressed in *G. changii* could not be functionally annotated either. Altogether these findings suggest that *G. gracilis*, and other red macroalgae, may harbour several unique genes previously uncharacterized with a biotic stress response. Regardless of the evidence suggesting that activated defence mechanisms are conserved in macroalgae and higher organisms (Chapter 1), mechanisms of defence in marine organisms may deviate from the expected model. Further scientific investigation is required to characterize the defensive roles of these novel genes. Classification of the differentially expressed *G. gracilis* genes according to molecular function (Table 2.6) revealed that six major groups were active after 24 hours of exposure to disease elicitors (Figure 2.9). These included genes involved in stress responses, metabolism, respiration, cell structure, protein processing or synthesis and DNA modification. Down-regulated genes encoded various protein products including thioredoxin, asparaginase, phosphoglycolate phosphatase, peroxiredoxin-like, cysteine desulphurase and polyubiquitin. Genes that were up-regulated encoded protein products for a tRNA-dihydrouridine synthase, a high light inducible protein, a SNF2 family chromodomain-helicase DNA-binding protein

and CHCH domain containing proteins. These findings have therefore proved that exposure to disease elicitors lead to transcriptional regulation of various genes in the macroalga *G. gracilis*. In addition, the study provided important clues with regard to the specific function of some of the transcriptionally regulated genes.

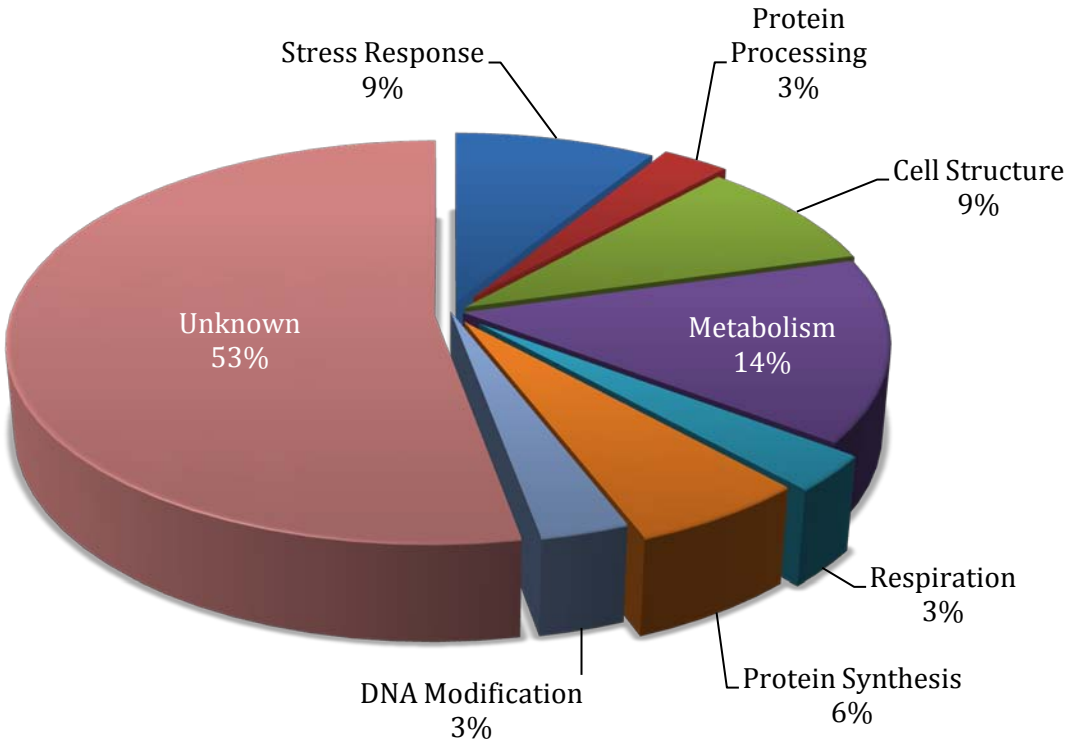


Figure 2.9 Representative pie chart of the distribution of molecular functions (Lee *et al*. 1999) associated with *G. gracilis* genes significantly differentially regulated after exposure to defence elicitors after 24 hours. In total, n = 50 cDNA inserts were sequenced. Redundant sequences were consolidated and the gene was only counted once which reduced the number of sequences used to create this pie chart to *n* = 34.

Another notable observation was that several differentially expressed genes originated from the 18 day nitrogen limitation EST library. This finding suggested a possible overlap in the genetic pathways activated when *G. gracilis* is challenged by biotic (pathogen attack) or abiotic (nitrogen limitation) stresses which is further supported by Mantri *et al*. (2010) and Fujita *et al*. (2006) who propose possible cross-talk between the various signalling pathways active during biotic and abiotic stresses. Furthermore, the choice to include both EST libraries in the construction of the microarray ensured enrichment of stress-related genes which would not necessarily be expressed under normal (non-stress) conditions. However, identification of transcriptionally regulated genes depended solely on whether or not those genes were present on the microarray. This point highlighted one limitation of this study, namely the relatively low density of the microarray. If a gene that was highly regulated at 24 hours post exposure to disease elicitors was not represented as a cDNA probe on the microarray, it would not have been detected in this study. Microarray studies of the red macroalga *Gracilaria changii* (Ho *et al*., 2008; Teo *et al*., 2007) contained almost double the number of cDNA probes (~3300). Collén *et al*. (2006) constructed a cDNA microarray comparable in size to this study (~1920 cDNAs) with the exception that the cDNAs had been sequenced and redundancy was minimized which enabled a more informed selection of genes for the construction of the microarray. Another limitation that arose due to the small density of the microarray was that the complete transcriptome could not be assessed. In contrast, a microarray study of the brown seaweed *Ectocarpus siliculosus* assayed expression of 12250 cDNAs obtained by sequencing clones from six different stress libraries (Dittami *et al*., 2009). Such large coverage of the *E. siliculosus* genome ultimately represented the entire genome and was able to highlight 'the complete transcriptome picture'. Analysis of individual genes, as well as clustering of coordinated genetic responses to the stress under investigation (for comparing independent microarray experiments), can be easily performed with high density microarrays. This point does not suggest that the current study is flawed in any way, but merely implies that clustering of transcriptionally regulated genes according to function is limited with a small data set (34 differentially expressed *G. gracilis* genes). In order to characterize the defence transcriptome of *G. gracilis* in its entirety, the complete nuclear (and plastid) genome will ultimately be required to create higher density microarrays. Nevertheless, this current *G. gracilis* microarray study (i.e. 34 significantly differentially expressed genes) was compared to a microarray study published by Schenk *et al*. (2000).

The latter study utilised cDNA microarrays to examine transcriptional changes in *A. thaliana* following a 72 hour fungal pathogen infection (*Alternaria brassicicola*) or after a 24 hour exposure to the defence-related signalling molecules salicylic acid, methyl jasmonate and ethylene. All the *A. thaliana* genes (~2375) present on the microarray were functionally annotated by Schenk *et al*. (2000), whereas sequence information was only established for the differentially expressed *G. gracilis* genes in this study (Table 2.6). Since the cDNA microarrays in both studies did not represent the entire genome of either organism, it was decided to assess whether differentially expressed *G. gracilis* genes could be matched to a homologous *A. thaliana* gene in order to compare expression patterns in response to the respective biotic stresses. This approach was obviously limited because it failed to (and could not) take into account possible gene isoforms, nor could it predict whether the protein product of each 'homologous' gene had the same sub-cellular localization and biochemical function. Moreover, the biotic stress treatments in the two experiments were fundamentally different. Even though experimental evidence suggests that evolutionarily conserved components of the defence mechanisms of *G. gracilis* and higher plants may be similarly activated (or repressed), it cannot be assumed that homologous genes would exhibit similar transcriptional regulation under the different conditions tested in the two studies. In other words, comparative transcriptome analyses of *G. gracilis* and *A. thaliana* that were both exposed to methyl jasmonate or pathogen infection would be more conducive to drawing valid biological conclusions. Collén *et al*. (2006) recently demonstrated altered transcriptional patterns in another red macroalga (*C. crispus*) following exposure to methyl jasmonate. However, a similar study has not yet been conducted with *G. gracilis*. Thus, similarities and discrepancies in the transcriptional regulation of homologous genes in the two studies are not unexpected. For example, an *A. thaliana* cytosolic thioredoxin (Accession number T43570) was transcriptionally up-regulated in response to fungal infection as well as exposure to salicylic acid and methyl jasmonate, but repressed by ethylene (Schenk *et al*., 2000). A second *A. thaliana* cytosolic thioredoxin (Z18522) was down-regulated in response to fungal infection but up-regulated upon exposure to methyl jasmonate and ethylene. In this study, a *G. gracilis* cytoplasmic thioredoxin was down-regulated after 24 hours of exposure to disease elicitors. Polyubiquitin genes displayed significant transcriptional regulation in both studies. One *A. thaliana* homologue (H37347) was down-regulated in response to all treatments. Similarly, a *G. gracilis* polyubiquitin

homologue was down-regulated in response to the presence of disease elicitors. In contrast, a different *A. thaliana* polyubiquitin homologue (N64950) was up-regulated during pathogen infection and exposure to salicylic acid and methyl jasmonate but down-regulated upon exposure to ethylene (Schenk *et al*., 2000). Genes encoding polygalacturonase exhibited significant transcriptional regulation in both studies. An *A. thaliana* homologue (R29820) was up-regulated during pathogen infection and after exposure to salicylic acid, methyl jasmonate and ethylene. In contrast, the *G. gracilis* homologue was down-regulated after 24 hours of exposure to elicitors. A *G. gracilis* serine protease-like gene transcript was found to be up-regulated in the current study. While serine protease-like gene homologues were not present on the *A. thaliana* microarray, mRNA transcripts encoding other proteases, i.e. a cysteine protease (T04773) and an aspartic protease (T75975), displayed significant transcriptional regulation during pathogen infection and exposure to defence-related signalling molecules (Schenk *et al*., 2000). No other *A. thaliana* genes were identified in the Schenk *et al*. (2000) study that are homologues of the remaining differentially expressed *G. gracilis* genes identified in this study (Table 2.6), further emphasizing the importance of comprehensive genome coverage in microarray experiments. Nevertheless, it is possible to extract biologically relevant data from comparative microarray studies involving *G. gracilis* if homologous genes are present on both sets of cDNA microarrays and/or if transcriptional changes in response to similar biotic or abiotic stresses are assessed.

It was obvious that comparison between the current *G. gracilis* microarray study (i.e. 34 significantly differentially expressed genes) and an *A. thaliana* microarray study (Schenk *et al*., 2000) was limited due to the low density of the *G. gracilis* microarray. Therefore, the transcriptional response of *C. crispus* following exposure to methyl jasmonate exposure (Collén *et al*., 2006) was compared to the transcriptional response observed in *A. thaliana* following methyl jasmonate exposure. It was thought that any similarities in transcriptional regulation patterns between another red macroalga (*C. crispus*) and higher plants would support the similarities observed between *G. gracilis* and higher plants. Cheong and Choi (2003) demonstrated that the genes up-regulated by methyl jasmonate in higher plants typically included those involved in jasmonate synthesis, cell wall formation, secondary metabolism and stress or defence proteins. Following exposure to methyl jasmonate, transcriptional up-regulation of potential stress genes was similarly observed in *C. crispus*

(Collén *et al*., 2006). These included glutathionine *S*-transferae (GST), heat shock protein (HSP), and an oxidoreductase protein (Taki *et al*., 2005). In a different study, jasmonic acid, a member of the jasmonate class of plant hormones, was also shown to cause an increase in mRNA transcript levels corresponding to epimerase, GST and antioxidant genes (catalase and manganese superoxide dismutase (Mn-SOD)) in *A. thaliana* (von Rad *et al*., 2005) which was corroborated in *C. crispus* (Collén *et al*., 2006). Interestingly, a putative stress gene of *G. gracilis* (high-light inducible protein) was transcriptionally up-regulated following exposure to disease elicitors for 24 hours. However, two other genes annotated to stress responses, i.e. thioredoxin and peroxiredoxin-like, were both down regulated after 24 hours. In addition to the genes transcriptionally up-regulated following exposure to methyl jasmonate in *A. thaliana* and *C. crispus*, the transcription of various other genes was down-regulated. These included genes predominantly involved in energy conversion and general metabolism which suggested that exposure to methyl jasmonate resulted in a proclivity toward defence rather than growth in red macroalgae and higher plants (Collén *et al*., 2006). Similarly, a pattern of transcriptional down-regulation for *G. gracilis* genes involved in general metabolism (i.e. chloroplast hypothetical protein, asparaginase, putative ATP synthase CFO subunit I, putative cysteine desulphurase, GDP-fucose transporter and PSD) (Table 2.6) was observed following exposure to disease elicitors which corroborated the parallels observed between *A. thaliana* and *C. crispus* following exposure to methyl jasmonate. However, the specific characterisation of methyl jasmonate signalling in *G. gracilis* is required to draw more conclusive results. In addition, a greater proportion of *G. gracilis* stress genes need to be represented in future microarray experiments.

2.3.7 Possible roles of functionally annotated *G. gracilis* genes in the context of activated defence

Functionally annotated genes which displayed differential expression and significant E-values ($< 1e^{-06}$) for sequence homology will each be briefly discussed in order to propose possible biological roles for their protein products in an activated defence response in *G. gracilis*. Obviously, genes associated with unknown biological functions or hypothetical protein products cannot be discussed further. Differential expression of these un-annotated

genes suggests that *G. gracilis* may have several unique genes which constitute its activated defence responses. For the sequenced genes that were not significantly differentially expressed after 24 hours of exposure to disease elicitors, any published scientific evidence that links them to activated defence in other organisms warranted further investigation because qPCR analyses were planned to not only validate the microarray results for a few genes, but also to assess transcriptional regulation at earlier time-points relative to 24 hours of exposure to disease elicitors (Chapter 3).

2.3.7.1 Significantly differentially expressed G. gracilis genes

Thioredoxin proteins belong to the thioredoxin family which include all proteins that exclusively encode a thioredoxin (TRX) domain. Thioredoxins are small (~12 kDa), ubiquitous proteins capable of altering the redox state of target proteins via reversible oxidation of its active site (a dithiol present in a CXXC motif) partially exposed on the surface of the protein. The active site reduces disulphide bonds in the target protein which results in the formation of a disulphide bond in the active site of thioredoxin. Thioredoxin is then converted back to its active form by a thioredoxin reductase enzyme which uses NADH or ferredoxins as reducing equivalents (Baumann and Juttner, 2002). Scientific evidence has implicated thioredoxins in several cellular processes, including regulation of various enzymes (Arnér and Holmgren, 2000; Holmgren, 1989), modulation of transcription factors (Hirota *et al*., 1999; Schenk *et al*., 1994), as hydrogen and electron donors (Holgrem, 1989), protection of cellular components against oxidative damage via the formation of disulphides or for the direct reduction of reactive oxygen species and other antioxidant proteins such as peroxiredoxin (Arnér and Holgrem, 2000; Chae *et al*., 1994). A *G. gracilis* cDNA clone exhibiting significant sequence homology to a cytosolic (type *h*) thioredoxin from *Griffithsia japonica* was down-regulated (M = –0.67) after 24 hours of exposure to disease elicitors. Cytoplasmic localization of the *G. gracilis* thioredoxin was confirmed using bioinformatics services available at http://www.expasy.ch. ChloroP, Predotar and TargetP all failed to detect transit peptides for localization to plastids, the endoplasmic reticulum or mitochondria (Appendix C.5). Multiple forms of thioredoxin *h* exist in plants which can be further classified into three distinct groups, designated I, II and III (Gelhaye *et al*., 2005). In

Arabidopsis thaliana, between five and eight distinct isoforms have been identified (Laloi *et al*., 2004; Meyer *et al*., 2002 and Rivera-Madrid *et al*., 1995), while three have been detected in wheat (Serrato *et al*., 2001 and Gautier *et al*., 1998). To date, the specificities and functions of the different isoforms in plants remain largely unknown (Maeda *et al*., 2003). However, certain physiological functions have been ascribed to type *h* thioredoxins in plant cells. Wheat thioredoxins have been shown to be involved in seed germination (Besse and Buchanan, 1997), for the specific reduction of wheat gliadins and glutenins (Kobrehel *et al*., 1992), for the activation of thiocalsin, a calcium dependent protease (Besse *et al*., 1996) and to reduce inhibitors of α-amylase and limit dextrinase (Wong *et al*., 1996). In tobacco and *Arabidopsis*, thioredoxin *h* mRNA was highly expressed in rapidly growing cells linking it to proper cellular development (Marty *et al*., 1993; Rivera-Madrid *et al*., 1995). Mouaheb *et al.* (1998) demonstrated that specific *Arabidopsis* type *h* thioredoxins were able to restore sulphate assimilation and confer hydrogen peroxide tolerance in yeast cells. Laloi *et al.* (2004) showed that a cytosolic thioredoxin in *A. thaliana* (AtTRXh5) was directly involved in the response to pathogens and oxidative stress. In mammals, the cytosolic thioredoxin has also been shown to be essential in oxidative stress (Arnér and Holmgren, 2000). Baumann and Juttner (2002) suggest that type *h* plant thioredoxins are part of the larger NADPH/NADP-thioredoxin reductase/thioredoxin or (NTR/Trx) system in which NADPH acts as the source of reducing power for the NADPH-thioredoxin reductase catalyzed reduction of thioredoxin. In the context of this study (defense response of *G. gracilis*), the cytoplasmic thioredoxin regulated by disease elicitors could be essential for protection against oxidative stress or in the regulation of other defence proteins. In support of a role in protection against oxidative stress, Weinberger *et al*. (1999) showed that agar oligosaccharides rapidly induce release of ROS in *G. gracilis*.

Several cDNA clones, all down-regulated (M = –0.81) after 24 hours of exposure to disease elicitors, displayed significant homology to a polygalacturonase (pG1) protein. Polygalacturonases belong to one of the largest hydrolase protein families and catalyze the hydrolysis of specific linkages in galacturonides and other polysaccharides (Markovic and Janecek, 2001; Torki *et al*., 2000). In land plants, cellulose (β-1,4 linked glucose units) is the

only structural polysaccharide providing support to the cell wall. In marine algae, cellulose accounts for a much smaller proportion (Frie and Preston, 1961). Algal cellulose is rarely a pure β-1,4 glucan and often contains sugars other than glucose, e.g. xylose (Mackie and Preston, 1974). Instead of cellulose, the skeletal polysaccharides in macroalgal cell walls are typically xylans (predominantly β-1,3 linked) and mannans (predominantly β-1,4 linked) which provide structural support and essential physio-chemical properties (Baldan *et al*., 2001; McCandless, 1981). Furthermore, in red macroalgal cell walls, the intercellular matrix in which the skeletal fibres are embedded is composed of flexible chains of differently sulphated galactans whose structure is based on alternating β-D-galactopyranose and α-L-galactopyranose residues (Murano, 1995; Knutsen *et al*., 1994). This polymer is referred to as agar which in some instances may be charged in *Gracilaria* species, suggesting a possible role in controlling ion exchange between the medium and the cell cytoplasm (Iyer, 2002). The main biological function of agar however, is to maintain the structural integrity of algal cell walls and to provide mechanical strength to the algal thalli (MacLachlan, 1985). The functions and regulation of cell wall hydrolytic enzymes, such as polygalacturonases, are therefore essential since they are responsible for the cleavage of bonds between polymers that constitute the cell wall. In plants, polygalacturonases are involved in various physiological functions which include cell elongation, removal of cells at the root tips, decay and softening of fruit, abscission of organs, seed germination and other developmental processes (Roberts *et al*., 2002; Cosgrove. 2000; del Campillo, 1999; Rose and Bennett, 1999; Hadfield and Bennett, 1998). Fleurence (1999) highlighted the ability of polygalacturonases to enzymatically degrade algal cell walls. In the context of a *G. gracilis* defence response, increased activity of polygalacturonases may structurally alter cell walls for the purposes of cell-cell adhesion and cell wall loosening in order to discard pathogen-infected cells or cells targeted for apoptotic cell death. After 24 hours of exposure to disease elicitors, the mRNA transcripts for pG1 were all transcriptionally down-regulated. If this down-regulation was associated with a decreased amount of pG1 protein in the macroalgal cell, it could have resulted in strengthening of the cell wall to counter-act agarolytic enzymes released by the pathogen. This is an important point because the mechanism of activating a defence response in this experiment was based on the action of agarolytic pathogens which secrete enzymes capable of degrading the host's agar cell wall. Transcriptional repression of *G. gracilis* polygalacturonase after 24 hours of exposure to disease elicitors may therefore be interpreted

as a host defence response which results in strengthening the cell wall to inhibit pathogen proliferation from the site of infection.

A cDNA clone with significant homology to a novel asparaginase was transcriptionally down-regulated (M = –0.77) after 24 hours of exposure to disease elicitors. Asparaginase has been described as a major nitrogen transport and storage compound in higher plants (Atkins *et al*., 1975). L-asparaginase catalyzes the cleavage of L-asparagine to aspartic acid and ammonia. The amino-acid L-asparagine is required in several important metabolic processes such as the citric acid cycle or glucogenesis (Dunlop and Roon, 1975). Asparaginase has also been implicated in the mobilization of the amide nitrogen for utilization in protein synthesis and in the synthesis of the amino acid (asparagine) in higher plants (Paul and Cooksey, 1981). Accumulation of free amino acids in diseased plants tissues has been reported (Uritani and Stahmann, 1961; Rohouringer *et al*., 1958; Van Gundy and Walker, 1957). In the context of an activated macroalgal defence response, Collén *et al*. (2007) observed transcriptional repression of genes involved in general metabolism. A similar biological phenomenon could occur in *G. gracilis*, i.e. repression of nitrogen metabolism genes. Furthermore, in support of this hypothesis, all genes functionally annotated to metabolism were found to be transcriptionally down-regulated (Table 2.5).

Phosphoglycolate phosphatase (PGPase) is a key enzyme required for photorespiration in photosynthetic organisms. Phosphoglycolate is synthesized through the activity of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) and is essential for the growth of photosynthetic organisms in the light (Mamedov *et al*., 2001; Bellanger and Ohgren, 1987). PGPase catalyzes the hydrolysis of phosphoglycolate. PGPase-deficient mutants isolated from the green alga *Chlamydomonas reihnhardtii* were not able to grow under ambient air-equilibrated conditions but required higher levels of CO_2 (Suzuki *et al*., 1990). This was attributed to the accumulation of phosphoglycolate during photosynthesis which strongly inhibited the enzyme triosephosphate isomerase (Suzuki *et al*., 1999; Norman and Colman, 1991; Anderson, 1971; Wolfenden, 1970). The regulatory mechanisms of PGPase activity

remain unclear although scientific evidence implicate regulation by environmental factors such as CO_2 concentration and light (Suzuki, 1995; Marek and Spalding, 1991; Suzuki *et al*., 1990; Baldy *et al*., 1989). This study has therefore shown a possible role of PGPase in activated defence. After 24 hours of exposure to the disease elicitors, PGPase was transcriptionally down-regulated (M = –1.29) almost three fold. In other words, the rate of photosynthesis could be decreased in response to a pathogen attack in order to inhibit growth and proliferation of the pathogen.

Peroxiredoxins (PRXs) are a ubiquitous family of multifunctional antioxidant thioredoxin-dependent peroxidases (Immenschuh and Baumgart-Vogt, 2005) required to protect cells against oxidative stress and mediate signal transduction cascades that use hydrogen peroxide as a second messenger molecule (Dietz, 2003; Wood *et al*., 2003a). The first PRX identified was the thiol-specific antioxidant protein (TSA) required for antioxidant protection of glutamine synthetase activity and DNA integrity (Kim *et al*., 1998; Chae *et al*., 1994). Furthermore, thioredoxin was shown to act as its electron donor (Chae *et al*., 1999; Chae *et al*., 1994) making it thioredoxin-dependent. The alkylhydroperoxide reductase AhpC was subsequently identified (Storz *et al*., 1989). Deletion of *AhpC* in *E. coli* caused hypersensitivity of the cells to killing by cumene hydroperoxide (Dietz, 2003). Changes in phosphorylation and cellular redox state have been shown to regulate peroxiredoxins. Plant PRXs can generally be divided into four three distinct classes: 1-Cys Prx, 2-Cys Prx, Type II Prx and PrxQ (Dayer *et al*., 2008; Dietz, 2003). All share the same catalytic mechanism where the active-site cysteine residue is oxidized to a sulphenic acid by a wide variety of peroxide substrates (e.g. hydrogen peroxide, organic hydroperoxides and peroxynitrite) (Hoffman *et al*., 2002). All peroxiredoxin isoforms contain at least one conserved catalytic cysteine residue. At least one putative cytosolic isoforms can use both thioredoxin and glutaredoxin as an electron donor, but the chloroplastic isoforms depend on reduced thioredoxin (Rouhier and Jacquot, 2002). A cDNA clone displayed significant homology to the peroxiredoxin (PRX)-like 2 family (from here on referred to as PRX2F). PRX2F proteins show sequence similarity to PRXs. Members of this group contain a cysteine-X-X-cysteine (CXXC) motif (similar to thioredoxin) which was conserved in the *G. gracilis* cDNA clone. Although the second

cysteine in the motif corresponds to the active cysteine of PRX, PRX2F lacks two amino acid residues present in the catalytic triad of PRX (Marchler-Bauer *et al*., 2008). In the context of macroalgal activated defence, Weinberger *et al*. (1999) showed that after exposure to agar or agar oligosaccharides, *G. gracilis* rapidly responded via an oxidative burst, increased respiration and halogenating activity. Therefore, transcriptional activation of antioxidant genes, such as PRX2F, is an expected physiological response. However, at 24 hours post exposure to disease elicitors, the PRX2F gene appeared to be more than 2-fold repressed (M = –1.17). Since elicitors resulted in a rapid (within minutes) release of ROS in *G. changii* (Weinberger *et al*., 1999), it is possible that transcriptional activation of the *G. gracilis* PRX2F gene occurred long before the 24 hour time-point which was tested in the microarray experiment. This hypothesis will be further investigated in Chapter 3.

Ubiquitin is a highly conserved protein (composed of 76 amino acids) found in all eukaryotic organisms. It may exist freely, in repeating units (poly-ubiquitin) or covalently attached to target proteins from the nucleus, cytosol or cell surface (Welchman *et al*., 2005). The process of ubiquitination thus refers to the formation of a covalent bond between the C-terminus (Gly_{76}) of ubiquitin and a lysyl amino group in the target protein (Smalle and Vierstra, 2004). Poly-ubiquitination has been implicated in non-lysosomal ATP-dependent selective proteolysis (Zhou and Ragan, 1995); plant signalling pathways mediating responses to hormones, light, sucrose, developmental cues; and recent evidence suggests that ubiquitination may also play an important role in plant disease resistance (Devoto *et al*., 2003; Ellis *et al*., 2002). Mono-ubiquitination of a target protein may result in non-proteolytic changes such as histone modification, protein activation and localization or protein-protein interactions (Schnell and Hick, 2003). The biochemical process of ubiquitination is operated by a multi-enzyme system that consists of ubiquitin-activating (E1), -conjugating (E2) and –ligating (E3) enzymes (Glickman and Ciechanover, 2002). Polyubiquitin gene expression has been detected in incompatible interactions and is required for the recycling of proteins as a result of oxidative stresses (Hammond-Kosack and Jones, 1996). Therefore, since a cDNA clone with significant sequence homology to polyubiquitin was repressed approximately two-fold (M = – 0.96) after 24 hours of exposure to disease

elicitors, this *G. gracilis* polyubiquitin may function contrary to the previously established roles of other polyubiquitin proteins during activated defence. Alternatively, 24 hours after exposure to the disease elicitors may have been too late to detect transcriptional up-regulation of this gene. Thus, elucidation of its molecular function will require further scientific investigation.

A cDNA clone with significant sequence homology to a tRNA-dihydrouridine synthase protein was observed. After 24 hours, the tRNA-dihydrouridine synthase was transcriptionally up-regulated (M = 0.73) in response to exposure to disease elicitors. The protein tRNA-dihydrouridine synthase catalyzes the synthesis of dihydrouridine, a modified base found in the D-loop of most tRNAs (Bishop *et al*., 2002). It is most likely required for the correct folding of tRNAs and indirectly involved in the synthesis of proteins since tRNAs transfer specific amino acids to growing polypeptide chains during translation. Despite transcriptional activation of this enzyme in *G. gracilis* during an activated defence response, its specific role remains unclear.

A cDNA clone displaying significant sequence homology to both a cysteine desulphurase as well as a glycerol-3-phosphate dehydrogenase was transcriptionally down-regulated (M = –0.93) after 24 hours of exposure to disease elicitors. Cysteine desulphurase is a pyridoxal 5'–phosphate-dependent enzyme that catalyzes the conversion of L-cysteine to L-alanine and sulphane sulphur via the formation of a protein-bound cysteine persulphide intermediate on a conserved cysteine residue. Evidence suggests that bacterial cysteine desulphurases are important for the biosynthesis of Fe-S clusters, thiamine, thionucleosides in tRNA, biotin, lipoic acid and NAD (Marquet, 2001; Begley *et al*., 1999). They are also involved in cellular iron homeostasis and in the biosynthesis of selenoproteins. Since they are capable of providing various biosynthetic pathways with sulphur / selenium-containing compounds, they could be essential for the production of cofactors or for the bioconversion of compounds (Mihara and Esaki, 2002) required in the activated defence response of *G. gracilis*. Glycerol-3-phosphate dehydrogenase is found in several eubacteria and

eukaryotes. It catalyzes the reversible conversion of dihydroxyacetone phosphate to glycerol-3-phosphate which is an important step in the flow of carbon between glycolytic intermediates and glyceride lipids (Wang *et al*., 1994). Yeast cells grown under osmotic stress were shown to accumulate glycerol (Brown, 1978). Glycerol was thought to act as a protective agent for osmosensitive proteins since yeast cells grown in high salt concentrations exhibited an increased intracellular glycerol concentration. In addition, an increased activity of cytoplasmic glycerol-3-phosphate dehydrogenase was observed relative to yeast cells grown in normal salt concentrations (Rep *et al*., 1999; Blomberg and Adler, 1989; Edgley and Brown, 1983). Therefore, glycerol-3-phosphate dehydrogenase activity has been linked to adaptation to abiotic stress. In the context of this study, glycerol-3-phosphate dehydrogenase has been shown to be transcriptionally regulated in the adaptive mechanisms to biotic stress in macroalgae as well.

A cDNA clone with significant sequence homology to a high light-inducible protein was transcriptionally up-regulated (M = 0.66) after 24 hours of exposure to disease elicitors. High light-inducible proteins (HLIPs) have been implicated in photo-protection during exposure to high irradiance (Salem and van Waasbergen, 2004). This is achieved by absorption of light and subsequent charge transfer and separation of the excitation energy to photochemical reaction centres (Heddad and Adamsak, 2000). The major photosystem II antenna for Rhodophyta is the phycobilisomes (Heddad and Adamsak, 2000). Phycobilisomes are made up of the photosynthetic centre polypeptide D1 and the D2 polypeptide which forms a heterodimeric reaction centre (He *et al*., 2001). In a previous microarray study on *G. gracilis* (Lebi, 2006), the author showed that a HLIP mRNA transcript was down-regulated after 18 days of nitrogen limitation. In contrast, this study revealed that a HLIP mRNA was transcriptionally induced (M = 0.66) after 24 hours of exposure to disease elicitors. Weinberger *et al*. (1999) reported that prolonged exposure to neoagarohexaose (disease elicitor) caused bleaching of *G. changii* thallus tips. When exposure to elicitors was combined with exposure to light, a more severe, enhanced production of ROS was evident. There are several ways in which excess, absorbed light energy can be harmful to photosynthetic organisms. For example, accumulation in light-

harvesting complexes and reaction centres may promote the formation of ROS which may damage cellular components and disrupt cell viability (Halliwell and Gutteridge, 1999). However, there is a rapid restoration of photosystem II function following excess light-damage which implies regulation of the repair system (Melis, 1999). Repair processes include degradation of damaged D1 proteins, *de novo* synthesis of D1 on chloroplast ribosomes, processing of newly synthesized D1, association of D1 with chlorophyll and its reaction centre partner (D2 protein) and assembly of the heterodimeric complex with other photosystem II polypeptides (Melis, 1999; Critchley and Russell, 1994; Aro *et al*., 1993). Both algae and vascular plants have evolved mechanisms for photo-acclimation that favour survival in high light (Demmig-Adams and Adams, 1992) which include changes in the composition of light-harvesting complexes (Niyogi, 1999; Durnford and Falkowski, 1997), redistribution of excitation energy between photosystems (Fujita, 1997; Biggins and Bruce, 1989) and stabilization of photosynthetic membranes (Harvaux and Niyogi, 1999). In this context, the transcriptional up-regulation of a high light-inducible mRNA transcript seems valid if the effect of prolonged exposure (24 hours) to disease elicitors is compounded by exposure to light which results in increased ROS production. Furthermore, several protein families related to light harvesting complex II proteins, such as the early light-inducible proteins in higher plants, ferns and algae as well as HLIPs in cyanobacteria have been shown to play important roles in adapting to high light-stress, cold stress, nutrient deprivation and photo-oxidative stress (Montane *et al*., 1999; 1997; Adamska, 1997).

A cDNA clone with significant sequence homology to a SNF2 family chromodomain-helicase DNA-binding protein was transcriptionally up-regulated (M = 0.81) in response to exposure to disease elicitors. SNF2 family chromodomain-helicase DNA-binding proteins are a group of highly conserved ATP-dependent enzymes. Their regulation is highly coordinated and involves the recruitment of transcriptional machinery together with modifications in the structure of chromatin. Proteins are divided into two distinct classes, namely, those that mediate post-translational histone modification and those that utilize energy derived from ATP hydrolysis to alter histone-DNA contacts within the nucleus (Marfella and Imbalzano, 2007; Phelan *et al*., 2000; Corona *et al*., 1999). In certain cases,

chromatin remodelling machinery may function as transcriptional co-activators via the alteration of chromatin fluidity that allows transcription factors to access packaged DNA. In other cases, transcription can be hindered when chromatin is tightly re-assembled (Kingston and Narlikar, 1999). Other physiological roles of chromatin-remodelling factors include DNA repair or homologous recombination (Lusser and Kadonaga, 2003). In the context of a *G. gracilis* defence response, transcriptional up-regulation of a chromodomain-helicase DNA binding mRNA transcript could indirectly effect the transcription of genes by altering chromatin structure.

GDP (Guanosine 5-diphosphate)-fucose transporters are essential components in fucose metabolism. A cDNA clone displayed significant sequence homology to such a protein was transcriptionally down-regulated (M = –0.78) after 24 hours of exposure to the disease elicitors. GDP-fucose is synthesized in the cytosol via two pathways, namely the salvage pathway and the *de novo* pathway (Moriwaki *et al*., 2007). Free L-fucose is converted to GDP-fucose by GDP-pyrophosphorylase while newly synthesized L-fucose is generated via three reactions catalyzed by GMD (GDP-mannose 4, 6-dehydratase) and FX (GDP-4-keto-6-deoxy-mannose-3, 5-epimerase-4-reductase) (Ohyama *et al*., 1998; Sullivan *et al*., 1998; Tonetti *et al*., 1996). GDP-fucose is subsequently transported from the cytosol to the Golgi lumen by GDP-fucose transporters to fucosylate various oligosaccharides and proteins. The reaction product, GDP, is converted by a luminal nucleotide diphosphatase to guanosine-5-monophosphate (GMP) and inorganic phosphate. The GMP is then exported back into the cytosol via an anti-port system that is coupled with the transport of GDP-fucose whereas the phosphate may be removed from the Golgi lumen via the Golgi anion channel (Hirschberg *et al*., 2001; Nordeen *et al*., 2000). Fucosylation is thus the process of adding of fucose sugar units to a target molecule and is a form of glycosylation. It is one of the most common modifications involving oligosaccharides such as glycoproteins and glycolipids (Hirschberg *et al*., 1998). Furthermore, fucosylation of glycoproteins may regulate the biological functions of signalling molecules or receptors (Miyoshi *et al*., 2007). Glycosylation, sulphation and phosphorylation of secretory and membrane-bound proteins, proteoglycans and lipids all take place in the lumens of the Golgi apparatus and endoplasmic reticulum and

is essential for post-translational modification of many cellular components (Berninsone and Hirschberg, 2000; Hirschberg *et al*., 1998). In the context of an activated *G. gracilis* defence response, transcriptional repression of the GDP-transporter gene in response to disease elicitors may indirectly affect the degree of post-translational modification of proteins and other cellular components present in the cell in order to activate the defence response.

A cDNA clone with significant sequence homology to a glycolipid transfer protein domain was transcriptionally down-regulated (M = –0.71) after 24 hours of exposure to disease elicitors. Nonspecific lipid transfer proteins (ns-LTPs) represent a protein family that is ubiquitous in plants (Kader, 1997; 1996). These proteins are generally soluble, cysteine-rich, relatively small (ranging between 91 – 95 amino acids) and are able to bind hydrophobic molecules such as fatty acids or fatty acid derivatives (Blein *et al*., 2002; Guerbette *et al*., 1999). The nsLTP fold is characterized by a domain composed of four α-helices held together by a network of four conserved disulphide bridges. This fold presents a large internal tunnel-like cavity which can accommodate different types of lipids (Salcedo *et al*., 2007). These proteins are able to exchange glycerophospholipids or galactolipids between membranes *in vitro* to organelles which cannot synthesize phospholipids (Kader, 1996; Yamada, 1992). In plants, there may potentially be several different isoforms of ns-LTPs present in different tissues (Dubreil *et al*., 1998; Vignols *et al*., 1994). Evidence suggests that they may also be synthesized as precursor proteins with N-terminal signal localization sequences for targeted secretion to areas such as the cell wall (Kader *et al*., 1996; Bernhard *et al*., 1991; Sterk *et al*., 1991; Tchang *et al*., 1988) (Salcedo *et al*., 2007; Garcia-Olmedo *et al*., 1995). Biological functions of proteins containing nsLTP folds include the transport of cuticular components (Meijer *et al*., 1993; Sterk *et al*., 1991); formation of protective surface layers against fungal or bacterial pathogens to inhibit growth (Cammue and Broekaert, 1996; Molina *et al*., 1993; Terras *et al*., 1992) and participation in oxidative metabolism of glyoxysomes (Tsuboi *et al*., 1992). Furthermore, all ns-LTPs isolated from plant species to date (e.g. maize, barley, spinach, *A. thaliana*, radish and broccoli) display antibiotic activity with high degrees of specificity (Molina *et al*., 1993). The expression of *lipid transfer protein* genes is transcriptionally regulated in response to pathogen infection (Molina *et al*.,

1993; Garcia-Olmedo *et al*., 1995). In a separate study, Moldonado *et al*. (2002) implicated a putative LTP in systemic resistance signalling of *A. thaliana*. However, since the *G. gracilis* gene encoding a glycolipid transfer protein domain identified in this study was observed to be transcriptionally repressed in response to disease elicitors, it is possible that it possess a different function. Alternatively, transcriptional up-regulation of this gene may have occurred prior to the 24 hour sampling time point following introduction of the disease elicitors to the *G. gracilis* culture medium. Thus, elucidation of its molecular function of this protein will require further scientific investigation.

2.3.7.2 G. gracilis genes that failed to meet the criteria for significant differential expression

Several cDNA clones that were not deemed significantly differentially expressed after 24 hours of exposure to disease elicitors according to the selection criteria (*P*-value < 0.05 and/or ($\mu - 1.96(\sigma) > M\text{-value} > (\mu + 1.96(\sigma))$), still exhibited altered transcription levels. This suggested that they might have been biologically relevant in the activated defence response of *G. gracilis*. A few were sequenced but four in particular (Table 2.6; marked with *), displayed significant homologies to genes associated with activated defence in other organisms. Furthermore, in certain instances, transcriptional regulation of these genes was shown to occur within the first few hours of the activated defence responses (and not at 24 hours). As such, if the hypothesis that activated defence in macroalgae and higher eukaryotes shared some common traits, these *G. gracilis* genes could not be overlooked.

A cDNA clone with significant sequence homology to a putative serine protease-like protein with endonuclease activity was transcriptionally up-regulated (M = 0.35) in *G. gracilis* in response to the disease elicitor. Proteases are classified as exo-, endo-, oligo- and omega-peptidases, all of which are able to hydrolyze peptide bonds (Rawling and Barret, 2004). Exopeptidases act only near the ends of polypeptide chains whereas endopeptidases act within the polypeptide (Antao and Malcata, 2005). Proteases can be further divided into sub-classes, namely, serine proteases which possess a serine residue in

the active site; cysteine and aspartic proteases which possess a cysteine residue or aspartic acid in the active site, respectively; and metalloproteases which require a metal ion for activity. Serine and cysteine proteases form covalent enzyme/substrate complexes via nucleophilic attack on the carbonyl group of the peptide bond whereas aspartic and metalloproteases do not (Dunn, 2001). Serine proteases were previously thought to be rare in plants but recent evidence suggests the opposite. Plants such as barley (*Hordeum vulgare*) (Fontanini and Jones, 2002), maize (*Zea mays*) (Goodfellow *et al*., 1993), *Arabidopsis thaliana* (Hamilton *et al*., 2003) and several others (Antao and Malcata, 2005) contain serine proteases. Plant serine proteases have been implicated in physiological processes (Palma *et al*., 2002) which include microsporogenesis (Kobayashi *et al*., 2002), pathogen recognition (Schaller, 2004), stress responses and the hypersensitive response (Tornero *et al*., 1997, 1996), signal transduction (Yano *et al*., 1999), senescence (Zhoa *et al*., 2000; Uchikoba *et al*., 1998) and protein processing/degradation (Barnaby *et al*., 2004; Müntz *et al*., 2001). Therefore, in the context activated defence in *G. gracilis*, transcriptional regulation of a gene encoding a serine-protease like protein definitely required further characterization.

Another *G. gracilis* cDNA clone with significant homology to a phosphoserine phosphatase (PSP) was transcriptionally up-regulated (M = 0.23) after 24 hours of exposure to disease elicitors. PSP proteins are classified as hydrolases that act on phosphoric monoester bonds and belong to the family of haloacid dehalogenase (HAD) superfamily (Wang *et al*., 2001). PSP is responsible for the biosynthesis of D-serine through its ability to dephosphorylate O-phospho-L serine in the following mechanism: O-phospho-L serine + H_2O = D serine + phosphate (Wang *et al*., 2001). Enzymes in this superfamily contain three highly-conserved sequence motifs: motif I: DXDX[T/V][L/V]; motif II: [S/T]XX; and motif III: K-[G/S][D/S]XXX[D/N] (Wang *et al*., 2001). The first Asp in motif I is the residue that gets phosphorylated (Collet *et al*., 1998). Motif III is particularly conserved between PSPs and P-type ATPases (GDGXXD), whereas motif III of HAD adopts a different sequence (SSXXD) (Aravind *et al*., 1998). The *G. gracilis* PSP protein sequence was aligned to three other PSP sequences (*Homo sapiens*, accession number NP_004568; *Arabidopsis*

thaliana, accession number NP_973858; and *Methanohalophilus mahii*, accession number NC_014002) (Appendix C.3). The multiple sequence analysis was able to detect all three protein motifs within the *G. gracilis* PSP DNA sequence. It is well established that reversible phosphorylation of proteins influences several biological processes in eukaryotes via the opposing actions of protein kinases and protein phosphatases (Wang *et al*., 2002; Smith and Walker, 1996). Phosphorylation of eukaryotic proteins mainly occurs on serine and threonine residues and to a lesser extent on tyrosine residues (Shenolikar, 1994). The serine/threonine phosphatases (PPases) specifically catalyze dephosphorylation of phosphoserine and phosphothreonine in protein substrates. PPases are classified into four sub-groups, namely, PP1, PP2A, PlP2B and PP2C (Cohen, 1989). Most PPases are composed of more than one catalytic subunit and share high sequence similarity (except for PP2C) but they are very different from tyrosine phosphatases (Luan, 1998). To date, almost all protein phosphatases are ubiquitous components of the signalling pathways triggered by the abiotic and biotic stresses that animal, yeast and plant cells may encounter (Xu *et al*., 1998; Smith and Walker, 1996). In particular, reversible phosphorylation is essential for activation and inactivation of mitogen-activated protein kinases (MAPKs) in the early steps of these pathways (Schweighofer *et al*., 2004; Mizoguchi *et al*., 1997). Through the use of specific kinase inhibitors, the authors demonstrated the role for kinases in eliciting a response to the pathogen. Furthermore, specific phosphatase inhibitors of PP1 and/or PP2A prevented the oxidative burst from being constitutively active under normal conditions (Chandra and Low, 1995; Suzuki and Shinshi, 1995; Mackintosh *et al*., 1994). It is therefore possible that PSP plays an important role in the defence signalling pathways of *G. gracilis*.

A cDNA clone that was transcriptionally down-regulated (M = –0.63) after 24 hours of exposure to disease elicitors, displayed significant sequence homology to a bacterial phosphotidylserine decarboxylase (PSD), as well as to putative trans-membrane proteins and a protein of unknown function DUF1254. PSD is an integral trans-membrane protein that catalyzes the formation of phosphatidylethanolamine (PtdEtn) from phosphatidylserine (PtdSer). It specifically decarboxylates a lipid-linked form of the serine moiety and is

essential in phospolipid metabolism of prokaryotic and eukaryotic organisms (Voelker, 1997). In prokaryotic organisms, PSD is present in the cytoplasmic membrane while in eukaryotic organisms, two forms of the enzymes exist, namely PSD1 (associated with the inner mitochondrial membrane) (Zinser *et al*., 1991) and PSD2 (associated with Golgi and vacuolar compartments) (Trotter and Voelker, 1995). Phospholipid-derived molecules appear to be more than just structural components of membranes. Scientific evidence has linked them to co-factors for membrane-enzymes, signal pre-cursors and even signalling molecules in plant defence (Laxalt and Munnik, 2002). Phospholipids generated in specific membrane domains may serve as docking sites for cytosolic proteins. Upon binding, the target protein may be directly activated through a conformational change or indirectly activated by being placed in proximity to various kinases (Laxalt and Munnik, 2002). Alternatively, some phospholipids may alter membrane properties to promote vesicle formation, membrane recycling or secretion (Burger, 2000). Transcriptional regulation of PSD therefore suggests an important role in the defence response of *G. gracilis*.

A cDNA clone which displayed significant sequence homology to alanine-glyoxylate aminotransferase (AGT) was transcriptionally down-regulated (M = – 0.56) in response to exposure to disease elicitors after 24 hours. Leaf peroxisomes are thought to contain at least four different types of aminotransferases constituted by serine, glutamic acid, alanine and asparagine (Rehfield and Tolbert, 1972). Although the plant alanine aminotransferase isoforms also occur in various sub-cellular compartments such as mitochondria and the cytoplasm, alanine-glyoxylate with aminotransferase activity is primarily located in the peroxisomes (Liepman and Olsen, 2003). Peroxisomes are ubiquitous eukaryotic organelles involved in various oxidative reactions and in the metabolism of fatty acids and other metabolites (Gabaldón *et al*., 2006). Moreover, Otter *et al*. (1992) and Penther (1991) proposed that constitutive expression of AGT is required for general amino acid metabolism. Other evidence implicates light-dependent expression of AGT with potential roles in photorespiration and abiotic stress protection (drought, high light intensities and hypersalinity) (Seki *et al*., 2002; Wingler *et al*., 2000; Son and Sugiyama, 1992; Son *et al*., 1991). Taler *et al*. (2004) proved that glyoxylate aminotransferase plays a direct role in

plant defence. Transgenic melon over-expressing the genes *At1* and *At2* (similar to a serine-glyoxylate aminotransferase and an *A. thaliana* alanine-glyoxylate aminotransferase, respectively) displayed enhanced activity of glyoxylate aminotransferases and significant resistance against pathogen (downy mildew) infection. Song *et al*. (2004) recently identified two other aminotransferases (AGD2 and ALD1) in *A. thaliana* that appeared to be required for plant development and disease resistance. AGD2 mutants exhibited an increased resistance to bacterial infection and growth defects whereas ALD1 mutants exhibited enhanced susceptibility to infection. The authors concluded that AGD2 catalyzed the synthesis of one or more amino-acids (most likely lysine) or an unknown development regulator / signal molecule produced through lysine synthesis. Alternatively, ALD1 was necessary to generate salicylic acid, a widely accepted defence signalling molecule (Audenaert *et al*., 2002; Delaney *et al*., 1994; Raskin, 1992). Since the *G. gracilis* alanine-glyoxylate aminotransferase identified in this study was transcriptionally repressed after 24 hours, its role in activated defence remains unclear. Alternatively, the 24 hour time-point could have been too late to detect significant induction of expression of this gene. Therefore, further investigation of transcriptional regulation of this gene within the first few hours of exposure to disease elicitors would clarify its role in the activated defence response of *G. gracilis*.

2.4 CONCLUSION

This is the first study, to our knowledge, that has used cDNA microarray technology to elucidate the transcriptional regulation of genes involved in the defence response (biotic stress) of the red macroalga *G. gracilis*. Although mRNA is not the end product of a gene, transcriptional regulation of a gene provides a good starting point for further investigation when genomic data is lacking for the organism being studied (Ptashne and Gann, 2002; Brazma and Vilo, 2000). The results of this study revealed that after 24 hours of exposure to disease elicitors, several genes were differentially expressed in *G. gracilis*. However, a large proportion of the genes sequenced (53%) could not be functionally annotated. This observation highlighted the existence of genes unique to red macroalgae as well as the need to improve the available genetic resources to reveal how they cope with biotic stresses. Due

to the inherent systematic technical variability associated with the process of manufacturing and subsequent analysis of microarrays, the potential to overshadow the biological variation that the tool intends to measure was a factor (Rockett and Hellmann, 2004). This was especially evident when the Cy3 (green) fluorescent dye was used to label RNA targets for hybridization to the cDNA probes. Normalization of the data removed these unwarranted influences and allowed comparison of gene expression data across three separate biological repeats. Nevertheless, before the differentially expressed genes identified in this study can be accepted with certainty, gene expression data has to be validated independently of the microarray experiment. Corroborating microarray data usually takes the form of an alternative method of quantitating mRNA abundance, such as qPCR analysis. In addition, other microarray studies in red macroalgae (*Collén et al*., 2007; Collén *et al*., 2006) suggests that the most marked changes in gene expression in response to a biotic stress occur within the first 6 hours of application of the stress. Therefore, validation of gene expression data for a select number of genes will be discussed in Chapter 3. Genes selected for validation will be based on scientific evidence clearly linking them to biotic stress responses in other organisms. In addition, transcriptional regulation between 0 and 24 hours of exposure to disease elicitors will be assessed to characterize transcriptional regulation of the selected genes more comprehensively.

CHAPTER 3

Validation of microarray experiment and investigation of transcriptional regulation using real-time PCR

3.1 INTRODUCTION

Analysis of gene expression patterns using cDNA microarrays enabled the identification of transcriptionally regulated *G. gracilis* genes upon exposure to disease elicitors. The general consensus for identification of differentially expressed genes was previously based on a fold change greater than two-fold induction or repression of the mRNA transcript. An opposing school of thought contradicted this notion by arguing that a gene could be considered differentially expressed if a change in its expression level (including less than two-fold differences) led to a biologically significant effect in the cell (McCarthy and Smyth, 2009). An alternative approach for identifying differential gene expression was based on statistical testing which was independent of fold-change but subject to the type I error or false positives. Bearing in mind the limitations of these approaches to identifying differentially expressed genes, the current microarray study employed a combination of both. However, as outlined in Chapter 2, interpretation of the microarray data may have been influenced by each of the steps prior to data analysis. For example, it was possible for the complex cDNA probe preparations to cross-hybridize non-specifically to closely related sequences as a result of repetitive elements, poly(A)-tails, common motifs or other unknown sequences in the DNA (Chuaqui *et al*., 2002). In addition, low-intensity hybridization signals, which are typically challenging to interpret, could have influenced the observed microarray results (Rajeevan *et al*., 2001a). Because of the problems associated with interpretation of microarray data, transcriptional regulation of the putative defence genes identified in the microarray experiment (Chapter 2) had to be validated using an alternative molecular technique. An inherent challenge associated with attempts to validate microarray data though, is the identification of an appropriate molecular tool which is both reliable and sufficiently sensitive to detect a wide range of changes in gene expression while bearing in mind the limitations of the validation method chosen. In this regard, several molecular tools have been previously used but each has their limitations. For instance, northern blot hybridizations and RNAse

protection assays are time-consuming and require large amounts (~ 5 µg) of RNA (Rajeevan *et al*., 2001b). When RNA is a limiting factor, conventional semi-quantitative reverse transcription-PCR may overcome this problem but quantitation remains difficult and relies on end-point analysis of the PCR product (Siebert, 1997; Wittwer *et al*., 1997; Higuchi *et al*., 1993). On the other hand, real-time PCR (qPCR) is relatively quick, inexpensive and only requires pico- to nanogram amounts of RNA (Chuaqui *et al*., 2002). Various fluorescence-based chemistries are compatible with qPCR and the technique is sensitive enough to quantitate the accumulation of specific mRNAs via analysis of the efficiency and rate of each PCR reaction (Walker, 2002; Wittwer *et al*., 1997; Higuchi *et al*., 1993). Disadvantages of qPCR include the significant amount of time and effort required to optimize amplification parameters; the absolute necessity for high quality RNA which may otherwise affect cDNA synthesis; and the influence of residual genomic DNA in RNA preparations which may inflate the measured gene expression levels. Real-time PCR is therefore a combination of three distinct steps: (i) the reverse transcriptase-dependent conversion of RNA into cDNA, (ii) amplification of the cDNA using PCR and (iii) detection and quantitation of amplification products in real-time (Gibson *et al*., 1996). Individual qPCR reactions are characterized based on the point at which fluorescence first rises above a defined threshold background fluorescence cycle (Ct). The more target present in the starting cDNA material, the lower the Ct value (Nolan *et al*., 2006). The measured gene expression data requires normalization in order to minimize the influence of the inherent limitations of the qPCR molecular technique. Typically, internal controls or reference genes (RGs) whose expression remains constant across all the RNA samples assessed are suitable candidates for data normalization. For comprehensive reviews on qPCR, see Nolan *et al*. (2006), Wong and Medrano (2005), Bustin and Nolan (2004) and Bustin (2002).

In this chapter, the use of qPCR to independently validate the gene expression data observed in the microarray experiment (Chapter 2) is described. Five separate genes which were differentially expressed at 24 hours post-exposure to disease elicitors were selected for validation, i.e. thioredoxin, peroxiredoxin-like 2 family, phosphoserine phosphatase, putative serine protease and phosphotidylserine decarboxylase. The specific selection of these five *G. gracilis* genes was based on published scientific evidence that implicated them in activated

defence mechanisms in other organisms. In this current qPCR study, experiments were designed to assess gene expression over a pre-defined time-course, i.e. during the first 30 minutes as well as during the first 24 hours of exposure to disease elicitors. Finally, the approach used to identify appropriate reference genes for data normalization will also be outlined.

3.2 MATERIALS AND METHODS

All media and solutions used in this study are listed in Appendix A.

All DNA oligonucleotide primers were synthesized in the Department of Molecular and Cell Biology, University of Cape Town.

3.2.1 Seaweed sample acquisition for exposure to disease elicitors

G. gracilis samples were sourced from the Jacobsbaai Sea Products abalone farm, Cape Town, South Africa. Deionized water was used to remove any sandy or slimy sediment from the thallus surface while any visible epiphytes were detached by hand. The seaweed was maintained in a 20 L plastic tank supplied with seawater. Water movement was achieved by pumping compressed air into the tank through a plastic airline. Temperatures were maintained at 18 – 20°C, and illumination was provided by cool white fluorescent tubes at 7400 Lux with a 16:8 light-dark cycle.

G. gracilis thalli weighing 1.0 ± 0.05 g fresh weight were transferred to nine 0.5 L Erlenmeyer flasks containing 250 ml artificial seawater (ASW) enriched with 1/3 strength PES (0.6% v/v) (Appendix A.1). All seaweed samples were acclimatized in a temperature controlled growth room for three days under a 16:8 light-dark cycle at 18°C. Water movement and aeration was achieved by pumping compressed air into each flask through plastic airlines. After this acclimatization period, all growth media was discarded and flasks containing seaweed were grouped as follows: one flask was assigned ‘time 0 control’, four flasks were assigned ‘experiment at 1 hour, 8 hours, 12 hours and 24 hours’ respectively and four flasks were assigned ‘control at 1 hour, 8 hours, 12 hours and 24 hours’ respectively. Two hundred and fifty millilitres of ASW supplemented with 1/3 strength PES (0.6% v/v) was added to each of the control flasks while, 250 ml of ASW containing disease elicitors (see Chapter 2; Section 2.2.2 for protocol) supplemented with 1/3 strength PES (0.6% v/v) was added to each of the experimental flasks. *G. gracilis* in the ‘time 0 control’ flask was

immediately removed, flash-frozen in liquid nitrogen and stored in 0.1 g amounts at –70 °C until RNA extraction. Seaweed samples in the remaining experimental and control flasks were incubated for their corresponding time period under a 16:8 light-dark cycle at 18°C, subsequently flash frozen in liquid nitrogen and stored in 0.1g amounts at –70 °C until RNA extraction. In the context of this study, a biological repeat was defined as a set of 9 flasks containing *G. gracilis* (1 time 0 control + 4 experiment + 4 control). In total, three biological repeats (BRs) were prepared for each qPCR experiment.

In addition to the microarray validation qPCR experiments, the study by Weinberger *et al*. (1999) prompted the incorporation of a shorter time-course qPCR study to assess transcriptional regulation of the five putative defence genes. Because exposure to agar and agar digestion products resulted in the release of reactive oxygen species in *G. conferta* and other *Gracilaria* species as well as the release of hydrogen peroxide into the growth media as early as 3 min post exposure to disease elicitors, a 30 min time-course experiment assessed gene expression at 0, 5, 15 and 30 min post exposure to disease elicitors. The experimental design replicated the protocol outlined above except that sampling time-points were altered accordingly.

3.2.2 RNA extraction

3.2.2.1 Twenty four hour time-course experiment

RNA was extracted from the experimental and control *G. gracilis* samples in order to synthesize cDNA from each seaweed sample over the 24 hour time-course. All solutions were prepared with 0.1% Diethylpyrocarbonate (DEPC) water (Appendix A.3.2.1) and autoclaved for 20 min at 121°C prior to use. All glassware and plasticware was treated with chloroform, followed by 100% ethanol, a wash in DEPC-water to remove RNAses and finally autoclaved at 121°C for 40 min prior to use.

Total RNA required for the microarray experiment had been extracted according to the protocol outlined by Azvedo *et al*. (2003) (Section 2.2.3). Ideally, subsequent validation experiments should have utilized the same RNA samples but unfortunately, these were completely exhausted in the microarray experiment. Surprisingly, sufficient amounts of high quality RNA could not be isolated from the new seaweed samples (Section 3.2.1) using the protocol of Azvedo *et al*. (2003). It was thought that the disease elicitors (agar oligosaccharides) were co-precipitating with the RNA, resulting in decreased RNA yields. As an alternative, the peqGOLD Plant RNA Kit (Optima Scientific) was tested and proved successful. Total RNA was isolated from 0.1 g of each of the *G. gracilis* samples in the 24 hour time-course experiment according to the manufacturer's instructions and re-suspended in 50 µl DEPC-H_2O. Genomic DNA (gDNA) was eliminated via an on-column DNAse-treatment step incorporated in the kit's RNA isolation protocol.

3.2.2.2 Thirty minute time-course experiment

One disadvantage of the peqGOLD Plant RNA isolation kit was its high cost. Consequently, an alternative method was required to isolate RNA from the seaweed samples in the 30 min time-course experiment. A modified Trizol (Invitrogen, Life Technologies) RNA isolation protocol was used in combination with the buffers remaining from the peqGOLD kit. Seaweed samples (0.1 g) taken at each time-point were homogenized in 600 µl extraction buffer (Buffer RPL, peqGOLD kit). The entire sample was then loaded onto a QIAshredder column (Qiagen Ltd.) as per the manufacturer's instructions to remove cellular debris. To the RNA-containing supernatant, 140 µl of Buffer SP (peqGOLD kit) was added to remove contaminating polysaccharides by precipitation and centrifugation at 12 000 rpm for 5 min. The supernatant was then removed and aliquoted into a sterile 1.5 ml microfuge tube. RNA was then extracted by adding Trizol reagent as per the manufacturer's instructions. Following extraction, total RNA samples were purified using butanol and diethylether according to the protocol outlined by Krebs *et al*. (2009) and re-suspended in 50 µl DEPC-H_2O.

3.2.2.3 Assessment of DNAse efficacy and presence of contaminating gDNA

In order to ensure that all gDNA was eliminated from the RNA samples prior to cDNA synthesis, the efficacy of the on-column DNAse-treatment using the peqGOLD Plant RNA isolation kit was assessed. This step was required due to the possibility that *G. gracilis* genes examined in this qPCR study could be amplified from residual gDNA using the synthesized gene-specific qPCR primers, leading to an exaggerated measure of gene expression. Therefore, a PCR reaction was set up to assess whether the high copy number 18S rRNA gene could be amplified using total RNA (DNAse-treated) as template. Total RNA (450 ng) was added to a PCR reaction mix containing 2.5 µl of 10X reaction buffer, 1.5 µl 25 mM $MgCl_2$, 0.5 µl each of 18S rRNA forward and reverse primers (Appendix B.1.3) and 1 U of *taq* polymerase (Super-Therm polymerase, Southern Cross Biotechnology) made to a final volume of 25 µl with nuclease-free H_2O. PCR reactions were performed in a Bioer XP thermal cycler (Separation Scientific) according to predefined cyclic conditions: initial 3 min denaturation followed by 35 cycles of 30 s denaturation at 94°C, 30 s annealing at 59°C and 1 min extension at 72°C. The reactions were terminated with an additional incubation of 7 min at 72°C and a cooling step at 4°C for 10 min. A positive control consisted of 10 ng of the 18S rRNA gene cloned into a plasmid while nuclease-free H_2O was added to the no template control (NTC). Entire PCR reactions were electrophoresed on a 1 % TAE (Appendix A.2) agarose gel. If the high-copy number 18S ribosomal RNA gene (expected size ~ 1.8 kb) was not amplified by PCR, it could be deduced that residual gDNA possibly remaining in the RNA preparation would not affect the measured expression levels of the gene of interest since they would have a lower copy number in comparison to the highly abundant 18S rRNA gene.

Despite the manufacturer's claim that total RNA isolated by Trizol should be completely free of residual gDNA, the presence of gDNA was still assessed prior to cDNA synthesis. Thus, PCR amplification of the high-copy number 18S rRNA gene was assessed using 450 ng of total RNA as template. One modification was made: instead of using the 18S rRNA primer set, a different forward primer (5' ATATGCGAAA GCATTTCCCAATCTC 3') was used which would yield an amplicon of 277 bp. PCR reactions were performed in a Bioer XP thermal cycler (Separation Scientific) according to predefined cyclic conditions: initial 3 min

denaturation followed by 35 cycles of 30 s denaturation at 94°C, 30 s annealing at 59°C and 1 min extension at 72°C. The reactions were terminated with an additional incubation of 7 min at 72°C and a cooling step at 4°C for 10 min. The positive control consisted of 10 ng of the 18S rRNA gene cloned into a plasmid while nuclease-free H_2O was added to the no template controls (NTC). Entire PCR reactions were electrophoresed on a 1 % TAE (Appendix A.2) agarose gel. Similarly, if the high-copy number 18S ribosomal RNA gene was not amplified by PCR, it could be deduced that residual gDNA possibly remaining in the RNA preparation would not affect the measured expression levels of the gene of interest since they would have a lower copy number in comparison to the highly abundant 18S rRNA gene.

3.2.3 Quantitative and qualitative analysis of RNA

Quantitative analysis of RNA was performed using a Nanodrop spectrophotometer (Thermo Scientific). Absorbance readings were measured at 260 and 280 nm, with one unit of absorption at 260 nm representing 40 µg ml^{-1} of RNA. Ratio measurements at 230, 260 and 280 nm were used to assess the purity of the RNA samples. RNA integrity was determined by assessing a 1 µg sample on a 1.2% formaldehyde-agarose gel (Appendix A.3.2.5). Formaldehyde-agarose gel electrophoresis was performed in 1 X MOPS (Appendix A.3.2.6) at 70V as described by Sambrook *et al*. (1989). RNA was subsequently visualized on a 254 nm UV-transilluminator.

3.2.4 Conversion of total RNA to complementary DNA (cDNA synthesis)

Complementary DNA was synthesized using the ImProm-II Reverse Transcription System (Promega) according to the manufacturer's instructions. For each time-point, 3 µg of total RNA was incubated with $Oligo(dT)_{15}$ (1.0 µg) for 5 min at 72°C in heating block. While cooling on ice, a master mix was prepared to constitute a final concentration of 1X reaction buffer, 3 mM $MgCl_2$, 0.25 mM dNTPs, 2.5 U RNAse inhibitor, 2 U reverse transcriptase and nuclease-free water to a final volume of 40 µl. The reverse transcription was continued by heating the reaction mixture to 25°C for 5 min, then 42 °C overnight (~ 14 hours). Reverse transcriptase was inactivated by incubating the reaction mixture at 70°C for 15 min. A no

reverse transcriptase control (No RT) was set up as above with the exclusion of reverse transcriptase. For each RNA sample, cDNA synthesis was performed in duplicate, subsequently pooled and stored at –70 °C until further use.

3.2.5 Designing of qPCR primers

The DNA sequences (Appendix E.1) of five putative defence genes (peroxiredoxin-like or PRX2F, phosphoserine phosphatase, putative serine-protease, phosphotidylserine decarboxylase and thioredoxin) were used to design primers. Open-source software, FastPCR (http://primerdigital.com/fastpcr.html), was used to design all gene-specific qPCR primers. Since the method of cDNA synthesis utilized the Oligo$(dT)_{15}$ primer, it was necessary to design primers toward the 3' ends of the genes of interest (Table 3.1). All primers were designed to have annealing temperatures of approximately 60°C.

Table 3.1 Gene specific qPCR primers for five putative defence genes and the expected amplicon sizes.

Gene Name (Clone ID)	Forward Primer (5'-3')	Reverse Primer (5'-3')	Size (bp)
PRX2F (233)	GCTGATCTGGACTACACGAAGA	TACGCCGACGCTACGATA	116
PSP (991)	AAAGCCACGATGCAGGTT	TCACCCTTAGAGTCTCACAA	112
PSD (897)	GATTGTCGAACGTATTGTGTAGG	CCTTGTCATGCCAAACACA	120
Serine protease-like (428)	GCATTTCCCTGCTTGAAAGG	AACTGCGGCGTCAATCTG	134
Thioredoxin (4)	GAGGACTGGCTGGTGAACTT	CCGTTCACGAAGATGATGA	136

3.2.6 Identification of reference genes for normalization of qPCR data

In order to identify genes whose expression did not change under either experimental or control conditions, genes with M-values (log_2 fold change) equal or close to zero were selected from the microarray experiment. Seven such genes met these selection criteria (clone identities: 245, 329, 363, 376, 408, 460 and 801). Glycerol stocks of *E. coli* XL1 Blue transformants (Section 2.2.1.1) with plasmids harbouring cDNA fragments corresponding to possible reference genes were grown at 37°C on Luria agar supplemented with 30 $\mu g.ml^{-1}$ chloramphenicol. Single colonies were used to inoculate 5 ml Luria broth (Appendix A.1.2) supplemented with 30 $\mu g.ml^{-1}$ Cm. Bacterial cultures were incubated overnight at 37^{o}C on a shaker at 100 rpm. Plasmid DNA was extracted according to Sambrook *et al*. (1989) (Appendix C.1). Plasmid DNA was re-suspended in 50 µl sterile distilled water and quantitated with a Nanodrop spectrophotometer. Additionally, DNA samples were electrophoresed on a 1.2% TAE agarose gel to verify plasmid quality and integrity. All cDNA inserts were PCR amplified using M13 universal primers (Appendix B.1.1) using standard cycling parameters (Section 2.2.1.1). PCR products were subsequently electrophoresed on a 1.2% TAE agarose gel to verify amplification of a single insert prior to DNA sequencing. Plasmids that met this quality criterion (clones 245, 329 and 801) (data not shown) were purified with the E.Z.N.A® Cycle-Pure Kit (peqLab Biotechnologie GmbH) according to the manufacturer's instructions. Plasmid DNA was stored as 100 $ng.\mu l^{-1}$ stocks and sent for DNA sequencing (Macrogen Inc., South Korea). DNA sequence data (Appendix E.2) were subjected to BLASTx and BLASTn algorithms (http://www.ncbi.nlm.nih.gov/Blast.cgi) (Section 2.2.9) in order to functionally annotate the potential reference genes where possible. Reference gene-specific qPCR primers were designed to the 3' ends of the DNA sequences using FastPCR (Table 3.2). In addition, primers designed to amplify an internal region of the 18S rRNA gene were also generated based on the *G. gracilis* 18S rRNA gene sequence (Appendix E.3).

Table 3.2 Gene specific qPCR primers for four reference genes with expected amplicon sizes.

Gene Name & Clone ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	Size (bp)
Hypothetical protein (245)	CTGCGAGAGCGGTACCAAATCG	GCCTTAGCTGGTGGTCTGG	120
Sugar epimerase (329)	TTTCTCAGGCTATGAGACCGAGC	ACTATGACGGGCACGAACATCC	114
Hypothetical protein (801)	AAACAGCGGGATGCCAATACC	ACCGCTCGTGGACGCTACC	131
18s rRNA	ATATGCGAAAGCATTTCCCAATCTC	GGAGAAGTCGTAACAAGGTTTCCGTAG	277

3.2.7 Real-time PCR

Real-time PCR analyses of the 24 hour and 30 min time-course experiments were used to assess the transcriptional regulation of five putative defence genes (GOIs) and four reference genes (RGs). All qPCR reactions were performed using the SYBR green chemistry (SensiMix dT, Quantace) on a Rotor-Gene 6000 system (Corbett Lifescience). The following PCR cycling parameters were optimized as follows: an initial enzyme activation step at 95 °C for 10 min followed by 40 cycles of 95°C for 10 sec, 58 – 60°C for 15 sec and 72°C for 20 sec. Fluorescence readings were acquired after the extension step. The optimal annealing temperature for GOIs 4, 233 and 428 was 59°C while 897 and 991 required a higher annealing temperature of 60°C. The optimal annealing temperature for RGs 245 and 801 was 60°C, while the annealing temperature for 329 and the 18S rRNA primer set was 59°C and 58°C, respectively. Triplicate qPCR reactions were performed in a reaction volume of 12.5 µl, each containing 0.25 µl of 50 x SYBR Green solution, 6.25 µl of 2 x SensiMix and 0.25 µl of 10 µM gene specific primers. The relative amounts of cDNA for each GOI and RG during the two time-course experiments was determined from 1 µl of pooled cDNA (in separate qPCR reactions) from each time-point and treatment (experiment and control). Control qPCR reactions for each gene included a no template (NTC) as well as the aforementioned

No RT control. Plasmid DNA harbouring the cDNA insert for each gene was used to construct a standard curve where Ct values were plotted against the Log template amount. A ten-fold serial dilution series of each plasmid (1 ng down to 0.00001 ng) was constructed. The RotorGene software (Version 1.7; Corbett Research) automatically generated qPCR reaction efficiencies (E) as well as correlation efficiencies (R^2-value) for each primer set. Ct values for the time-course experiments for each GOI and RG were extrapolated from the corresponding standard curve. Reaction efficiencies were calculated using the equation: $E = 10^{(-1/slope)} - 1$. Ideally, the efficiency of any PCR reaction should be 100%, i.e. a doubling of the amplicon after each amplification cycle. However, most PCR reactions do not typically display ideal amplification efficiencies and as such, calculations without an appropriate correction factor may overestimate the starting mRNA concentration (Liu and Saint, 2002). As a result, reaction efficiencies between 90 – 110% (0.9 – 1.1) which correspond to a slope of –3.1 to –3.6 in the standard curves are generally considered optimal (Wong and Medrano, 2005). The specificity of each PCR amplification was assessed using dissociation (melt) curves in which PCR reactions were heated from 72 to 95°C, rising by 1°C between each step with a hold period of 5 s. In addition to technical replication, the expression of each gene at a given time-point under each treatment was assessed using three separate biological repeats. Gene expression data obtained from three biological replicates were averaged and subsequently used in data analysis and relative quantitation of mRNA transcripts.

3.2.8 Normalization and relative quantitation of gene expression

Before gene expression could be analysed, RGs that were stably expressed across each time-point and treatment (experiment and control) had to be identified. RGs were required to normalize variation in measured gene expression as a result of technical bias associated with the qPCR technique. Stable expression of RGs was assessed using the geNORM VBA applet for Microsoft Excel (Vandesompele *et al*., 2002; http://medgen.ugent.be/genorm/). The geNORM program calculated a gene expression stability measure (M) by comparing RG expression to all other RGs tested. RGs with the highest variation were excluded in a stepwise manner which enabled a ranking of RGs according to their expression stability. After data normalization, gene expression was analysed using a relative quantitation model based on Pfaffl (2001) (Figure 3.1). Two parameters for each GOI and RG, namely, qPCR

amplification efficiency and Ct value as calculated in the Rotor-Gene series 6000 software (Version 1.7), were required to establish the relative expression of each gene. A value of 1 was added to all calculated qPCR efficiencies as per the Pfaffl model. Relative quantitation was achieved by selecting time-point 0 as the calibrator sample. In other words, gene expression at each sample point in the two time-course experiments was represented as n-fold change relative to time 0.

$$\text{Ratio} = \frac{(E_{\text{target}})^{\Delta Ct \text{ target (control} - \text{sample)}}}{(E_{\text{reference}})^{\Delta Ct \text{ reference (control} - \text{sample)}}}$$

Figure 3.1 Equation representing the Pfaffl mathematical model for calculating relative expression ratios in qPCR. The ratio of the gene of interest (target) expressed in a sample versus a calibrator sample (control) in comparison to a reference gene (reference). E_{target} was the amplification efficiency of the target gene and $E_{reference}$ was the amplification efficiency of the reference gene. ΔCt target was the difference in Ct values of the target gene in the control and sample and ΔCt reference was the difference in Ct values of the reference gene in the control and sample. PCR amplification efficiencies (E) were calculated according to $E = 10^{[-1/slope]}$ (Pfaffl, 2001).

3.2.9 Statistical analysis and validation of microarray experiment

The normalized GOI expression levels (experiment and control) were expressed as an n-fold change relative to the GOI time 0 calibrator samples. The aim of this experiment was to determine whether the five putative defence genes were transcriptionally regulated as a consequence of exposure to disease elicitors. Since a control sample was included at each time point along the 24 hour time-course, a student's *t*-test was deemed appropriate to establish whether gene expression in the experimental samples changed significantly relative to the control samples. Data sets that were not normally distributed or that displayed unequal variances were transformed to meet the requirements for parametric statistical testing. Where data could not be adequately transformed, a non-parametric statistical test was used instead. In addition, all GOI expression values (n-fold changes) at 24 hours were $\log_2$ transformed and compared to the microarray results in order to validate their transcriptional regulation.

3.3 RESULTS

3.3.1 Extraction of total RNA

Two time-course experiments (24 hours and 30 min) were performed in order to characterize the transcriptional regulation of five putative defence genes in response to disease elicitors. *G. gracilis* samples at each sampling time-point and treatment (experiment and control) were collected and stored at –70°C until RNA was extracted. Total RNA preparations for both time-course experiments exhibited no signs of ribonuclease degradation and were deemed optimal for use in cDNA synthesis. As evident in Figures 3.2 – 3.3, the 28S and 18S ribosomal RNA bands were distinct and intact. In addition, several other distinct bands were detected in the total RNA preparations. Total RNA was re-extracted where ribonuclease degradation of the RNA was apparent (Figure 3.2, Panel B).

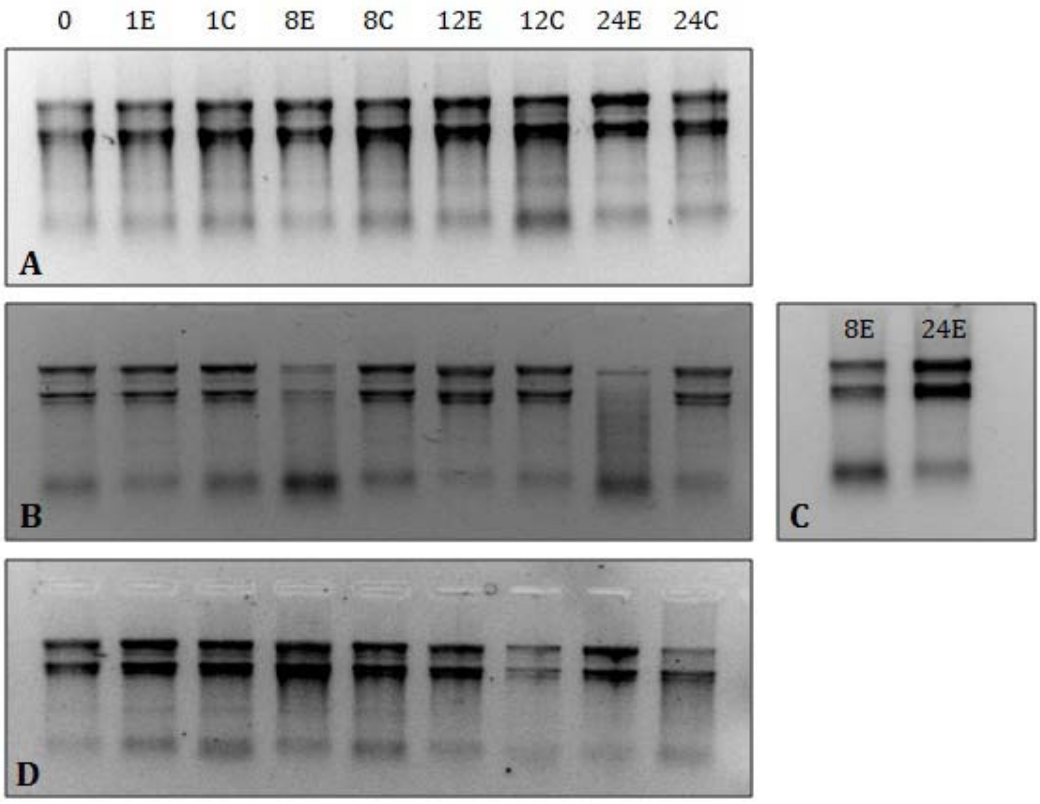


Figure 3.2 Assessment of RNA quality and integrity by electrophoresis of 1.2 µg total RNA through a 1.2 % formaldehyde-agarose gel. Time-course sampling points are represented by 0, 1, 8, 12 and 24 hours with experimental or control samples designated "E" or "C" respectively. Total RNA was isolated using the peqGOLD plant RNA isolation kit where (A) represents total RNA isolated from BR1, (B) represents total RNA isolated from BR2, (C) represents the RNA isolations that were repeated for samples 8E and 24E which initially exhibited signs of ribonuclease degradation and (D) represents total RNA isolated from BR3.

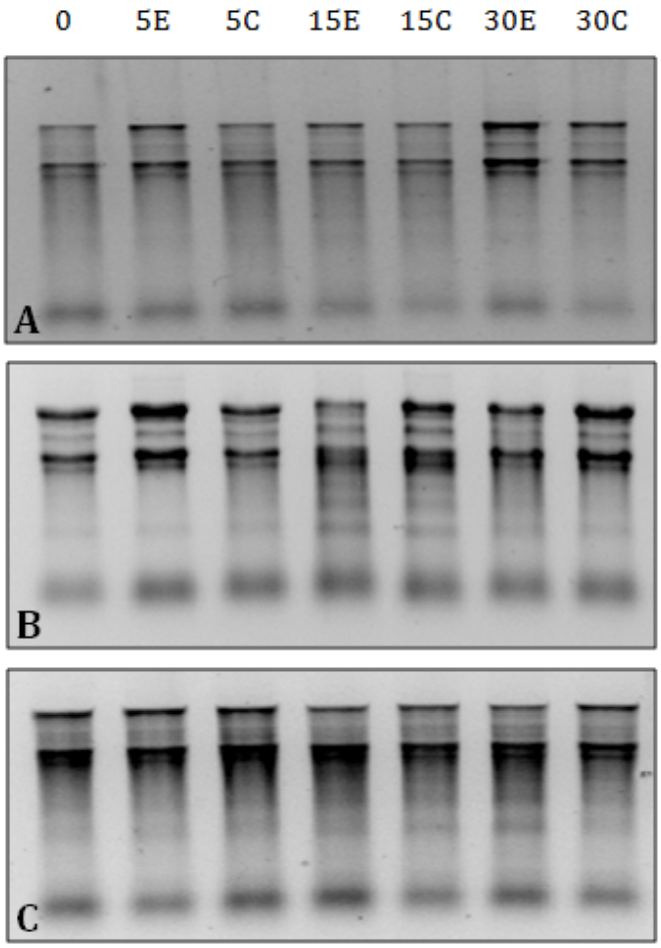


Figure 3.3 Assessment of RNA quality and integrity by electrophoresis of 1.2 µg total RNA through a 1.2 % formaldehyde-agarose gel. Time-course sampling points are represented by 0, 5, 15 and 30 min with experimental or control samples designated “E” or “C” respectively. In order to improve the purity of the total RNA preparations isolated using a modified Trizol reagent protocol, butanol and diethylether washes were incorporated into the protocol. Panel (A) represents total RNA isolated from BR1, (B) represents the total RNA isolated from BR2 and (C) represents the total RNA isolated from BR3.

3.3.2 Assessment of DNAse efficacy and genomic DNA contamination in total RNA preparations

Prior to cDNA synthesis, all total RNA preparations were assessed for residual gDNA contamination by PCR using 18S rRNA gene-specific primers. If residual gDNA was present in the RNA preparations, the 18S primers would have amplified the high copy number 18S rRNA gene. The on-column DNAse-treatment of RNA (24 hour experiment) proved effective in eliminating residual gDNA since the expected 1.8 kb amplicon was not PCR-amplified when total RNA was used as template (Figure 3.4A). A band of the expected molecular weight (1.8 kb) was amplified in the positive control which ruled out any problems associated with poor primer design. Total RNA preparations from the 30 min experiment, which were not DNAse-treated, were similarly tested using PCR to assess the levels of residual gDNA. Instead of using an 18S rRNA primer set designed to yield a 1.8 kb amplicon, a smaller amplicon of 277 bp was expected. As evident in Figure 3.4B, an amplicon of the expected size could not be detected when total RNA was used as template. However, a band of the expected molecular weight (277 bp) was amplified in the positive control which ruled out any problems associated with poor primer design. Non-specific products of larger molecular weight amplified in the positive control sample were attributed to excess plasmid DNA in the reaction and the high number of amplification cycles (35). It was therefore concluded that no residual gDNA contaminated total RNA preparations in either of the time-course experiments and that synthesis of cDNA could commence.

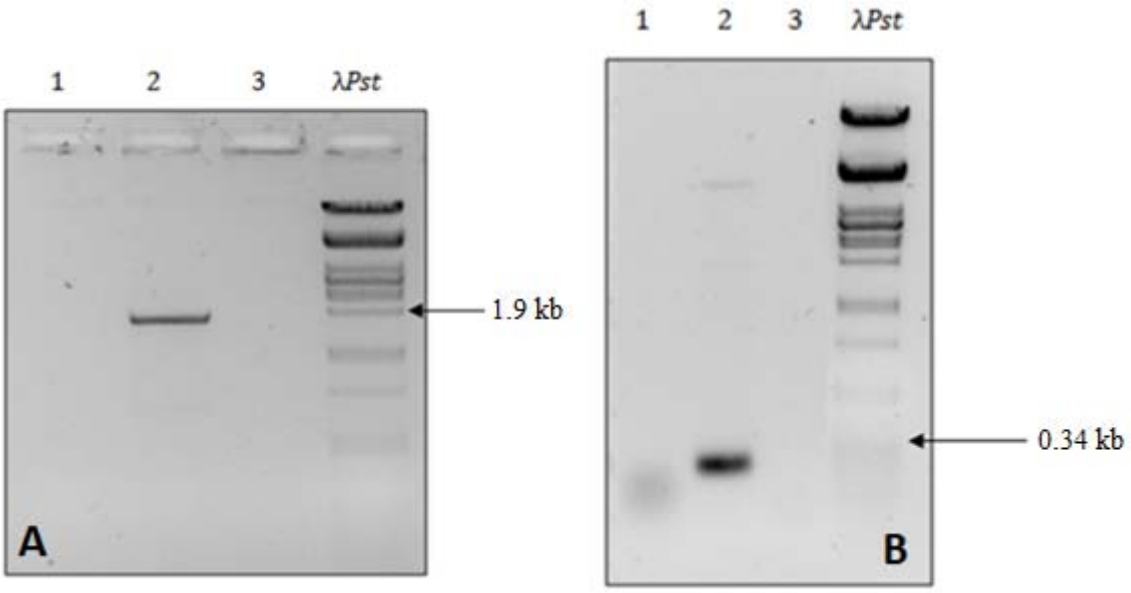
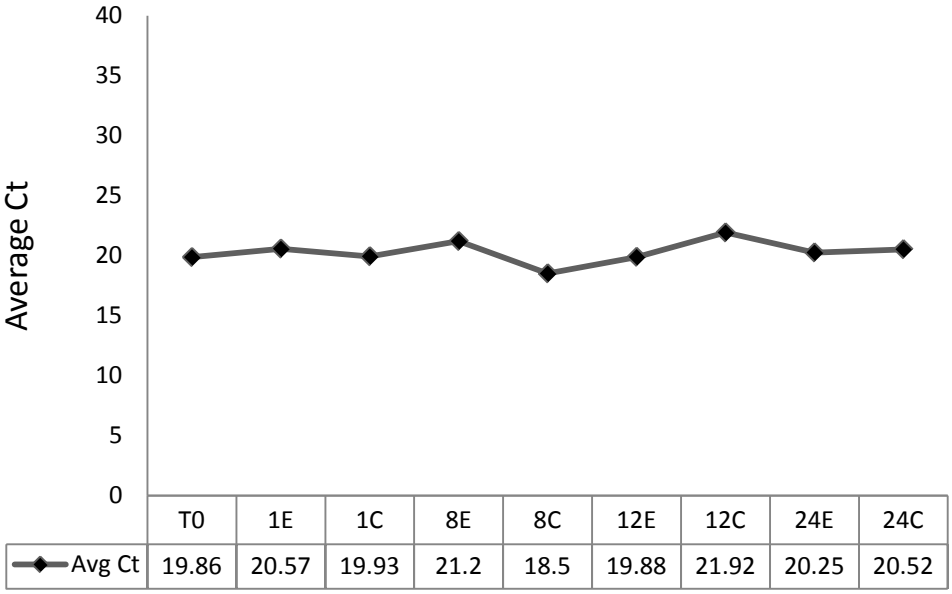


Figure 3.4 Agarose gel (1.2 %) depicting the PCR amplification of 18S rRNA gene product to test the amount of residual gDNA in the total RNA samples. Panel (A) represents the 18S rRNA PCR test for the 24 hour time-course experiment. Lane (1) represents amplification of 450 ng of DNAse-treated RNA, lane (2) represents the positive control which contained 10 ng of plasmid DNA and lane (3) represents the no template control (NTC). Panel (B) represents the 18S rRNA PCR test for the 30 min time-course experiment. Lane (1) represents the amplification off 450 ng of total RNA (not DNAse-treated), lane (2) represents the positive control which contained 10 ng of plasmid DNA and lane (3) represents the no template control (NTC). A λ*Pst* molecular weight marker was used to determine the size of amplification products. The whole PCR reaction volume (25 µl) was loaded into each lane of the gel.

3.3.3 Identification of reference genes for normalization of 24 hour qPCR data

In order to normalize the GOI expression data for the 24 hour time-course experiment, a stably expressed RG had to be identified. In the context of this study, a stable RG was defined as a gene whose expression (or Ct values) did not fluctuate at each of the time-points and under both treatments (experimental and control). Only one biological repeat was used to establish the most stable RG in order to minimize the expense of the experiment due to the high cost of the qPCR reagents. From the corresponding standard curve, PCR amplification efficiency and correlation efficiency for each RG was calculated automatically by the Rotor-

Gene series 6000 software (Version 1.7). The Ct values for a potential RG at each time-point and treatment were plotted to establish the extent of variation in gene expression (Figure 3.5 – 3.8). PCR amplification efficiencies for all RGs tested ranged between 0.8 and 1.2 except for the 18S rRNA gene which displayed an efficiency of 0.64. Consequently, the 18S rRNA gene was eliminated as a potential RG. Analysis of the Ct values for the three remaining RGs revealed that their expression was relatively stable since their Ct values varied by only two units across the time-course and treatments. Of the three remaining RGs, the geNORM Microsoft Excel applet proposed that RG 245 and 329 were the most stably expressed for the 24 hour time-course experiment. Since only one RG was required for data normalization using the Pfaffl relative quantitation model, RG 245 was selected as the RG for all three biological repeats.



	T0	1E	1C	8E	8C	12E	12C	24E	24C
Avg Ct	19.86	20.57	19.93	21.2	18.5	19.88	21.92	20.25	20.52

Figure 3.5 Graph representing the average Ct values for *RG 329* over the 24 hour time-course experiment for BR1. The x-axis represents the sampling time-point in hours while "E" and "C" denote experimental and control samples respectively (PCR amplification efficiency = 0.84; R^2 = 0.999).

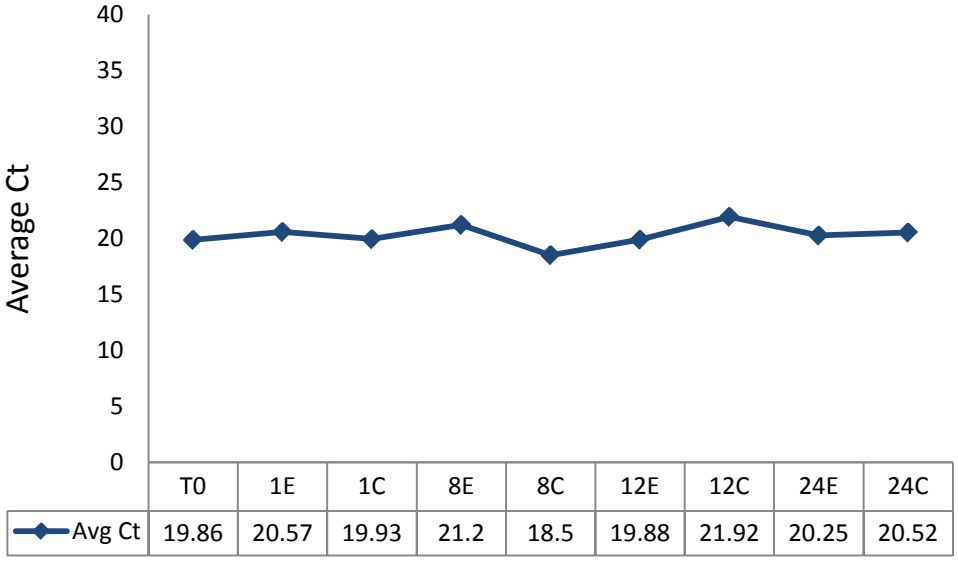


Figure 3.6 Graph representing the average Ct values for *RG 245* over the 24 hour time-course experiment for BR1. The x-axis represents the sampling time-point in hours while "E" and "C" denote experimental and control samples respectively (PCR amplification efficiency = 1.22; R^2 = 0.999).

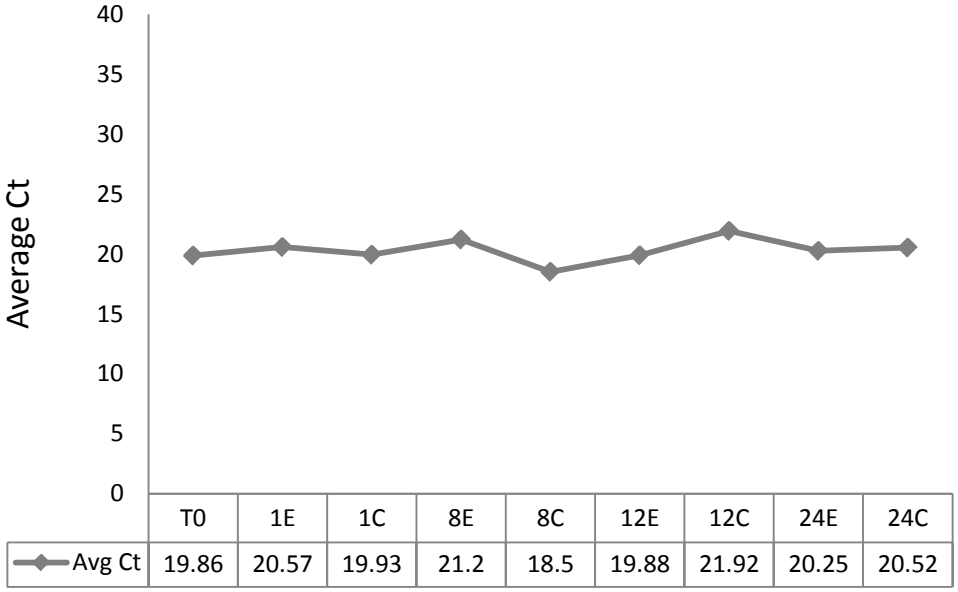


Figure 3.7 Graph representing the average Ct values for *RG 801* over the 24 hour time-course experiment for BR1. The x-axis represents the sampling time-point in hours while "E" and "C" denote experimental and control samples respectively (PCR amplification efficiency = 1.03; R^2 = 0.987).

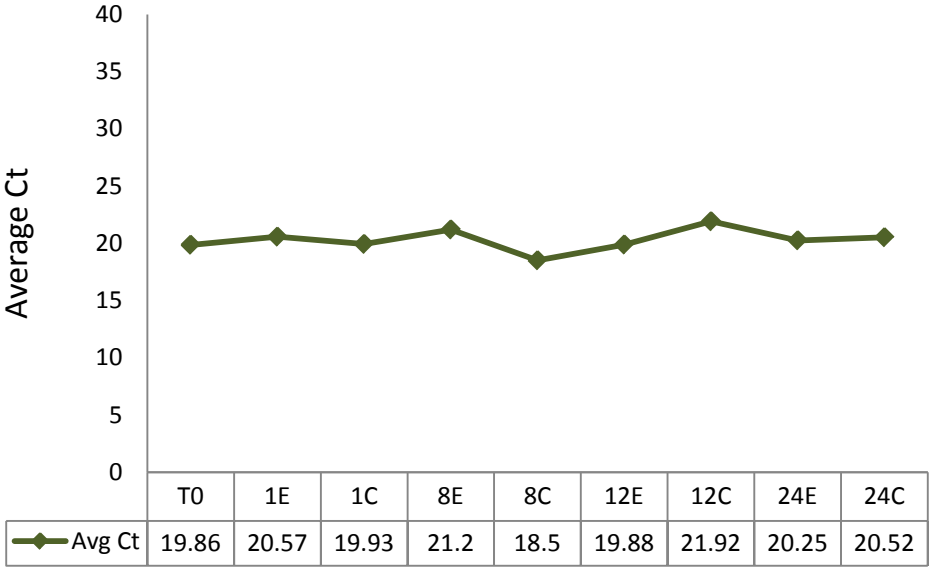


Figure 3.8 Graph representing the average Ct values for *18S rRNA RG* over the 24 hour time-course experiment for BR1. The x-axis represents the sampling time-point in hours while "E" and "C" denote experimental and control samples respectively (PCR amplification efficiency = 0.63; R^2 = 0.979).

3.3.4 Identification of reference gene for normalization of 30 min qPCR data

Reference genes 245 and 329 were identified as the two most stable RGs in the 24 hour time-course experiment. As a result, they were assessed as normalization controls for the 30 min time-course experiment. Similarly, only one biological repeat was used to establish the most stable RG in order to minimize the cost of the experiment. As for the 24 hour experiment, plasmid DNA was used to construct standard curves in order to calculate PCR amplification and correlation efficiencies. The Ct values for each time-point and treatment were plotted for the two RGs to establish the degree of variation in gene expression (Figure 3.9 – 3.10). The amplification efficiency for RG 245 and 329 was 1.2 and 0.8, respectively. Analysis of the Ct values for both RGs revealed that Ct values varied by only two units across the time-course and treatments with one exception, i.e. the relatively high Ct value observed in sample 30E for RG 245. Therefore, RG 329 was selected to normalize GOI expression data for all three biological repeats in the 30 min experiment.

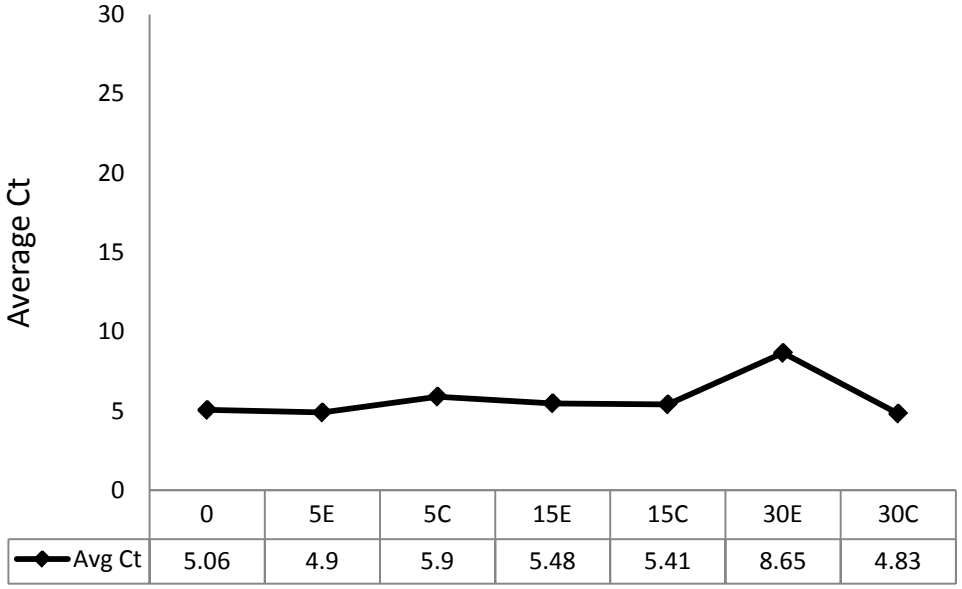


Figure 3.9 Graph representing the average Ct values for *RG 245* over the 30 min time-course experiment for BR1. The x-axis represents the sampling time-point in minutes while "E" and "C" denote experimental and control samples respectively (PCR amplification efficiency = 1.09; R^2 = 0.980).

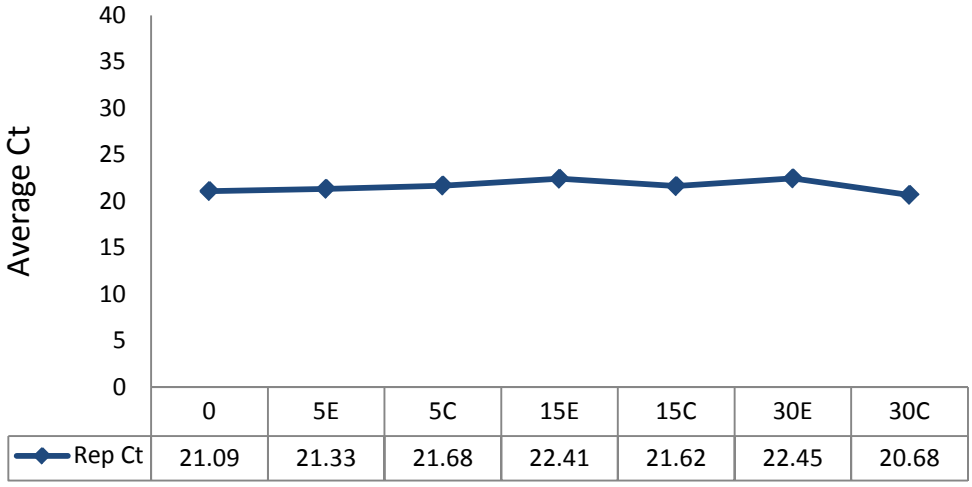
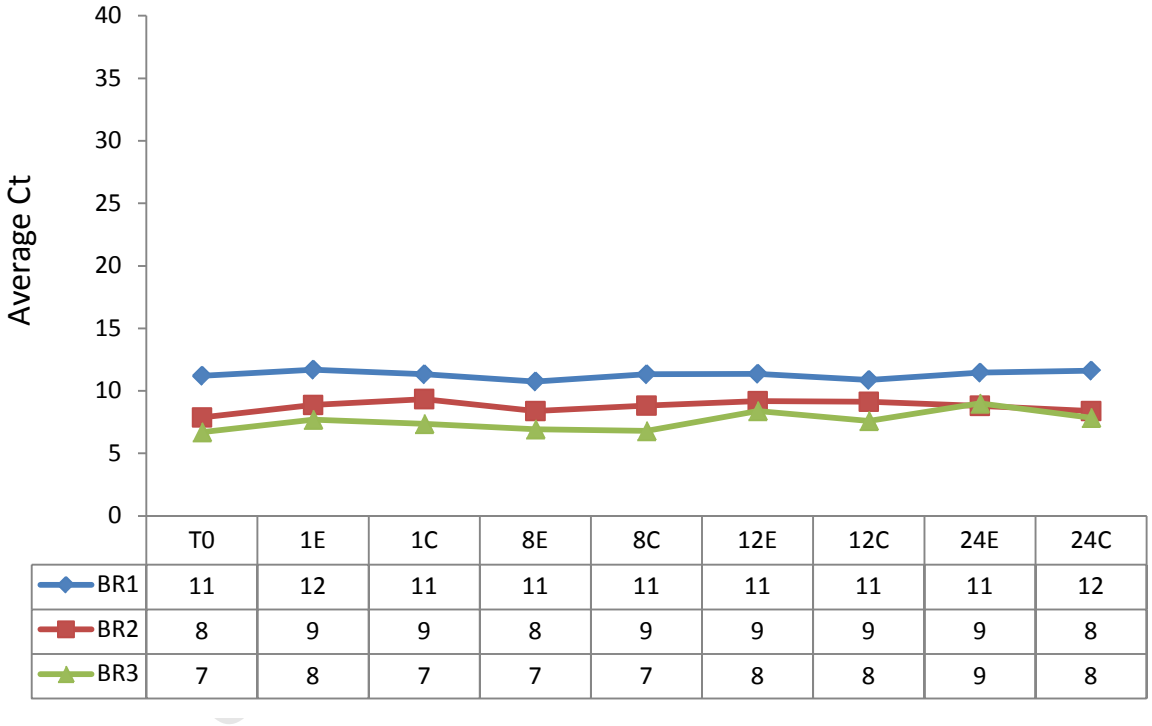


Figure 3.10 Graph representing the average Ct values for *RG 329* over the 30 min time-course experiment for BR1. The x-axis represents the sampling time-point in minutes while "E" and "C" denote experimental and control samples respectively (PCR amplification efficiency = 0.80; R^2 = 0.99).

3.3.5 Relative quantitation of putative defence genes during the 24 hour experiment

Before expression of the GOIs was analysed, RG 245 had to be confirmed as a stably expressed gene across all three biological repeats. As expected, RG 245 exhibited a gene expression pattern that did not fluctuate over the 24 hour time-course under experimental and control conditions. In addition, the Ct values for RG 245 were within two units of each other (Figure 3.11).



	T0	1E	1C	8E	8C	12E	12C	24E	24C
BR1	11	12	11	11	11	11	11	11	12
BR2	8	9	9	8	9	9	9	9	8
BR3	7	8	7	7	7	8	8	9	8

Figure 3.11 Graph representing the average Ct values of RG 245 expression in all three biological repeats over the 24 hour time-course experiment. The x-axis represents the sampling time-point in hours while "E" and "C" denote experimental and control samples respectively. The average Ct values for each biological repeat were determined automatically in the Rotor-Gene 6000 software from the respective RG 245 standard curves.

Dissociation (melt) curves were used to determine the PCR amplification specificity of each primer set tested (RG 245 and five GOIs). Each primer set was optimally designed with

respect to its amplification specificity. In each case, amplification of a single PCR amplicon product, represented by a single peak in each dissociation curve, was observed (Figure 3.12).

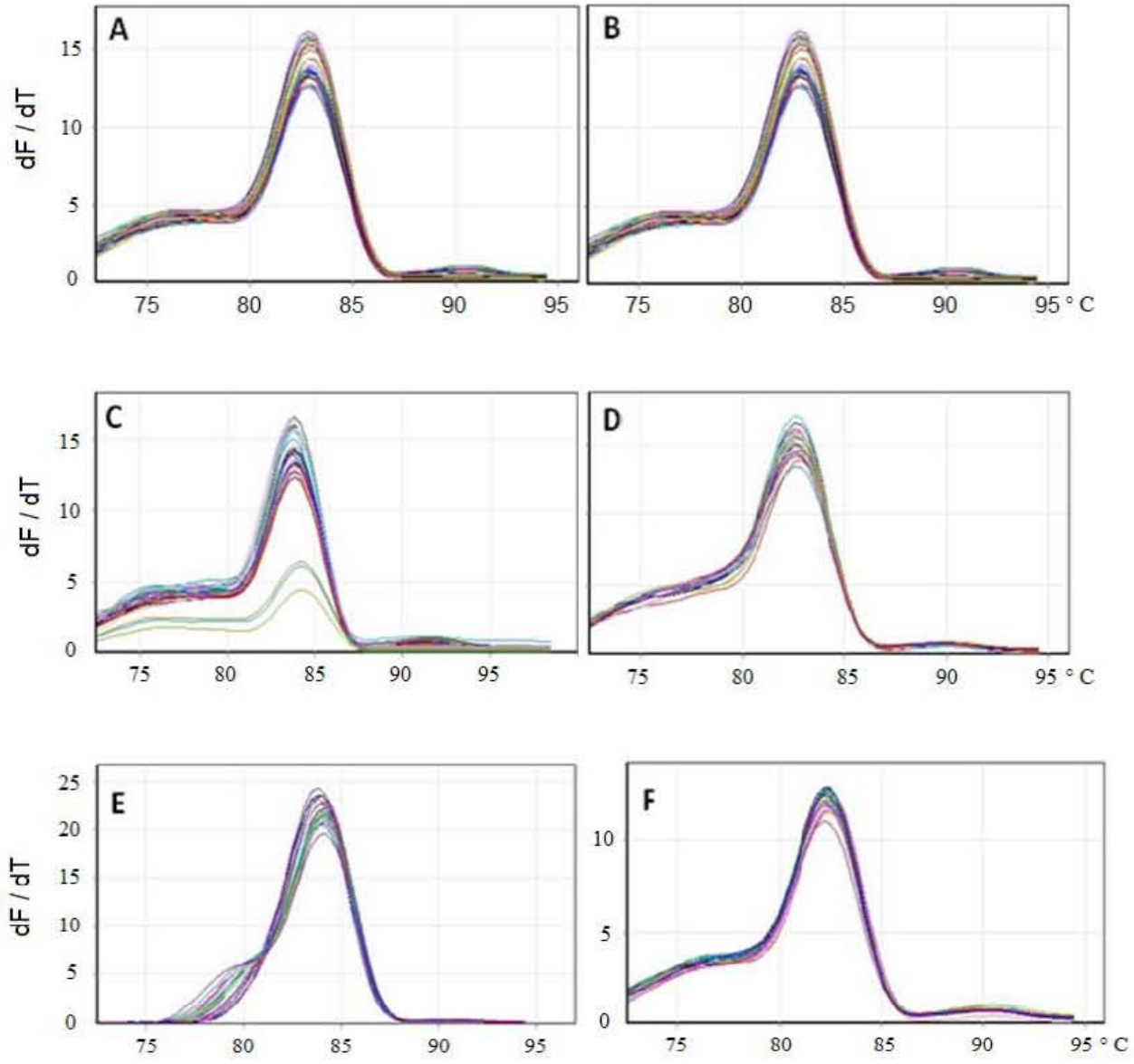


Figure 3.12 Melt curve analyses of GOI and RG primer sets used in 24 hour time-course experiment. The x-axis represents the range of melting temperature (75 – 95°C) applied to the qPCR amplicons. The y-axis represents the derivative of the raw data after smoothing (dF/dT) as calculated by Rotor-Gene 6000 software. (A) represents the melt curve for PRX2F; (B) PSP; (C) PSD; (D) serine-protease like; (E) thioredoxin and (F) RG 245.

Standard curves were prepared within each qPCR run for each gene (five GOIs and one RG), i.e. one standard curve was prepared per biological repeat instead of preparing one standard curve that was imported into the two remaining biological repeats. The standard curves were used to calculate the PCR amplification efficiency and correlation efficiency for each primer set. PCR amplification efficiencies ranged between 0.6 and 1.1 and were deemed to be satisfactory. As per the Pfaffl relative quantitation model, a value of one was added to all amplification efficiencies prior to relative quantitation (Table 3.3). Furthermore, the Pfaffl model was able to incorporate amplification efficiency into the relative quantitation calculation.

Table 3.3 PCR amplification efficiency and correlation efficiency for each GOI and RG 245 primer set for the 24 hour time-course experiment. Data was calculated from standard curves using the Rotor-Gene 6000 software. PCR amplification efficiencies are shown as means ± standard error, $n = 3$.

Gene of Interest (Clone ID)	PCR Amplification Efficiency	Correlation Efficiency
PRX2F (233)	1.73 (± 0.05)	0.999
PSP (991)	1.84 (± 0.02)	0.990
PSD (897)	1.62 (± 0.10)	0.960
Serine protease-like (428)	1.96 (± 0.01)	0.999
Thioredoxin (4)	2.00 (± 0.01)	0.990
RG 245	2.07 (± 0.07)	0.999

By using the Pfaffl mathematical model, relative changes in GOI expression were normalized using the expression of RG 245 and subsequently represented as n-fold changes relative to the time 0 calibrator sample. Inclusion of time-point controls (where *G. gracilis* was incubated for x hours without the addition of disease elicitors to the growth media) in the 24 hour time-course experiment was useful since the GOI expression pattern observed in the control samples could be assumed to be representative of normal expression of each GOI in *G. gracilis*. The relative expression of each putative defence gene in response to the presence or absence of elicitors will be presented separately in the next section.

Analysis of PRX2F gene expression revealed that the PRX2F mRNA transcript was regulated in response to exposure to disease elicitors (Figure 3.13). Data was $\log_{10}$ transformed such that the requirements of parametric statistical tests were met. In the experimental samples, an increase in gene expression was observed as early as 1 hour after exposure to disease elicitors. This trend appeared to be maintained over the duration of the 24 hour experiment. A *t*-test comparing the mean gene expression between the experimental and control samples at 1 hour after exposure to elicitors failed to detect statistically significant changes in PRX2F expression. The power of the performed test (0.158) was below the desired power of 0.800. Low power is typically as a result of too few replicates or large variability in the data set (Williams *et al*., 1997). It was therefore possible that a type II error (false negative) was committed. Similarly, a *t*-test was employed to analyse PRX2F expression in the experimental and control samples at 8, 12 and 24 hours. The observed induction of PRX2F expression in the experimental sample was statistically significant ($P < 0.05$) at 8 hours but not at 12 and 24 hours. Once again, the power of the latter two tests was less than 0.800 which suggested that a type II error was possibly committed. It was concluded that the low number of biological repeats (three) in this study, together with the large standard error of the means, contributed to lowering the power of the statistical tests. Cumming *et al*. (2007) suggest that by increasing the number of biological repeats, narrower inferential error bars with more precise estimates of the true population values can be achieved. Nevertheless, a clear biological trend of increased PRX2F expression in response to exposure to disease elicitors was evident.

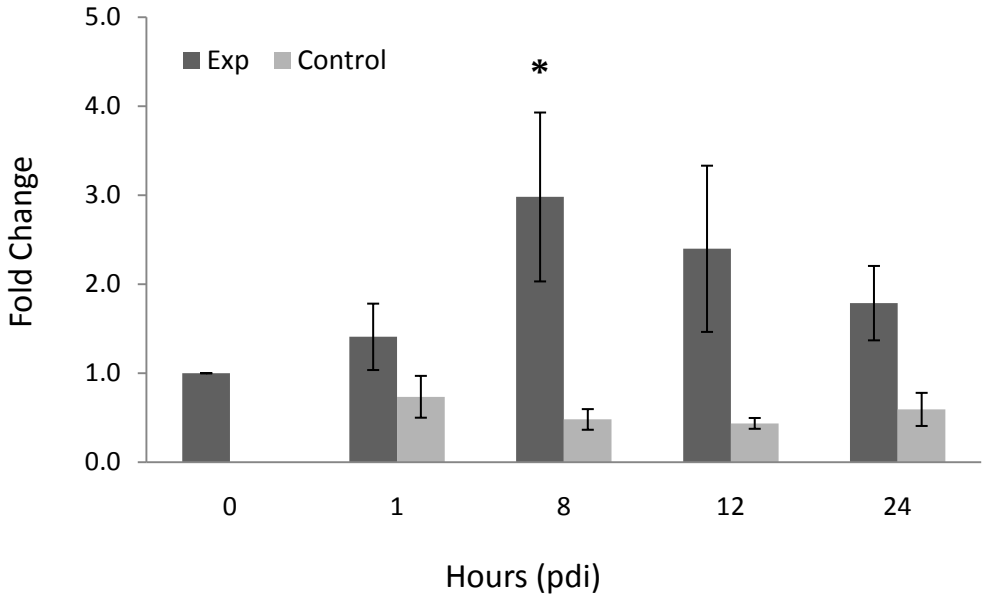


Figure 3.13 Gene expression data for PRX2F. The x-axis represents the respective time-points (hours) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests detected a statistically significant difference between the experimental and control samples at the 8 hour time-point ($P < 0.05$) (denoted by '*'). Data are shown as means ± standard error, $n = 3$ but were $\log_{10}$ transformed to meet the requirements of parametric statistical testing.

Statistical analysis of PSP gene expression failed to detect significant changes in gene expression between the experimental and control samples at each time-point (Figure 3.14). As observed for PRX2F expression, large error bars were also evident in the measured PSP expression. Similarly, *t*-tests were affected by low power (less than 0.800) which suggested that a type II error (false negative) may have been committed. In this context, a biological trend in the pattern of PSP expression was observed in the seaweed samples exposed to disease elicitors. Between 8 and 12 hours post exposure to elicitors, gene expression in the experimental samples was repressed by approximately two-fold relative to the time-point control samples. At 24 hours, expression between experimental and control samples appeared relatively equal to each other. Thus, the effect of disease elicitors on the transcriptional regulation of PSP was most significant between 8 and 12 hours. However,

more biological repeats are required to increase the power of the *t*-tests before the observed changes in gene expression can be attributed to the presence of the disease elicitor.

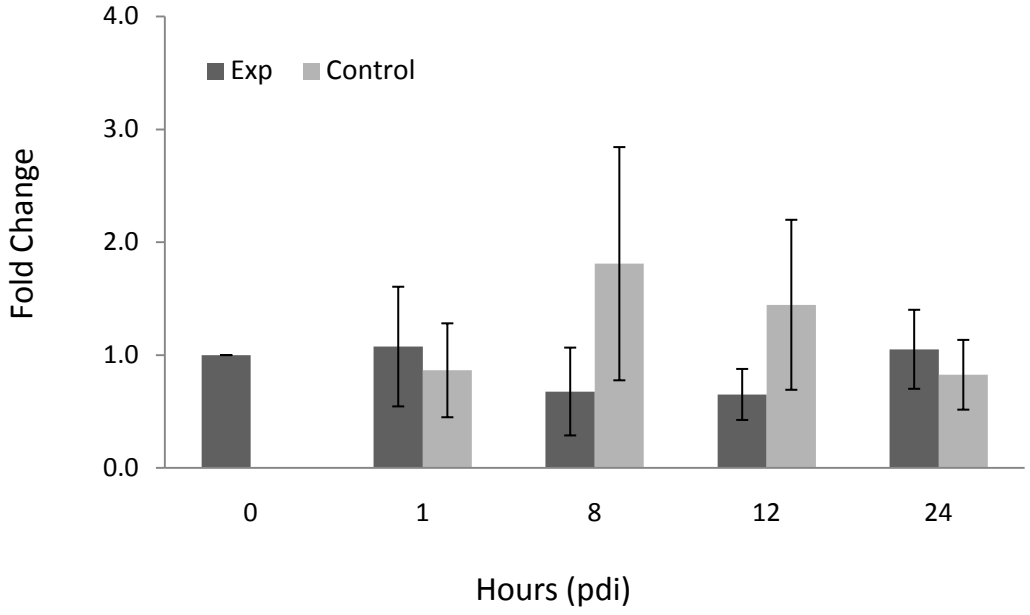


Figure 3.14 Gene expression data for PSP. The x-axis represents the respective time-points (hours) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests failed to detect statistically significant changes between the experimental and control samples at each of the time-points. Data are shown as means ± standard error, $n = 3$.

Analysis of PSD expression in the 24 hour time-course experiment revealed a very distinct repression in transcription in the samples exposed to elicitors (Figure 3.15). Even though PSD expression in the experimental sample at 1 hour appeared to be approximately 2-fold less than the control sample, a *t*-test failed to detect a statistically significant difference. However, the power of the performed test (0.676) was lower than 0.800 which suggested that the result was possibly a type II error. Similarly, the marked repression in PSD expression in the experimental sample relative to its time-point control at 8 hours failed to be statistically significant, but was also subject to a low power of the performed test (0.155). However, at 12 and 24 hours post exposure to elicitors, the observed repression of PSD expression relative

to the time-point controls was statistically significant ($P < 0.001$ and $P < 0.05$, respectively). Together these observations suggested that the repression of PSD in response to disease elicitors was a rapid transcriptional response in the activated defence of *G. gracilis*. Interestingly, expression of PSD in the control samples steadily increased over the 24 hour period. At 24 hours, PSD expression was approximately 7-fold greater than in the time 0 sample. However, the physiological reason for the observed increase in PSD expression in healthy seaweed over a 24 hour period remained unclear. It is possible that in healthy seaweed, the PSD protein product is necessary for various molecular functions within the cell. It would have been interesting to characterize transcriptional regulation beyond 24 hours to establish whether this increase in PSD expression was maintained. As stated in Chapter 2, PSD is an integral trans-membrane protein that catalyzes the formation of phosphatidylethanolamine (PtdEtn) from phosphatidylserine (PtdSer) (Voelker, 1997). Phospholipid-derived molecules appear to be more than just structural components of membranes. Scientific evidence has linked them to co-factors for membrane-enzymes, signal pre-cursors and even signalling molecules in plant defence (Laxalt and Munnik, 2002). Therefore, further research is required to establish how PSD functions under conditions in which *G. gracilis* is not exposed to elicitors.

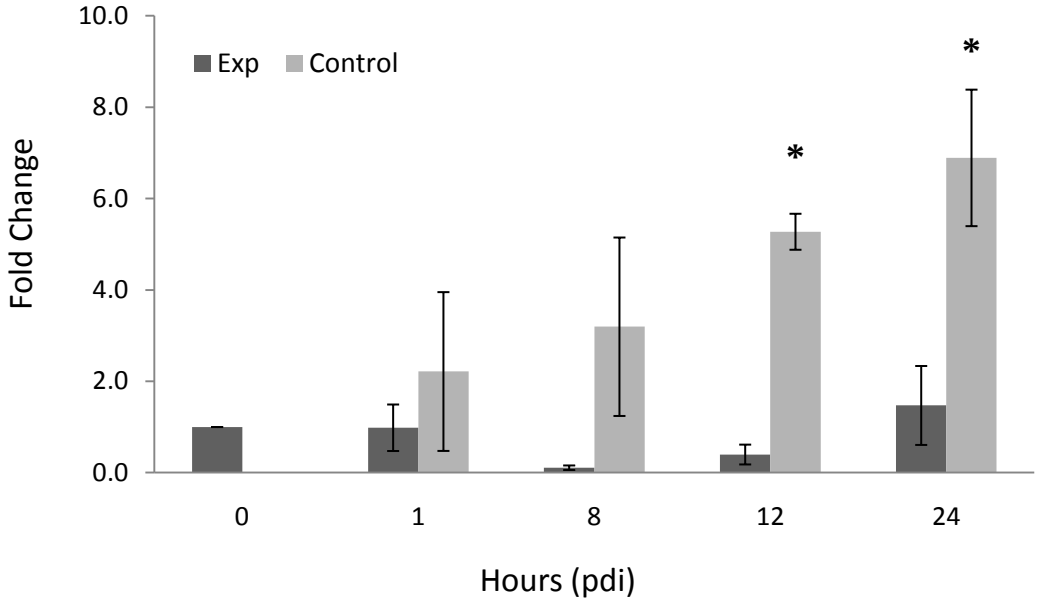


Figure 3.15 Gene expression data for PSD. The x-axis represents the respective time-points (hours) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests detected statistically significant differences in PSD gene expression in the experimental and control samples (denoted by '*'). At 12 hours, the statistical significance was at $P < 0.001$ and at 24 hours, the statistical significance was at $P < 0.05$. Data are shown as means ± standard error, $n = 3$.

Analysis of serine protease-like gene expression detected no statistically significant differences in gene expression over the 24 hour time-course. However, it appeared as if exposure to elicitors caused a repression of gene expression at 1, 8 and 12 hours relative to the control samples (Figure 3.16). Due to the variability in the data sets (large error bars) as well as the low power of the performed *t*-tests (< 0.800), it was possible that type II errors were committed during statistical analysis. More biological repeats are thus required to increase the power of the statistical tests as well as to reduce the variability in the data before the transcriptional regulation of this serine protease-like mRNA transcript, in response to disease elicitors, can be interpreted correctly.

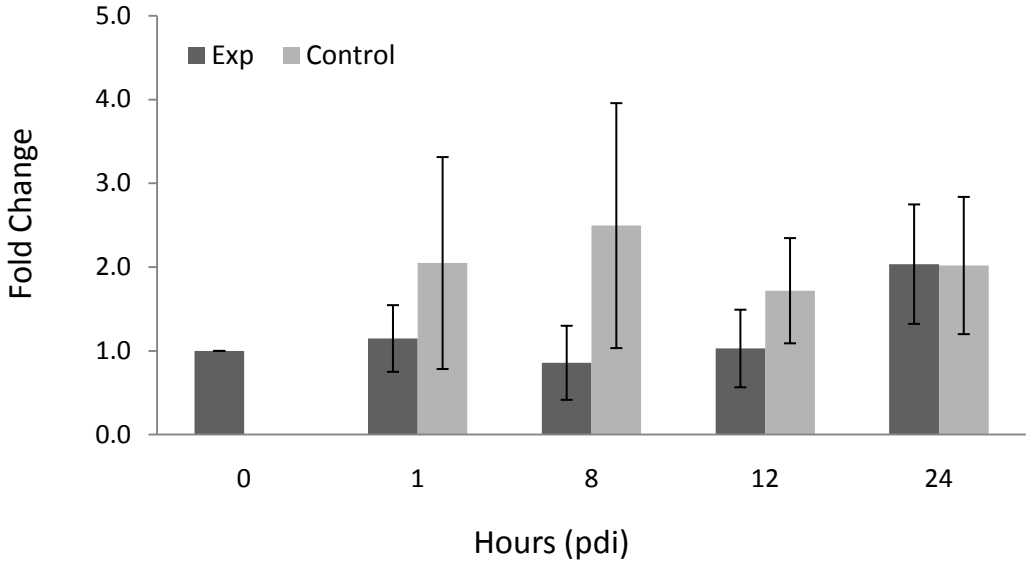


Figure 3.16 Expression data for the serine protease-like gene. The x-axis represents the respective time-points (hours) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests failed to detect statistically significant changes between the experimental and control samples at each time-point. Data are shown as means ± standard error, $n = 3$.

Similarly, it appeared that exposure to disease elicitors resulted in repression of thioredoxin expression relative to the control samples at the 8, 12 and 24 hour time-points (Figure 3.17). Although, transcription of the thioredoxin gene was repressed in the experimental samples by approximately 5-fold, 2-fold and 1.3-fold, respectively, *t*-tests failed to detect statistically significant differences in thioredoxin expression between the experimental and control samples at these time-points. As before, the low number of biological repeats (3) and the variability of the data affected the power of each *t*-test performed (< 0.800) which resulted in possible type II errors. More biological repeats are thus required to confirm whether thioredoxin is in fact significantly transcriptionally regulated upon exposure to disease elicitors.

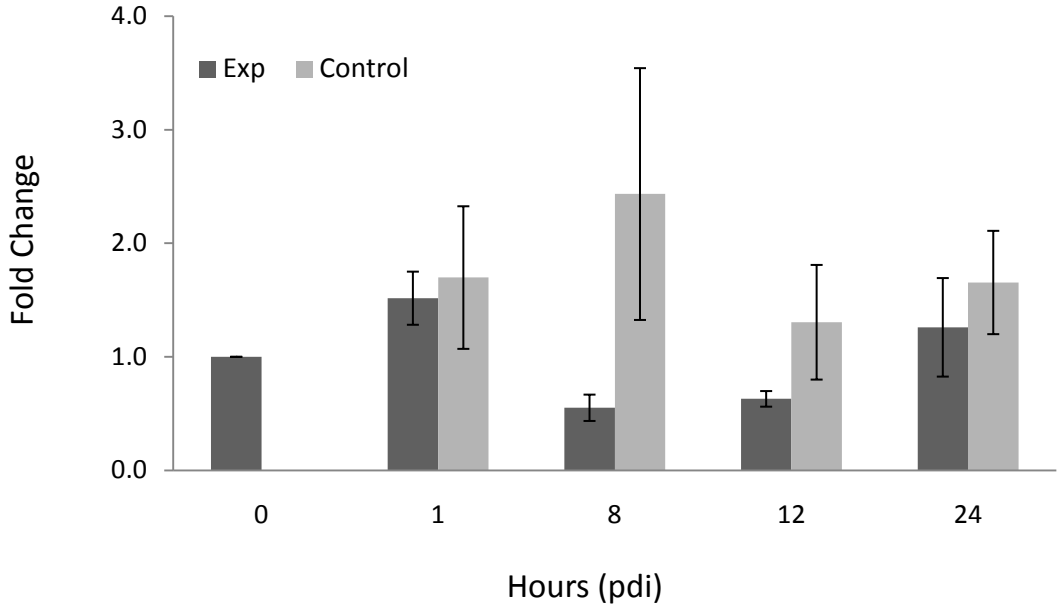


Figure 3.17 Gene expression data for thioredoxin. The x-axis represents the respective time-points (hours) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests failed to detect statistically significant changes between the experimental and control samples at each time-point. Data are shown as means ± standard error, $n = 3$.

3.3.6 Relative quantitation of putative defence genes during the 30 min experiment

In order to characterise transcriptional regulation of the five putative defence GOIs over a 30 min time-course, it was necessary to confirm that RG 329 was stably expressed in all three biological repeats. As expected, RG 329 expression did not fluctuate significantly in the control and experimental samples over the 30 min time-course (Figure 3.18). In addition, the Ct values for RG 329 in BRs 2 and 3 were within three units of each other while the Ct values of RG 329 expression in BR1 did not differ by more than 4 units which was deemed acceptable for a reference gene.

	0	5E	5C	15E	15C	30E	30C
BR1	20	20	23	21	19	21	19
BR2	21	21	22	22	22	22	21
BR3	20	19	19	19	21	19	19

Figure 3.18 Average Ct values of RG 329 across all three biological repeats over 30 minutes. The x-axis represents time-point (hours) where "E" and "C" denote experimental and control samples respectively. Average Ct values for each biological repeat were determined automatically in the Rotor-Gene 600 software from the respective RG 329 standard curves.

Normalization and statistical analyses of gene expression for all five GOIs was performed similarly to the 24 hour experiment. Amplification and correlation efficiencies were calculated from the standard curves for each primer set and were deemed optimal with values ranging between 1.75 and 2.0 (Table 3.4). As per the Pfaffl relative quantitation method, a value of one was added to all amplification efficiencies prior to data analysis. Relative changes in GOI expression were represented as an n-fold change relative to the time 0 calibrator sample. Similarly, the inclusion of time-point controls (where *G. gracilis* was incubated for x min without the addition of disease elicitors to the growth media) in the 30 min time-course experiment was important because it was assumed that they would be representative of normal expression for that GOI in *G. gracilis*. In total, three BRs were performed to determine the relative changes in GOI expression over the 30 min time-course

experiment. The dissociation curves for all five GOI primer sets already proved the high specificity of the PCR reactions as evident by amplification of a single PCR product (Figure 3.12). Similarly, PCR specificity of the RG 329 primer set was assessed which revealed high amplification specificity (Figure 3.19).

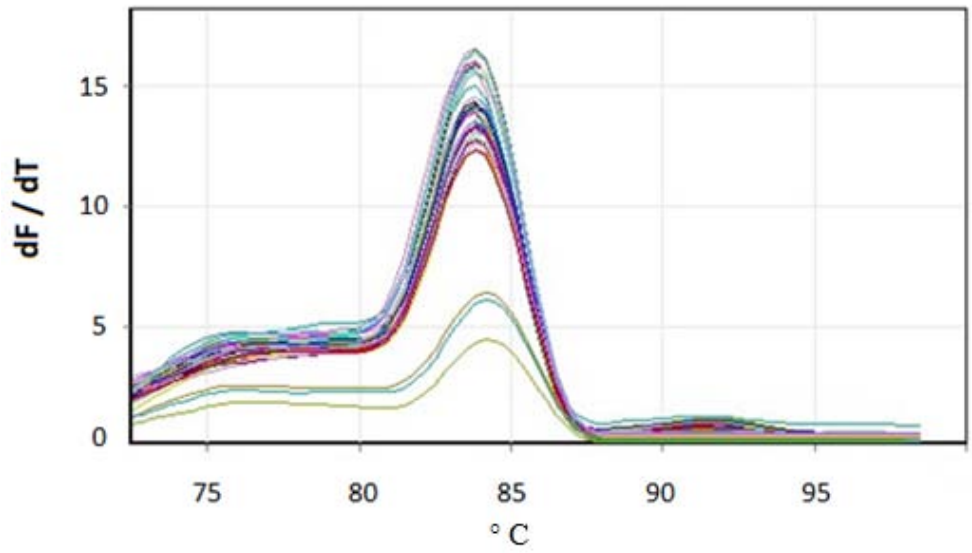


Figure 3.19 Melt curve analyses for RG 329. The x-axis represents the range of melting temperatures (75 – 95°C) applied to the qPCR amplicons. The y-axis represents the derivative of the raw data after smoothing (dF/dT) as calculated using Rotor-Gene 6000 software.

Table 3.4 Amplification and correlation efficiencies for each GOI primer set and RG used for normalization (30 min experiment). Data was calculated from standard curves using the Rotor-Gene 6000 software. PCR amplification efficiencies are shown as means ± standard error, $n = 3$.

Gene of Interest	Amplification Efficiency	Correlation Efficiency
PRX2F	1.75 (± 0.01)	0.998
PSP	1.80 (± 0.02)	0.990
PSD	1.84 (± 0.02)	0.999
Serine protease-like	1.91 (± 0.02)	0.999
Thioredoxin	1.99 (± 0.02)	0.870
RG 329	1.84 (± 0.01)	0.994

Statistical analyses applied to the gene expression data sets for all five GOIs detected no statistically significant changes in gene expression as a consequence of exposure to disease elicitors during the first 30 min (Figure 3.20 – 3.24). Despite the failure to detect statistically significant changes, the expression of three putative defence genes appeared to be altered in certain experimental samples. For example, PRX2F expression in the experimental sample at 15 min post disease induction (pdi) was approximately 2.5-fold greater than its time-point control (Figure 3.20). PSD expression in the experimental samples at time-points 5, 15 and 30 min pdi all appeared distinctly different relative to their control time-points (Figure 3.22). Thioredoxin expression in the experimental samples at 15 and 30 min pdi was approximately 3-fold repressed and 3-fold induced relative to the respective time-point controls (Figure 3.24). Regardless of these possibly biologically relevant transcription patterns, the low

power of the *t*-tests and the high variability in standard error of the means were a function of the low number of biological repeats employed in the experiment (n = 3). It was therefore possible that type II errors were committed. As a result, it was concluded that more biological repeats were required to correctly establish and interpret the effects of disease elicitors on transcription of the five GOIs tested in the 30 min time-course experiment.

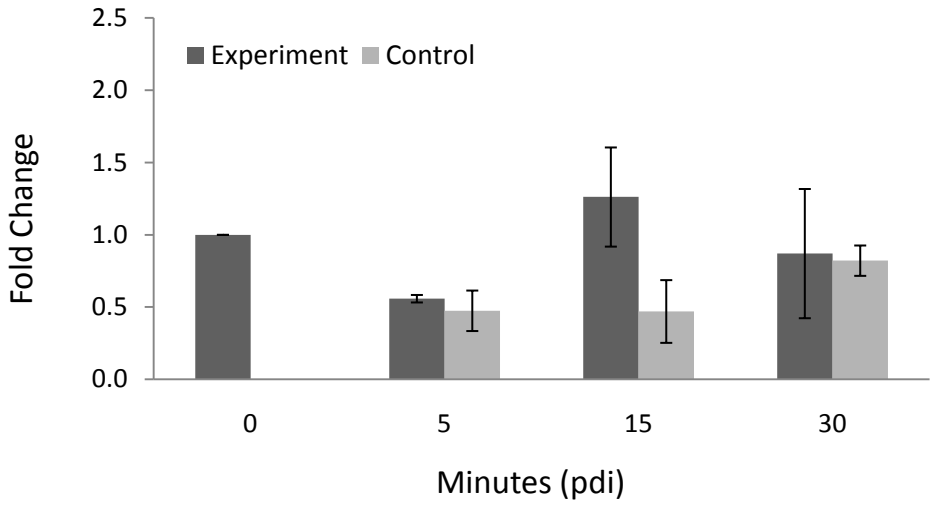


Figure 3.20 Gene expression data for PRX2F over the 30 min time-course experiment. The x-axis represents the respective time-points (min) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests failed to detect statistically significant changes between the experimental and control samples at each time-point. Data are shown as means ± standard error, n = 3.

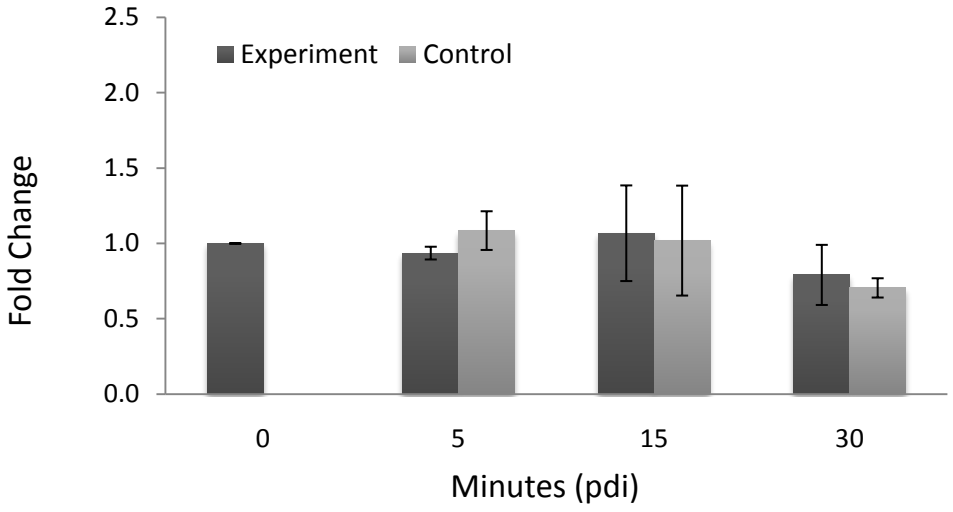


Figure 3.21 Gene expression data for PSP over the 30 min time-course experiment. The x-axis represents the respective time-points (min) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests failed to detect statistically significant changes between the experimental and control samples at each time-point. Data are shown as means ± standard error, $n = 3$.

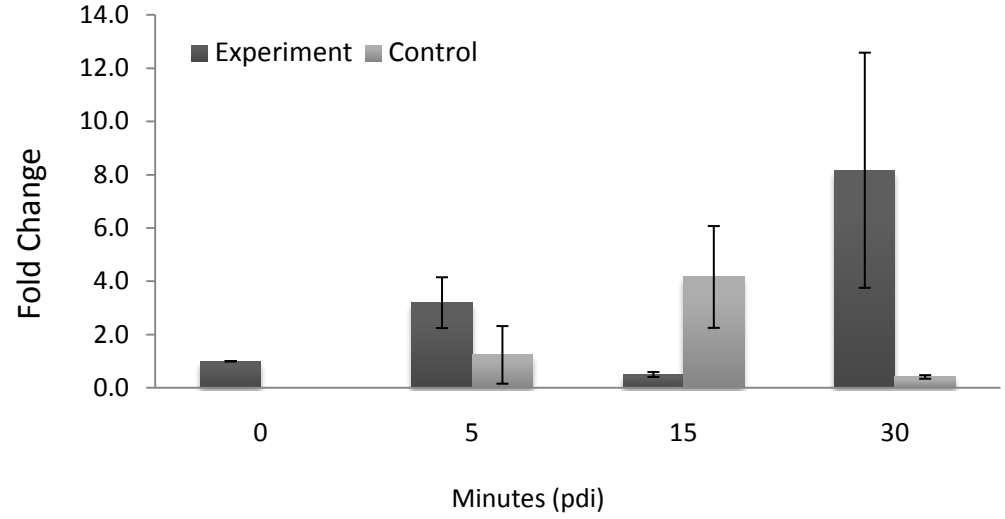


Figure 3.22 Gene expression data for PSD over the 30 min time-course experiment. The x-axis represents the respective time-points (min) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests failed to detect statistically significant changes between the experimental and control samples at each time-point. Data are shown as means ± standard error, $n = 3$.

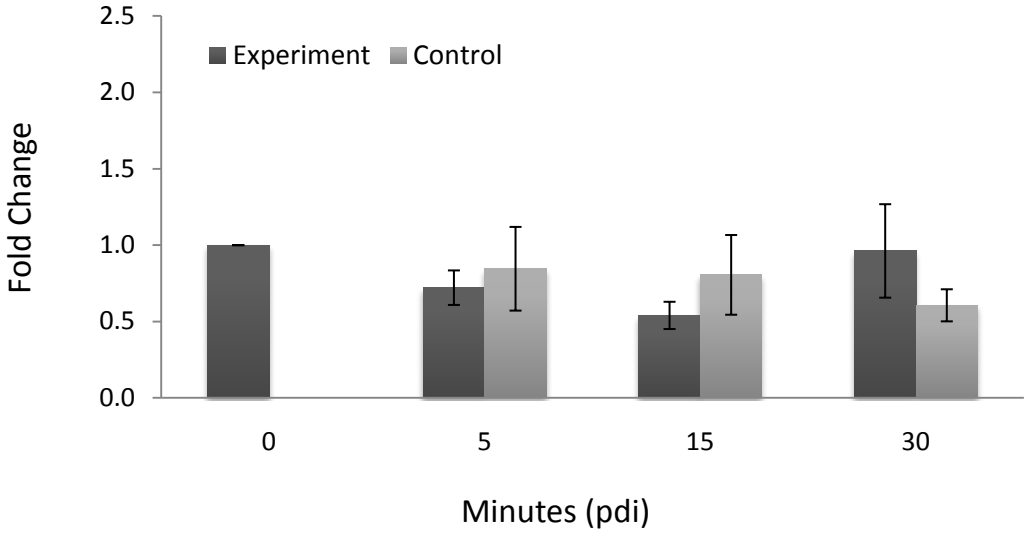


Figure 3.23 Expression data for serine protease-like gene over the 30 min time-course experiment. The x-axis represents the respective time-points (min) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests failed to detect statistically significant changes between the experimental and control samples at each time-point. Data are shown as means ± standard error, $n = 3$.

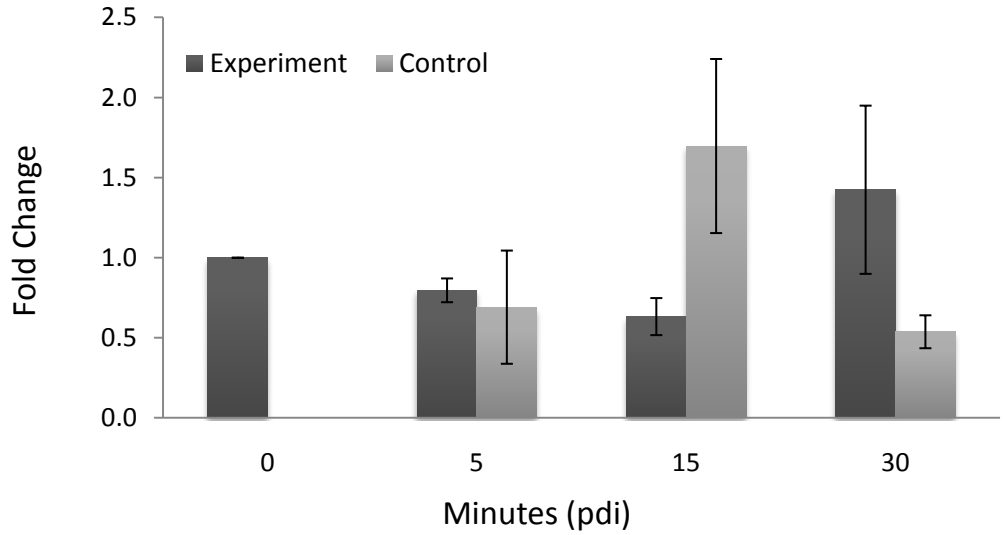
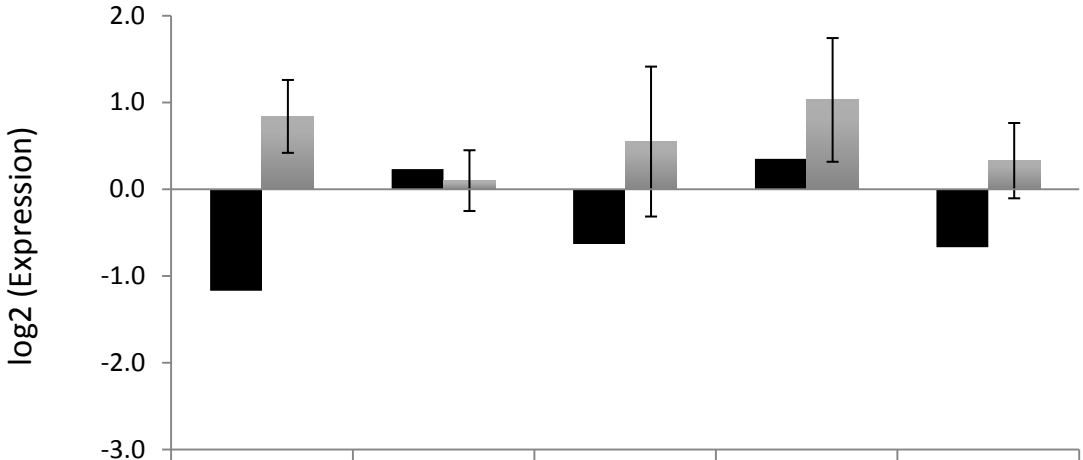


Figure 3.24 Gene expression data for thioredoxin over 30 min. The x-axis represents the respective time-point (min) post disease induction (pdi). The y-axis represents fold change in expression relative to time 0. Parametric *t*-tests failed to detect statistically significant changes between the experimental and control samples at each time-point. Data are shown as means ± standard error, $n = 3$.

3.3.7 Validation of microarray experiment

Expression of the five putative defence genes after 24 hours of exposure to disease elicitors was compared between the microarray and qPCR experiments. In the microarray experiment, gene expression after 24 hours was represented as a fold-change relative to time 0. In addition, this fold-change was $\log_2$ transformed such that two-fold induction of gene expression was represented by an M-value of 1 and a two-fold repression was represented by an M-value of –1. Gene expression in the qPCR experiments was also represented as a fold-change relative to time 0, but in order to validate the microarray results the qPCR gene expression data was $\log_2$ transformed (Figure 3.25). Expression of two genes (i.e. the PSP gene and the serine-protease like gene) was successfully validated as both molecular techniques detected transcriptional up-regulation after 24 hours. However, qPCR failed to validate the microarray data regarding expression of the PRX2F, PSD and thioredoxin genes. It is important to note that large error bars were associated with the qPCR gene expression data which are indicative of large variation within the data.



	PRX2F	PSP	PSD	Serine protease-like	Thioredoxin
log2(M-value)	-1.17	0.23	-0.63	0.35	-0.67
log2(qPCR)	0.84	0.10	0.55	1.03	0.33

Figure 3.25 Gene expression at 24 hours post exposure to disease elicitors as determined by microarray (dark blue bars) and qPCR (grey bars) analysis. The relative changes in gene expression were calculated as ratios relative to the 0 time-point control. The y-axis represents $\log_2$ transformation of relative fold-change. qPCR data are shown as means ± standard error, $n = 3$.

3.4 DISCUSSION

3.4.1 Validation of microarray experiment

Quantitative real-time PCR (qPCR) is commonly used to validate gene expression results obtained in microarray analysis (Morey *et al*., 2006). In addition to validating the microarray results (Chapter 2), the qPCR experiments presented in this chapter assessed transcriptional regulation of five putative defence genes during the first 24 hours of exposure to disease elicitors. A positive correlation between the two molecular techniques was observed for only two out of five genes analyzed. Failure to validate the transcriptional regulation of the three remaining genes was unexpected but not surprising. The most significant factor that may have contributed to the observed discrepancy between the microarray and qPCR experiments was the fact that mRNA used in either experiment was from different biological samples. This could not be avoided since all the RNA was exhausted in the microarray experiment. However, literature suggests that it is not uncommon that the two molecular techniques result in disagreement (Morey *et al*., 2006). In one study, the red macroalga *C. crispus* was exposed to methyl jasmonate over a 24 hour time-course experiment and differentially expressed genes were identified using microarray analysis (Collén *et al*., 2006). qPCR was used in order to validate the observed transcriptional regulation for several of these genes. The authors observed a discrepancy in the expression data obtained with the two molecular techniques for one-third of the genes (2 out of 6) examined. A similar discrepancy in qPCR validation was observed in a different microarray study on the red macroalga *G. changii* (Teo *et al*., 2009). This phenomenon is not unique to organisms whose genomes are not completely sequenced (such as *G. gracilis*). For example, the correlation between microarray and qPCR results obtained from an optimized commercial human genome microarray was low because approximately 13 – 16% of transcripts displayed conflicting transcriptional regulation patterns (Dallas *et al*., 2005). There are several reasons as to why this discrepancy between microarray and qPCR techniques may arise. It is well established that both molecular tools have inherent limitations that may significantly influence the data obtained from each method (Morey *et al*., 2006; Wurmbach *et al*., 2003; Bustin, 2002; Chuaqui *et al*., 2002; Freeman *et al*., 2002; Yang *et al*., 2002). Essentially, these limitations are due to biases arising from biological and technical factors. For example, the quality and integrity of RNA is a major factor toward achieving accurate results. The effect of RNA amplification

must also be taken into account since the total RNA used in the microarray analysis was amplified whereas qPCR experiments used unamplified total RNA. However, published evidence proposes that T7-based amplification systems closely approximates the original RNA sample for gene expression analysis using cDNA microarrays or qPCR (Vermeulen *et al*., 2009; Degrelle *et al*., 2008; Zhao *et al*., 2002). Thus, the effect of using amplified RNA in the microarray experiment versus unamplified RNA in qPCR validation was unlikely the cause of the observed discrepancy in gene expression. Furthermore, the RNA extraction methods used were different between the microarray and qPCR experiments. Consequently, varying degrees of contaminating substances such as polyphenols or polysaccharides may have affected cDNA synthesis for each molecular technique differently. Microarray experiments are susceptible to the effects of fluorescent dye-biases and non-specific hybridizations whereas relative quantitation in qPCR experiments is susceptible to the formation of primer dimers and the reduction in efficiency of *taq* polymerase in the later amplification cycles. Another important variable was that data normalization between the two techniques was fundamentally different. Microarray data was normalized using linear models while qPCR experiments relied on stable expression of one or more reference genes. While the pitfalls contributing to discrepancies between the two methods may potentially be significant, most sources of error can be controlled through proper experimental design, appropriate controls and normalization of the data (Morey *et al*., 2006). The fact that transcriptional regulation of two genes was successfully validated using qPCR suggests that both molecular tools (cDNA microarrays and qPCR) can be used with confidence to identify differentially expressed genes. The observations in this qPCR study also further emphasize the importance of performing an adequate number of biological repeats as evident by the large error bars which may have influenced the interpretation of validation of transcriptional regulation.

3.4.2 Analysis of transcriptional regulation of five genes involved in defence response

Before discussing the gene expression patterns observed in the qPCR experiments for each putative defence gene, it is important to bear in mind the context in which they were acting, i.e. the current model of activated defence in macroalgae (outlined in detail in Chapter 1,

Section 1.5). The question therefore becomes: Based on what is already known about activated defence in macroalgae, are the changes in gene transcription as a result of exposure to an endogenous elicitor physiologically relevant? In order to propose physiological roles for each putative defence gene based on its pattern of transcription, only the long (24 hours) time-course qPCR experiment will be discussed in detail.

The effect of disease elicitors on transcriptional regulation of the putative defence genes over the 30 min time-course experiment was not as distinct as those observed during the 24 hour time-course experiment. All five genes displayed fluctuations in expression as a consequence of exposure to disease elicitors over the 30 min period, but no statistically significant changes were detected due to the low power of the performed *t*-tests. This suggested that 30 min may have been too short a time period to detect significant transcriptional changes in the samples exposed to disease elicitors or that several more biological repeats must be performed before conclusive results can be obtained. On the other hand, the results of the 24 hour time-course qPCR experiments (although subject to large variability in the gene expression data and consequently low power of *t*-tests) clearly demonstrated that significant changes in gene expression occurred in *G. gracilis* as a result of exposure to agar and agar oligosaccharides (disease elicitors). Furthermore, secondary signalling (most likely via the production of reactive oxygen species) had to take place in order to cause the observed transcriptional changes in expression of genes involved in the defence response of *G. gracilis*. In theory, if the agar oligosaccharides added to the surrounding growth media were not effective disease elicitors, gene expression in the experimental and control samples would have been approximately equal. In addition to proving that the putative defence genes were transcriptionally regulated, several stably expressed reference genes were identified and used to normalize the qPCR data. The identification of reference genes was a significant finding for two reasons: (i) reference genes had not been previously identified for *G. gracilis* qPCR experiments and (ii) without the application of appropriate normalization strategies, the qPCR results would have been compromised. However, it was evident that the small number of biological repeats (3) resulted in significant variation in the measured gene expression values (represented by large error bars). Consequently, statistical analyses often detected no statistically significant changes in gene expression. As previously discussed in Chapter 2, a lack of statistical significance or relatively small changes in gene expression does not

necessarily equate to a lack of biological relevance in *G. gracilis*. Cumming *et al*. (2007) highlighted the error in over-interpreting a lack of statistical significance. If the *P-values* generated in statistical tests are greater than a certain cut-off (0.05 in this case), it would obviously be concluded that no statistically significant effect was evident but it should not be concluded that the effect is zero. The biological effect may have been small or the experiment was not repeated enough times to reveal it. Therefore, where a lack of statistically significant changes in GOI expression was detected, data was interpreted in the context of a biological trend instead.

Upon closer inspection of each GOI's pattern of transcriptional regulation during the 24 hour time-course experiment, incorporation of corresponding time-point controls in the experimental design proved useful. In contrast to the microarray experiment which only assessed gene expression at 24 hours after exposure to disease elicitors relative to time 0, the qPCR experiment offered a more detailed pattern of gene expression. The most significant changes in gene expression were generally observed between 8 and 12 hours after exposure to the elicitors. This observation was corroborated by two separate studies. Zou *et al*. (2005) used microarrays for expression profiling of soybean in response to infection with the pathogen *Pseudomonas syringae*. The most number of transcriptional changes for the genes analyzed were observed 8 hours post bacterial infection. Collén *et al*. (2006) exposed the red macroalga *C. crispus* to methyl jasmonate. Using microarrays, the authors observed the most dynamic changes in gene expression after 6 hours post addition of methyl jasmonate to the growth media. The time-points at which gene expression was assessed were 0, 2, 4, 6, 12 and 24 hours post addition of methyl jasmonate. When the results of this study were directly compared to the Collén *et al* (2006) study, another interesting commonality emerged. A peroxiredoxin-like (CO650899) transcript was transcriptionally up-regulated in response to methyl jasmonate and increased expression of this gene was maintained over the 24 hour time-course. Similarly, expression of the PRX2F mRNA transcript from *G. gracilis* was induced as early as 1 hour post exposure to the agar oligosaccharide elicitor and maintained over 24 hours. Alignment of the two DNA sequences revealed a 60% sequence homology (Appendix C.4). As previously established in Chapter 2, the function of peroxiredoxin or peroxiredoxin-like proteins is to inhibit the damaging effects of ROS and to potentially regulate hydrogen peroxide-mediated signalling required for the regulation of stress

responses (Wood *et al*., 2003b; Foyer and Noctor, 2000). Although the generation of ROS production was not directly assessed in this study, exposure to agar oligosaccharides was previously shown to result in an oxidative burst (Weinberger *et al*., 2010; 1999) with an associated hypersensitive response (Weinberger *et al*., 1999) in *G. gracilis*. The hypersensitive response was quantitated by monitoring thallus tip bleaching. Although tip bleaching was not assessed in this study, the significant transcriptional up-regulation of the *G. gracilis* PRX2F gene confirmed that an oxidative burst did occur in response to the disease elicitors. Increased transcription of antioxidant genes during an activated defence response is supported by the study of Zou *et al.* (2005). The authors demonstrated that in response to bacterial infection of soybean, transcriptional up-regulation of defence, stress-related and oxidative genes took place which was thought to be required to mount a hypersensitive response (HR). Jabs (1999) proposed that the cell has at least three ways of responding to the generation of ROS. If a low dose of ROS is produced, antioxidant enzymes such as superoxide dismutase, catalases and peroxidases are induced. In the event of a more intermediate dose, plant cell death (PCD) pathways are activated. If ROS levels are high, death will occur via cell necrosis. Basically, this process of pathogen-induced defence and ROS production can be characterized as three distinct phases: (i) induction, (ii) effector and (iii) degradation. The first phase is a weak, short-lived oxidative burst that usually occurs within the first two hours of exposure. Physiological changes occur at the level of signalling, primary metabolism, cell wall structure and membrane metabolism. The second phase occurs in direct response to the first phase and normally involves a much stronger oxidative burst that leads to an HR. Physiological changes occur at the level of secondary metabolism and more complex signalling pathways to transmit and amplify the PCD message (Buckner *et al*., 2000). Studies indicate that the coordinating centre for the PCD message is the mitochondria. Changes in permeability of the mitochondrial membrane result in the opening of pores formed between adenosine nucleotide transporters in the inner membrane and porins in the outer membrane. The swelling of mitochondria during PCD is thought to occur due to this pore opening which leads to hyperosmolarity of the mitochondrial matrix and subsequent expansion of the matrix space. The formation of these pores is stimulated by various PCD signals including Ca^{2+} ions, H_2O_2 and proteases (Jabbs, 1999). In the final degradation phase, phenomena such as PCD, DNA fragmentation (via the recruitment and activation of proteases and nucleases) and the shutting down of photosynthesis are observed as a last attempt to limit pathogen proliferation. In this context (i.e. three distinct phases of an

activated defence response), the pattern of transcriptional up-regulation of the *G. gracilis* PRX2F in response to disease elicitors seems to agree. After 1 hour of exposure to the elicitors, up-regulation of the PRX2F gene was already detected (corresponds to the first phase of activated defence). After 8 hours, the expression of the PRX2F transcript was maximally induced which suggested that the second phase may have been activated in *G. gracilis*, i.e. the HR and PCD. During the third phase (12 – 24 hours), the concentration of the PRX2F mRNA transcript remained higher relative to the control samples. Before this proposed hypothesis (activation of all three phases in activated defence in *G. gracilis*) is accepted, translational regulation of the gene product (protein) in response to disease elicitors over a 24 hour time-course must be assessed. In a study by Horling *et al*. (2003), the effects of oxidative stress on the expression of seven different peroxiredoxin genes in *A. thaliana* leaves were clearly demonstrated. Leaf slices were incubated in the presence of H_2O_2 and other oxidative stressors for 4 hours. Expression of the peroxiredoxin genes was induced in each case, corroborating the hypothesis that expression of the *G. gracilis* peroxiredoxin-like gene observed in this study was induced in response to oxidative stress that occurred as a consequence of exposure to disease elicitors.

The most notable transcriptional changes in expression of the *G. gracilis* phosphoserine phosphatase (PSP) gene occurred between 8 and 12 hours in the seaweed samples exposed to disease elicitors. As established in Chapter 2, the formation and breakage of phosphate monoester bonds via the opposing actions of protein kinases and protein phosphatases potentially impact several biological processes in eukaryotic organisms (Wang *et al*., 2002). Phosphorylation of proteins via the action of various plant protein kinases have been shown to play essential roles in defence mechanisms, DNA replication (Muzi-Falconi *et al*., 1996), signal transduction, cell cycle control (Wilkinson and Millar, 2000; Parsons and Spencer, 1997), protein translation (Sarre, 1989) and energy metabolism (Stone and Walker, 1995; Bowler and Chua, 1994; Hubber *et al*., 1994; Budde and Randall, 1990). Protein kinases are thought to directly respond to second messengers such as Ca^{2+} making them primarily responsible for regulating the phosphorylation status of proteins. On the other hand, protein phosphatases are thought to reverse the effects of protein kinases (Smith and Walker, 1996). In this context, if the phosphorylation of defence proteins by protein kinases is required to activate them, an associated transcriptional down-regulation of the genes linked to reversing

phosphorylation (phosphatases) was an expected physiological response. This inference is supported by the study of Weinberger *et al*. (2005). The authors showed that protein kinase inhibitors prevented an oxidative burst in *Gracilaria* species exposed to elicitors while phosphatase inhibitors enhanced the production of ROS after elicitor perception. The observed maximal transcriptional repression of PSP between 8 and 12 hours was an interesting result because the same time-points were associated with the most significant transcriptional up-regulation of PRX2F (antioxidant). To further understand reversible phosphorylation during an activated defence response of *G. gracilis* however, transcriptional regulation of the protein kinases involved in the activated defence response must be assessed (not done in this study). Furthermore, transcriptional regulation of phosphatases and kinases in the presence of their specific inhibitors may further elucidate the role of PSP in *G. gracilis*. Alternatively, evidence suggests that plant phosphatases may not only function to counterbalance kinases but may also play key roles in signalling cascades of abscisic acid (ABA), pathogen or stress responses and developmental process (Luan, 1998).

Transcriptional regulation of phosphotidylserine decarboxylase (PSD) in response to exposure to disease elicitors was very distinct. Throughout the entire duration of the 24 hour time-course, a significant repression of PSD expression in the *G. gracilis* samples exposed to disease elicitors was evident. As discussed in Chapter 2, PSD is an integral trans-membrane protein that catalyzes the formation of phosphatidylethanolamine (PtdEtn) from phosphatidylserine (PtdSer). It acts to specifically decarboxylate a lipid-linked form of the serine moiety and is essential for phospholipid metabolism in both prokaryotic and eukaryotic organisms (Voelker, 1997). Phospholipids may alter the physical properties of membranes to increase or decrease ion flux and membrane transport, vesicle formation and endo- or exocytosis. Phospholipid transport among the organelles of eukaryotes constitutes one of the most fundamental processes of membrane biogenesis (Choi *et al*., 2005). In addition, phospholipid-derived molecules are more than just structural components as some have been implicated to function as signalling molecules in plant defence and co-factors for other membrane enzymes (Laxalt and Munnik, 2002). Therefore, the significant transcriptional regulation of the *G. gracilis* PSD by disease elicitors seems very interesting. In order to understand why the PSD gene is almost completely switched off during the activated defence response of *G. gracilis*, it may be useful to consider its mechanism of action. In addition, the

role of PtdSer (a substrate of PSD) and the use of the end-products of the PSD pathway may also aid in elucidating the role of PSD in activated defence. As outlined in Choi *et al*. (2005), PtdSer is synthesized in the ER or the mitochondria-associated membrane (MAM) in eukaryotes. In plants, PtdSer is synthesized via a completely different pathway compared to prokaryotes and yeast, i.e. plant PtdSer is synthesized via a calcium-dependent base-exchange reaction in which the head-group of an existing phospholipid is replaced with L-serine (Vance and Steenbergen, 2005). PtdSer is subsequently transported to either PSD1 (located in the mitochondria) or PSD2 (located in the Golgi apparatus) and between the MAM and mitochondria via the actions of ubiquitin-ligase and EF-hand, respectively, for conversion to PtdEtn. Following this conversion, PtdEtn in the mitochondria or Golgi apparatus is transported to other cellular organelles for synthesis of various phospholipids for membrane integrity or cellular signalling. Transcriptional repression of the PSD gene in response to disease elicitors can be explained as follows: apart from the role of PtdSer as a minor phospholipid in most biological membranes, it is also required for specific cellular functions (Vance and Steenbergen, 2005) such as the physiological activation of several proteins. Protein kinase C (PKC) and its various isoforms are activated via receptor-mediated hydrolysis of inositol phospholipids such as PtdSer. PKC plays a crucial role in the diverse signal transduction pathways that occur across membranes and regulates many Ca^{2+}-dependent processes (Nishizuka, 1986). If this physiological role applies to *G. gracilis* PSD, i.e. the requirement of PtdSer to activate protein kinase C, repression of the PSD gene could result in less of the PSD enzyme which converts PtdSer into PtdEtn. Mitogen-activated protein kinase (MAPK) cascades are conserved in all eukaryotic organisms (Errede *et al*., 1995) and transcriptional repression of PSD may therefore implicate MAPK signalling in macroalgae. In *A. thaliana*, approximately 20 different MAPKs have been identified in the genome (Zhang and Klessig, 2001). MAPK cascades in plants are critical components in normal growth and development, responses to pathogen infection, wounding, extremes in temperature, drought, salinity stresses, UV-radiation and ROS (Zhang and Klessig, 2001). MAPKs may even induce expression of defence genes and PCD (Zhang *et al*., 2000; Suzuki *et al*., 1999; Zhang and Klessig, 1998). PKC has also been directly implicated in oxidative stress signalling in various degenerative diseases and cancer in mammals since various antioxidant inhibitors blocked PKC-dependent cellular responses (Gopalakrishna and Janken, 2000). As discussed earlier, protein kinases are required for the activation of NADPH oxidase and the subsequent oxidative burst in *Gracilaria* species exposed to disease elicitors

since the response was blocked by protein kinase inhibitors (Weinberger *et al*., 2005). Therefore, the higher the concentration of PtdSer in *G. gracilis* during an activated defence response, the greater the potential for activation of these essential protein kinases. Switching off the gene (PSD) that is responsible for the conversion of PtdSer to PtdEtn would essentially enable the activation of these essential protein kinases. Although this study has not identified genes directly involved in MAPK cascades, the observed transcriptional repression of PSD expression in response to disease elicitors has implications for the activation of PKC. If this proposed biological function of PSD is not applicable to *G. gracilis* however, an alternative role may be as follows. Exposure of PtdSer on the outer surface of the plasma membrane is characterized as an early event leading toward PCD (O'Brien *et al*., 1997). This directly implicates PCD in the activated defence response of *G. gracilis*. In other words, repression of PSD expression may have prevented the conversion of PtdSer to PtdEtn such that more PtdSer could be exposed on the outer surface of the plasma membrane to promote PCD during the activated defence response of macroalgae.

A biological trend in the transcriptional regulation of the *G. gracilis* serine protease-like gene was evident in response to disease elicitors, i.e. between 1 and 12 hours, gene expression in experimental samples appeared to be repressed relative to their time-point controls. However, *t*-tests failed to detect statistically significant changes which suggested that several more biological repeats are required to clarify transcriptional regulation of this particular serine protease-like gene upon exposure to disease elicitors. As discussed in Chapter 2, the involvement of plant serine proteases in various physiological processes has been widely documented (Palma *et al*., 2002). Of particular relevance to the activated defence response of *G. gracilis* is the possible role of serine proteases (or serine protease-like proteins) in a HR and signal transduction pathways. In addition to these roles, serine proteases have also been implicated in protein processing which was documented in the last two (of the three) distinct phases of an induced defence response in plants (Jabs, 1999). Based on the transcriptional regulation of the *G. gracilis* antioxidant gene (PRX2F) it was deduced that *G. gracilis* most likely responded physiologically to the elicitors, i.e. an initial oxidative burst in the first hour of exposure followed by a stronger, prolonged oxidative burst between 8 and 24 hours. If the *G. gracilis* defence response moved into the final phase, a HR and PCD would have been expected to occur together with the activation of proteases such as nucleases and caspases.

Caspases are cysteine proteases that cleave their target proteins at specific aspartate residues (Buckner *et al*., 2000). They are fundamentally different from serine proteases which cleave serine residues in target proteins. However, caspase-like independent PCD pathways may also exist in plants. For example, the inhibition of plant PCD by class-specific protease inhibitors of cysteine-protease (non-caspase) and serine-protease suggest a possible role for non-caspase proteases in PCD (Coffeen and Wolpert, 2004; Woltering *et al*., 2002; Beers *et al*., 2000). In several plants, serine proteases of the subtilisin family are able to cleave caspase-specific substrates (Beers *et al*., 2000). Although scientific evidence supporting the role of caspase-like proteases in plant PCD is growing, there is still no direct evidence for these proteases playing a central role in PCD (Rotari *et al*., 2005; Beers *et al*., 2000). Studies like Tornero *et al*. (1997, 1996) evaluated the response of tomato plants to viral infection. The authors observed an induction of the HR and increased transcription of pathogenesis-related genes. Furthermore, the authors reported that PR-P69 expression was induced by other pathogens such as fungi and nematodes in addition to ethylene and salicylic acid which are molecules known to mediate defence responses in plants. In another study, Yano *et al*. (1999) proved that serine proteases were required in signal transduction pathways and for the activation of a HR and PCD in response to pathogen infection in tobacco cells. Xylanase from *Trichoderma viridae* induced an oxidative burst which was prevented by serine-protease inhibitors. In the context of an activated defence response in *G. gracilis*, there is no evidence to suggest that the putative serine protease-like gene identified in this study is involved in PCD, but because transcriptional regulation in response to disease elicitors was observed, the possibility should be explored further. Alternatively, this particular serine protease-like gene (and protein) could be required in signal transduction pathways or in the indirect regulation and cleavage of other serine-containing proteins more directly involved in defence response. Characterization of the translation regulation of this serine protease-like protein may offer more insight into its role in activated defence in *G. gracilis*.

Transcriptional regulation of a cytosolic thioredoxin from *G. gracilis* suggested that exposure to disease elicitors may have resulted in gene repression during the 24 hour time-course experiment. Similarly, due to the failure to detect statistically significant changes in gene expression, several more biological repeats are required to clarify transcriptional regulation of this particular thioredoxin upon exposure to disease elicitors. As previously discussed in

Chapter 2, cytosolic (type *h*) thioredoxins posses various physiological roles. For example, thioredoxins have been implicated in the regulation of certain enzymes, modulation of transcription, cellular development, act as possible hydrogen donors for other biosynthetic enzymes or are required for the protection of proteins from oxidative damage. If the *G. gracilis* cytoplasmic thioredoxin characterized in the current study was directly involved in oxidative protection, a gene expression pattern similar to that of PRX2F might have been expected. However, complicating this conjecture is the fact that several isoforms typically exist in an organism. Therefore, in order to clarify the physiological roles of cytoplasmic thioredoxins, the number of isoforms present in *G. gracilis* would have to be established. An experimental approach similar to that of Rivera-Madrid *et al*. (1995) could be attempted. The authors isolated five different clones encoding thioredoxin homologues from *A. thaliana* cDNA libraries. Southern blots were used to establish the number of genes present in the genome as well as the abundance or presence of thioredoxin-related sequences. The complexity of the thioredoxin system is maintained in higher plants (Gelhaye *et al*., 2005; Baumann and Juttner, 2002) and it is possible that the thioredoxin system is equally complex in *G. gracilis*. The pattern of thioredoxin gene expression in this qPCR study strongly suggested that it was an isoform that lacked significant antioxidant activity. If it did possess antioxidant activity, the assumed oxidative burst might have transcriptionally up-regulated gene expression particularly in the seaweed samples exposed to the disease elicitors which was not the case. To actually establish a physiological role for this specific cytosolic thioredoxin, it would be interesting to determine its *in vivo* target protein-complexes during the activated defence response. Yamazaki *et al*., (2004) identified the target proteins of an *A. thaliana* cytoplasmic thioredoxin which included proteins which displayed antioxidant activity (ascorbate peroxidase, germin-like protein and a monomeric type II peroxiredoxin). The transcriptional regulation of this thioredoxin could therefore possibly be linked to the PRX2F identified in this study. Another study (Verdoucq *et al*., 1999) demonstrated a link between thioredoxins and peroxiredoxins. When two thioredoxin genes were disrupted in the yeast *S. cerevisiae*, cell viability and growth were severely compromised. Expression of an *A. thaliana* type *h* thioredoxin (AtTRX3) in the thioredoxin-deficient yeast strain restored the wild-type phenotype to normal cell growth, improved its ability to grow on methionine sulphoxide and enhanced its H_2O_2 tolerance. Other protein targets of cytosolic thioredoxin include proteins involved in polypeptide biosynthesis (elongation factor-II and eukaryotic translation initiation factor 4A), protein degradation (the regulatory sub-unit of 26S

proteasome) and metabolism (alcohol dehydrogenase, fructose 1,6-bis phosphate aldolase-like protein, cytosolic glyceraldehydes 3-phosphate dehydrogenase, cytosolic malate dehydrogenase and vitamin B_{12}-independent methionine synthase). In addition, several target proteins were localized to the chloroplast (chaperonin 60, heat shock protein 70 and glutamine synthase). It is therefore possible that transcriptional regulation observed for this *G. gracilis* cytosolic thioredoxin could have indirectly impacted any of these potential target proteins which might play more direct roles in the activated defence response. The qPCR result for this thioredoxin has therefore raised more questions than answers which would be interesting to investigate further.

3.5 CONCLUSION

Macroalgae encounter a wide variety of biotic stresses and have evolved a variety of adaptive mechanisms. Disease tolerance and resistance is typically enhanced through physical adaptations as well through inducible physiological changes at the molecular and cellular level once a disease elicitor is detected. Following elicitor detection, a molecular response must be relayed via signal transduction pathways which eventually lead to changes such as defence gene expression or synthesis of anti-pathogenic chemical compounds. The macroalgal model of pathogen perception and activated defence has been shown to exhibit several similarities to vascular plants and other higher organisms. In this context, this chapter proved that the red macroalga *G. gracilis* (most likely) responded with an oxidative burst upon perception of an endogenous elicitor. Furthermore, the transcriptional changes observed had to be mediated by signal transduction pathways. The most significant changes in gene expression occurred during the first 8 to 12 hours post exposure to disease elicitors. The 30 min time-course experiment failed to reveal any significant changes in gene expression for the genes assessed. During the 24 hour time-course experiment however, transcription of the antioxidant gene, PRX2F, was significantly induced while thioredoxin generally displayed no transcriptional regulation in response to the elicitor. Phosphoserine phosphatase was transcriptionally repressed in response to exposure to disease elicitors which highlighted the importance of phosphorylation events during activated defence. Essentially, the transcriptional repression of phosphoserine phosphatase was a logical biological response

because phosphatases reverse the effect of kinases which are known to phosphorylate and activate various defence proteins. The significant transcriptional down-regulation of PSD, a gene responsible for the bio-conversion of PtdSer was an interesting observation. Literature suggests that PtdSer is required to activate PKC or other kinases involved in MAPK signalling cascades. Furthermore, PtdSer is directly involved in the initiation of PCD via its deposition into plasma membranes. Thus repression of the gene that is responsible for the bio-conversion of PtdSer would have theoretically resulted in a higher concentration of PtdSer in the cell and thus implicates the HR and PCD in the activated defence response of *G. gracilis*. The putative serine protease-like gene exhibited transcriptional regulation in response to disease elicitors as well but its physiological role in the defence response of *G. gracilis* remains unclear although scientific evidence suggests possible roles in PCD, signalling pathways or recycling and processing of serine-containing proteins. Characterization of the transcriptional regulation of these five putative defence genes was only the first step in understanding the presumably highly complex activated defence response of *G. gracilis*. Levels of the functional unit (protein product) required to effect a physiological change may not always positively correlate with the transcriptional changes observed for its gene. Consequently, the next chapter will attempt to establish whether there is a correlation between the transcriptional and translational regulation of two putative defence genes in *G. gracilis*.

CHAPTER 4

Investigation of translational regulation of two putative defence genes using western hybridization analysis

4.1 INTRODUCTION

Gene expression studies at the transcriptional level cannot be considered independently from translational regulation since proteins are the functional units in any given cell. The presence of protein product does not guarantee biological function but may be a better indication of activity than mRNA transcript detection. Consequently, analyses at both levels are required in order to elucidate the metabolic and regulatory processes in living organisms (Nie *et al*., 2006). The central dogma of molecular biology states that the information encoded in DNA is transferred to protein via mRNA molecules. However, comparative studies between mRNA and protein abundance suggest that correlation between the two is generally poor and that mRNA levels alone cannot be used to reliably predict protein abundance (Maier *et al*., 2009; Cox *et al*., 2005; Chen *et al*., 2002). This poor correlation has been ascribed to various biological variables in addition to the limitations associated with the molecular tools used to quantitate mRNA and protein molecules. These include northern blot analysis, qPCR and microarrays at the mRNA level, and 2D SDS-PAGE, mass spectrometry and western blot analysis at the protein level. Variation due to biological factors is primarily due to the fundamental differences along the path of gene transcription and translation into protein (Figure 4.1).

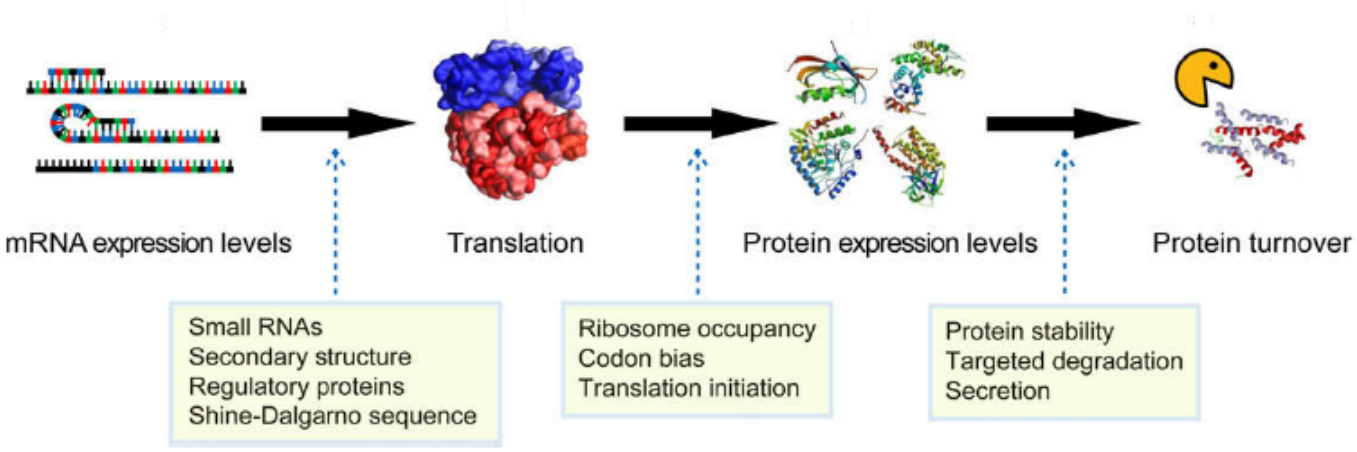


Figure 4.1 Central pathway of molecular biology depicting the biological parameters that influence the mRNA-protein correlation (Maier *et al*., 2009).

In Chapter 3, the transcriptional response of five putative defence genes of *G. gracilis* to disease elicitors was assessed. Gene transcription patterns cannot be assumed to positively correlate with their protein products. Thus, the aim of this chapter was to establish this relationship for two putative defence genes. The mRNA transcript for a peroxiredoxin-like (PRX2F) gene was shown to be significantly up-regulated in response to disease elicitors over a 24 hour period. The predicted protein product showed significant homology to peroxiredoxin and other antioxidant proteins when subjected to BLAST analysis. Previous studies have shown that *G. gracilis* responds with the release of reactive oxygen species (ROS) upon exposure to the breakdown products of its cell wall (Weinberger *et al*., 2010; 1999). As a result, increased levels of antioxidant proteins can be expected in *G. gracilis* samples exposed to disease elicitors in order to cope with and prevent oxidative damage. A mRNA transcript for a putative serine protease-like gene appeared to be repressed following exposure to disease elicitors (bearing in mind that statistical analysis failed to detect statistically significant changes in gene expression due to the low power of the tests and a possible type II error). Serine proteases have been implicated in defence signalling, programmed cell death and even recycling or regulation of other proteins more directly involved in the activated defence response of other organisms. Therefore, it was essential to

further assess the translational regulation of this serine protease-like gene in *G. gracilis* samples exposed to disease elicitors.

As highlighted in Chapter 1, genomic and proteomic information is lacking for non-model organisms such as *G. gracilis*. In such instances, proteomic studies involving western hybridization analyses are often compromised because commercially available antibodies yield low cross-reactivity with *G. gracilis* proteins. Although the PRX2F and serine protease-like protein sequences of *G. gracilis* could be aligned to those from fully-annotated genomes, sequence homology was not high enough to warrant the purchase of commercially available antibodies. An alternative to this problem was to use a protein expression system in which the coding region of the *G. gracilis* gene was cloned and expressed in *Escherichia coli*. Recombinant proteins were subsequently purified and used to raise highly specific polyclonal antibodies. This approach has been used successfully before in *G. gracilis* (Schroeder *et al*., 2003) and other non-model plant species such as *Xerophyta viscosa* (Maredza, 2007; Mowla, 2005).

In this chapter, the method employed to clone and express *G. gracilis* genes in an *E. coli* expression system will be discussed. To achieve this objective, the entire coding region for the *G. gracilis* PRX2F and serine protease-like genes had to be established. Cloning of the complete open reading frames of both *G. gracilis* genes ensured that the expressed recombinant proteins were not truncated, and thus, effectively represent native, functional *G. gracilis* proteins. Recombinant proteins were purified and used to raise polyclonal antibodies in rabbits. These polyclonal antibodies were subsequently used for quantitative assessment of each *G. gracilis* protein under two experimental conditions (exposure / no exposure to disease elicitors) over a 24 hour time-course experiment by means of western hybridization analysis. By assessing the amount of protein present in each *G. gracilis* sample, the correlation between mRNA and protein expression was established for both genes.

4.2 MATERIALS AND METHODS

4.2.1 Determination of full-length gene sequences

DNA sequences corresponding to the PRX2F and putative serine protease-like genes were previously determined (Chapter 2; Section 2.2.8; Appendix F.1). The SMART (Switching Mechanism at 5' end of RNA transcript) protocol, used to construct cDNA stress-libraries, was designed to preferentially enrich for full-length cDNAs by eliminating the need for T4 DNA polymerase. Conventional cDNA cloning procedures use the T4 DNA polymerase in combination with adaptor ligation to generate blunt cDNA ends after second-strand synthesis. Consequently, synthesized cDNA populations tend to lack the 5' ends of genes when compared to the original mRNA transcripts (Creator^{TM} SMART^{TM} cDNA library construction kit, Clonetech) (Appendix D.2).

Therefore, because the Creator^{TM} SMART^{TM} cDNA library construction kit essentially eliminated the potential for missing 5' ends in cDNA preparations, it was assumed that the full-lengths of both genes investigated in this study had been obtained, i.e. any truncated cDNAs lacking the 5' primer binding site would not have been converted to double-stranded cDNA and included in the cDNA library. However, to confirm that full length genes were present, the predicted open reading frame (ORF) of each DNA sequence was determined using DNAMAN (Version 4.13; Lynnon Biosoft). ORFs were subjected to protein homology searches using the BLASTp algorithms provided by NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and Uniprot (http://www.uniprot.org/) and subsequently aligned to the full-length ORFs of the ten most homologous sequences using ClustalW (http://www.uniprot.org/). Only four proteins with identities of "serine protease-like" were selected for alignment to the ORF that corresponded to the *G. gracilis* serine protease-like protein. In addition, both *G. gracilis* protein sequences were assessed for the presence of conserved protein domains to confirm full-length gene sequences.

4.2.2 PCR amplification of gene open reading frames

The *G. gracilis* PRX2F and serine protease-like coding regions were PCR-amplified from the original cDNA clones used to construct the microarray slide. Since the expression vector pET-29a (+) (Novagen, USA) was selected for gene cloning (Appendix D.3). PCR primers were designed to incorporate a 5' up-stream S-tag and a 3' down-stream His-tag into the expressed recombinant protein (Table 4.1). Forward primers were designed with a *Bam*HI restriction endonuclease site which directly replaced the start (ATG) codon of each putative defence gene. Reverse primers were designed with a *Hind*III restriction endonuclease site and directly replaced the stop codon of each putative defence gene. Prior to PCR amplification, the original cDNA clone (in a pDBR-LIB plasmid vector; Appendix D.1) was linearized using 1U of FastDigest *Bam*HI (Fermentas Life Sciences, Germany) as per the manufacturer's instructions. FastAP Thermosensitive Alkaline Phosphatase (Fermentas Life Sciences, Germany) was included in the reaction to prevent re-circularization of plasmid DNA.

Table 4.1 PCR primers used for amplification and cloning of the *G. gracilis* genes into the pET-29a (+) expression vector. Underlined, italicized sequences represent *Bam*HI and *Hind*III restriction sites in the forward and reverse primers, respectively.

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	Optimal Tm
PRX2F	GCG***GGATCC***ACTCGCTATGCCATTGC	GGC***AAGCTT***TCCCAATGGTG	60°C
Serine protease-like	GAG***GGATCC***ACTGTCTCTACCG	CGC***AAGCTT***CTGAGCAAGAGTC	59°C

The coding region of each gene was amplified in a Bioer XP thermal cycler (Separation Scientific) according to the following cycling parameters: 95°C for 3 min; 35 cycles of 95°C for 1 min, 60°C for 30 s, 72°C for 2 min and a final extension step of 7 min at 72°C. Each PCR reaction contained 1X $MgSO_4$ buffer, 0.2 mM dNTPs, 0.2 µM of each primer, 1.25 U of *Pfu* DNA polymerase (Fermentas Life Sciences, Germany), 4 mM $MgSO_4$ and 5 ng of linearized plasmid DNA in a final volume of 50 µl. PCR products were electrophoresed on a 1% (w/v) TAE (Appendix A.2) agarose gel and visually assessed for the amplification of a single band. PCR amplified cDNA was subsequently gel-extracted using a Biospin Gel extraction kit (Bioflux) according to the manufacturer's instructions and quantitated using a Nanodrop spectrophotometer (Thermo Scientific). All DNA was stored in TE buffer (pH 7.6) at 4°C until use.

4.2.3 Cloning of open reading frames into protein expression vector

One microgram of each gel-extracted PCR product was digested with FastDigest restriction enzymes *Bam*HI and *Hind*III (Fermentas Life Sciences, Germany) according to the manufacturer's instructions. Restriction enzymes were heat-inactivated for 10 min at 80°C followed by PCR purification using an E.Z.N.A® Cycle-Pure Kit (peQLab Biotechnologie GmbH) as per the manufacturer's instructions. Digested PCR products were re-suspended in TE buffer (pH 7.6), quantitated using a Nanodrop spectrophotometer (Thermo Scientific) and stored at 4°C. Similarly, 5 µg of the pET29a (+) expression vector was digested with both restriction enzymes before cloning. The digested PCR fragments and expression vector were ligated overnight at 4°C using a standard ligation protocol and T4 DNA ligase (Fermentas Life Sciences, Germany). Five microlitres of the ligation mixture was used to transform *E. coli* (DH5α) according to the method of Dagert and Ehrlich (1979). Bacterial cultures were grown overnight at 37°C on LA supplemented with 30 $\mu g.ml^{-1}$ Kanamycin (Kan) (Appendix A) to select positive transformants. Single colonies were used to inoculate 5 ml Luria broth (Appendix A.1.2) supplemented with 30 $\mu g.ml^{-1}$ Kan. Bacterial cultures were incubated overnight at 37^oC on a shaker at 100 rpm. Plasmid DNA was extracted according to Sambrook *et al*. (1989) (Appendix C.1), re-suspended in 50 µl sterile distilled water and quantitated with a Nanodrop spectrophotometer. Additionally, DNA samples were electrophoresed on a 1.2 % TAE agarose gel to verify plasmid quality and integrity. To

confirm that transformants were positive and cloned in the correct frame, three screening approaches were used: (i) the entire recombinant DNA fragments (S-tag + ORF + His-tag) were amplified using established cycling parameters (Section 4.2.2) with a few modifications: forward and reverse primers used were designed to the S-tag [5' – CGAACGCCAGCACATGGACAGC – 3'] and T7 terminator [5' – GCTAGTTATTGCTCAGCGG – 3'] regions, respectively; (ii) plasmid DNA was digested using restriction enzymes *Bam*HI and *Hind*III and electrophoresed on a 1% TAE agarose gel; and (iii) plasmid DNA corresponding to the positive transformants was sequenced in both directions using the S-tag and T7 terminator primers on an ABI 3730 XL DNA sequencer (Macrogen Inc., South Korea). Upon successful screening in all three approaches, plasmid DNA (pET29a:*G. gracilis* ORF) was used to transform a protein expression host (*E. coli* BL21) as above. Similarly, positive transformants were selected and screened using PCR.

4.2.4 Recombinant protein expression

Before small-scale recombinant protein expression trials were performed, the solubilities of both recombinant proteins in the *E. coli* expression host were predicted based on their amino acid sequences using the bioinformatics tools available at ProtParam (http://au.expasy.org/tools/protparam.html) and Recombinant Protein Solubility (http://www.biotech.ou.edu). This was required to devise an appropriate purification strategy for each recombinant protein. Small scale trial expression experiments (20 ml LB supplemented with 30 µg.ml^{-1} Kan) were then conducted to establish whether the recombinant proteins could be successfully expressed in *E. coli* BL21 cells. Bacterial cultures were grown to an optical density (600 nm) of 0.6 – 0.8 and recombinant protein expression was induced by adding IPTG (isopropyl-β-D-thiogalactopyranoside) to a final concentration of 1 mM to each flask. Samples (1 ml) of each bacterial culture were initially removed before adding IPTG (pre-induction of protein expression) and subsequently at 1, 2, 3, 4, 5, 6 and 24 hours after recombinant protein expression was induced. Samples were then centrifuged at 12 000 rpm for 5 min in order to separate the bacterial cells from the culture supernatant. Bacterial pellets were stored at –20°C until use. In addition to inducing recombinant protein expression via the addition of IPTG to the culture media, an auto-

inducing culture medium (Overnight Express Media ™, Novagen) which eliminated IPTG for protein expression was also tested.

In order to determine whether the recombinant proteins were expressed in the soluble or insoluble protein fractions in *E. coli*, bacterial pellets were re-suspended in 500 µl Lysis Buffer (Appendix A.3.4.1.1) and subjected to four freeze-thaw cycles in liquid nitrogen. Samples were then centrifuged at 12 000 rpm for 10 min. The resultant supernatant and pellet were separated and represented soluble and insoluble protein fractions, respectively. One hundred microlitres of 5X Sample Application Buffer (Appendix A.3.4.3) was added to each fraction and boiled for 5 min. The actual solubility of each recombinant protein was compared to its predicted solubility by assessing crude extracts of the soluble and insoluble fractions of total *E. coli* protein on SDS-PAGE (Laemlli, 1970). In addition to establishing whether the recombinant proteins were soluble or insoluble in *E. coli*, these trial experiments were used to determine the experimental parameters for optimal expression of recombinant protein.

After demonstrating successful expression of recombinant proteins in the small-scale study, expression was scaled up to maximize the amount of recombinant protein available for purification. Unfortunately the experimental parameters that proved optimal for the small-scale volume (20 ml culture) proved to be ineffective when the culture volume was scaled up to 500 ml. As a result, experimental parameters (incubation temperature, IPTG concentration and duration of recombinant protein expression) had to be re-optimised for the large-scale study. For expression of recombinant PRX2F, *E. coli* BL21 cells containing the expression vector were grown at 37°C in 500 ml LB supplemented with 30 µg.ml^{-1} Kan to an optical density of 0.6 – 0.8. Protein expression was induced by adding IPTG to a final concentration of 1 mM. The flask was then transferred to 25°C and incubated for a further 22 hours. Expression of the recombinant putative serine protease-like protein was performed using the same parameters except that the optimal temperature and duration of recombinant protein expression was 20°C and 23.5 hours, respectively. Bacterial cells were harvested by centrifugation at 8 000 rpm for 10 min at 4°C and subsequently stored at –20 °C. Prior to recombinant protein purification, the solubility of each recombinant protein was determined as per the protocol described for the small-scale expression study.

4.2.5 Affinity purification of recombinant proteins

All buffers used for purification were prepared on the day and de-gassed by filtration through a 0.22 µM filter under vacuum. Since each recombinant protein contained a 6X histidine-tag, affinity purification using immobilized metal-ion affinity chromatography was performed using a 5 ml HisTrap™ HP ready-to-use column, pre-packed with precharged Ni^{+} sepharose (GE Healthcare) linked to a HPLC system (Gilson 321 gradient pump, Gilson FC204 fraction collector and a Waters 484 UV detector all operated by UniPoint^{TM} Systems software). Purification of recombinant PRX2F took place under native conditions (Appendix A.3.4.1) as the expressed protein was soluble in *E. coli*. Briefly, bacterial cells were thawed on ice for 15 min and then re-suspended in 2 – 5 ml lysis buffer (Appendix A.3.4.1.1) per gram wet weight. Lysozyme was added to a final concentration of 1 $mg.ml^{-1}$ and the bacterial cells were incubated for a further 30 min on ice. The cell lysate was sonicated for 4 min at 300 W with a 15 s cooling delay between bursts followed by centrifugation at 10 000 rpm for 30 min at 4°C to pellet cellular debris.

Purification of recombinant serine protease like protein took place under denaturing conditions using 8 M urea (Appendix A.3.4.2) since it was expressed as an insoluble protein in *E. coli*. Briefly, bacterial cells were thawed on ice for 15 min. Prior to solubilisation in lysis buffer (A.3.4.2.1), the pellet was washed in Triton-X to solubilise the cell membranes as follows: three times in 1% Triton-X with a final wash in 0.1% Triton-X. Between washes, the bacterial pellets were centrifuged at 8 000 rpm for 15 min and the supernatant was discarded. The bacterial cells were then lysed by gentle vortexing, taking care to avoid foaming, for 1 hour at room temperature. Cell lysates were centrifuged at 10 000 rpm for 30 min at room temperature to pellet cellular debris.

The cleared lysates for both recombinant proteins were subsequently filter-sterilized using a 0.22 µM filter prior to purification in order to remove cellular debris and to prevent blockage of the Ni^{+} column. Recombinant proteins bound to the Ni^{+} column under native or denaturing conditions were washed to remove non-specific binding of *E. coli* proteins using the appropriate wash buffers and eluted using a gradient of imidazole (20 mm – 500 mM) (Appendix A.3.4). Eluted fractions were collected over a 30 min period while the Waters 484 UV detector simultaneously measured optical density (595 nm) and voltage (mV) as elution

off the column progressed. Eluted protein fractions with the highest optical densities were analyzed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to determine purity as follows: each fraction was mixed with an appropriate amount of 5X Sample Application Buffer and boiled for 5 min followed by SDS-PAGE. The purest fractions (as determined by the presence of a single protein band) were pooled, concentrated and de-salted using Amicon Ultra centrifugal filter columns (Millipore Corporation, U.S.A) with 10-kDa molecular weight size exclusion according to the manufacturer's instructions. Recombinant proteins were subsequently re-suspended in phosphate buffered saline (PBS) pH 7.4 (Appendix A.2.4) and quantitated using a Nanodrop spectrophotometer (Thermo Scientific). Recombinant proteins were stored as aliquots of 500 µg at –70°C until antibodies could be raised in rabbits. Prior to immunization, the purity of each recombinant protein was confirmed using western hybridization. Immunodetection was performed with polyclonal anti-histidine antibodies (GeneTex, Inc.) which cross-reacted specifically with the C-terminal poly-histidine tag fused to the recombinant proteins.

4.2.6 Antibody production and determination of antibody titre

Animal ethics clearance was obtained for this study (number 008/043). Two New Zealand rabbits were immunized with 500 µg of each recombinant protein together with Freund's incomplete adjuvant to raise polyclonal antibodies according to the protocol by Rybicki (1979) (performed by University of Cape Town, Animal Unit). Each animal was initially bled to obtain pre-immune sera and then injected with 1 ml of antigen emulsion sample on days 1, 8, 15 and 29. Bleeding took place on days 22 and every two weeks thereafter (days 37, 51 and 65) until a high titre antiserum was obtained. In order to test the pre-bleed and first four bleeds for the presence of anti-PRX2F and anti-serine protease-like antibodies, serial dilutions of purified recombinant protein were 'dot-blotted' in duplicate onto nitrocellulose membranes which were subsequently air-dried. Western hybridization analysis was performed as follows: To prevent non-specific binding of antisera, the membranes were blocked by immersion in 100 ml blocking buffer (Appendix A.3.5.1) for 2 hours at room temperature. The bleeds were diluted 1:100 in blocking buffer and each membrane was incubated separately in each bleed for 1 hour at room temperature. The membranes were washed four times in 100 ml 1X TBS-T (Appendix A.3.5.3) for 15 min each time. The

secondary antibody (peroxidase labelled goat anti-rabbit IgG (H + L) liquid conjugate at 1.0 $mg.ml^{-1}$ (KPL; Kirkegaard and Perry Laboratories, Inc.) was diluted 1:5000 in blocking buffer, added to the membranes and incubated for 1 hour at room temperature. The membranes were washed four times in 100 ml TBS-T as described previously. Freshly prepared (500 µl) developing substrate solution (TMB Membrane Peroxidase substrate; KPL) was added to each membrane. Colour development was allowed to proceed for a few minutes. Reactions were stopped simultaneously by washing each membrane in water.

4.2.7 Purification of antibodies by polyethylene glycol precipitation

Polyclonal antibodies were purified using a method adapted from Polson *et al*. (1964). One volume of antiserum was mixed with 2 volumes of borate buffered saline pH 8.6 (BBS; 35 mM Boric acid, 37.5 mM NaCl). Crushed polyethylene glycol (PEG) 6000 [14% (w/v)] was added to the antiserum and dissolved by gentle inversion. The solution was centrifuged at 12 000 rpm for 10 min at 4°C. The pellet was re-dissolved in the original volume of antiserum and 14% (w/v) PEG 6000 was again added and dissolved. The antibodies were pelleted as before and dissolved in half the original antiserum volume using PBS containing 60% glycerol. Purified polyclonal antibodies were stored in aliquots at –20°C.

4.2.8 Seaweed sample acquisition for exposure to disease elicitors

G. gracilis samples used to isolate protein for western hybridization analyses were those prepared in Chapter 3 (Section 3.2.1). Three biological repeats in total were tested. The experimental flasks contained seaweed samples grown in media supplemented with disease elicitors. Samples were removed at 1, 8, 12 and 24 hours post addition of the disease elicitor. Controls flasks contained seaweed samples grown in media lacking disease elicitors. As per the experimental flasks, samples were removed at 0, 1, 8, 12 and 24 hours.

4.2.9 Isolation of seaweed total protein

Seaweed (0.5 g) from each time-point along the 24 hour time-course experiment was ground in a cooled, sterile pestle and mortar using liquid nitrogen. Ground material was not allowed to thaw and placed in 1 ml of extraction buffer (0.5 M Tris pH 7.5; 10 mM EDTA; 1 % Triton-X and 2 % β mercaptoethanol). Samples were vortexed thoroughly and subsequently centrifuged for 5 min at 12000 rpm. Each supernatant was transferred into a sterile 2 ml microfuge tube and an equal volume of Tris-saturated phenol (pH 8) was added. Samples were vortexed vigorously for 10 min followed by centrifugation at 12000 rpm for 1 min. The aqueous phase (clear top layer) was removed and the protein phase (phenol phase) was re-extracted with an equal volume of extraction buffer. Samples were again vortexed followed by centrifugation at 12000 rpm for 1 min. The aqueous phase was discarded and the phenol phase (brown when isolating protein from *G. gracilis*) was split into three equal volumes to which 5 volumes of 0.1 M ammonium acetate in methanol was added. Protein was allowed to precipitate overnight at –20°C. Following precipitation, samples were centrifuged for 5 min at 12000 rpm and concentrated in one sterile microfuge tube. The protein pellet was washed with 500 µl 0.1 M ammonium acetate in methanol and then with 80 % ice-cold acetone. The pellet was air-dried for 10 min and then re-suspended in urea lysis buffer (ULB) (Appendix A.3.6.1). Protein concentrations were calculated using Bradford's protocol (1976), using the Bio-Rad Protein Assay Dye Reagent (Bio-Rad, Germany) with one modification: 0.1M HCl was included in each quantitation reaction to counteract the presence of urea. Dilutions of bovine serum albumin standard (Pierce, U.S.A) were used to construct a protein standard curve. Protein quality and integrity were assessed using SDS-PAGE while the remainder of the samples were stored at –70°C until western hybridization analyses could be performed.

4.2.10 Determination of antibody specificity and optimal amount of *G. gracilis* protein for western hybridization analyses

A decreasing range (50, 40, 30, 20 and 10 µg) of total protein extracted from the time 0 control sample was separated according to size by SDS-PAGE. A range of total protein was tested to determine the specificity of each polyclonal antibody and the amount of *G. gracilis*

protein required to prevent saturation of signal in western hybridization analysis. Protein electrophoresis was performed using a Mini-PROTEAN Tetra cell system (Bio-Rad. Germany) set to a 100V constant current until the dye front was eluted. SDS-PAGE gels were equilibrated in Towbin buffer (Appendix A.3.5.4) following electrophoresis. The separated proteins were transferred onto nitrocellulose transfer membrane (S&S Protan, PerkinElmer Life Sciences) using a Mini Trans-Blot cell (Bio-Rad, Germany) for 1 hour at 100V. Western hybridization analysis was performed as described in Section 4.2.6 with a few modifications: The highest titre of purified polyclonal antibodies (1°) (as determined by dot blot analyses) was diluted 1:1000 in blocking buffer without pre-absorption against *E. coli* BL21 total protein extract. The secondary antibody (2°) was diluted 1:50000. *G. gracilis* proteins were detected using the Immun-Star™ WesternC™ Chemiluminescent Kit (Bio-Rad, Germany) as per the manufacturer's instructions. Chemiluminescent signals were visualised using the Molecular Imager ChemiDoc XRS+ system (Bio-Rad, Germany).

4.2.11 SDS-PAGE and western hybridization analysis of the 24 hour time-course experiment

SDS-PAGE

Total protein extracted from each *G. gracilis* sample (time 0 in addition to the experimental and control samples at 1, 8, 12 and 24 hours) was separated according to size by SDS-PAGE (in duplicate). Stacking gels were prepared as follows: 4 % (w/v) acrylamide, Tris-Cl/SDS pH 6.8, 0.05 % (w/v) ammonium persulphate (AMPS), 0.01 % (v/v) N,N,N',N'-Tetramethylethylenediamine (TEMED) while resolving gels were prepared as follows: 10 % (w/v) acrylamide, Tris-Cl/SDS pH 8.8, 0.2 % (w/v) ammonium persulphate (AMPS), 0.04 % (v/v) TEMED. Acrylamide solution [40 % (w/v) (Sigma)] was used as the acrylamide stock solution. Tris-Cl/SDS buffer composition can be found in Appendix A.3.5. All protein samples for SDS-PAGE were prepared by mixing 5 X SDS application buffer (0.25 M Tris-HCl pH 6.8, 50 % glycerol, 10 % SDS, 0.05 % bromophenol blue and 0.5 M DTT) with the appropriate volumes of protein. The optimal amount (Section 4.2.10) of protein sample was loaded equally, as determined by the Bradford assay. Protein samples were electrophoresed at 100 V until the dye front was eluted. One protein gel was stained with Coomassie solution

[0.05 % (w/v) Coomassie Brilliant Blue R-250; 50 % methanol; 10 % glacial acetic acid; 40 % water] while the other was used for western hybridization analysis.

Western hybridization analyses

Following SDS-PAGE, the second protein gel was equilibrated in Towbin buffer. The separated proteins were then transferred onto nitrocellulose transfer membranes (S&S Protan, PerkinElmer Life Sciences) using a Mini Trans-Blot cell (Bio-Rad, Germany) for 1 hour at 100V. Western hybridization analyses were performed as described in Section 4.2.6 using optimal antibody concentrations (1° diluted 1:1000 and 2° diluted 1:50000). Primary antibodies were pre-absorbed against *E. coli* BL21 total protein extract to minimize non-specific hybridization. The *G. gracilis* PRX2F and serine protease-like proteins were detected using the Immun-Star^{TM} WesternC^{TM} Chemiluminescent Kit (Bio-Rad, Germany) as per the manufacturer's instructions. Chemiluminescent signals were visualised using the Molecular Imager ChemiDoc XRS+ system (Bio-Rad, Germany). Immunoblots were exposed for 120 s with one image captured every 10 seconds to prevent over-exposure of the signal.

Equal loading of protein was visually assessed by staining the nitrocellulose membranes with Ponceau S solution [0.5 % (w/v) Ponceau S stain (Sigma-Aldrich, Germany) and 1 % (v/v) glacial acetic acid] and experimentally verified by re-probing each membrane with an affinity purified polyclonal anti-rubisco antibody (1 $\mu g.\mu l^{-1}$; Agrisera) diluted 1:10000 in blocking buffer. The concentration of the rubisco protein was not expected to change in any of the *G. gracilis* total protein samples across the 24 hour time-course. Prior to re-probing, the membranes were stripped at 50°C for 30 min with agitation in a stripping buffer (0.2 % β-mercaptoethanol, 2 % SDS and 62.5 mM Tris-HCl pH 6.8) followed by five rinses in 1X TBS-T. Chemiluminescent detection and visualisation of the rubisco protein was performed as before. For all SDS-PAGE and western hybridization analyses, a pre-stained protein ladder (170 – 10 kDa; Fermentas Life Sciences, Germany) was used to size positive signals. Western hybridization analyses of the PRX2F and serine protease-like proteins were performed separately, i.e. one membrane was analysed for either protein, stripped and then re-probed for the loading control rubisco protein.

4.2.12 Densitometry analysis of western hybridization immunoblots

Quantity One Software [Version 4.5.2 (Bio-Rad, Germany)] operating a ChemiDoc XR molecular imager [Version 4.6.9 (Bio-Rad, Germany)] was used to determine the density of each chemiluminescent signal. Only the band of the expected molecular weight was assessed. The volumetric analysis tool was selected to calculate the density of each signal as it was able to map the contour of the entire signal as opposed to using a rectangle of fixed dimensions. This tool was specifically chosen because in certain cases, electrophoresis resulted in warping of the SDS-PAGE gels, resulting in anomalies in the shape of the protein bands following transfer to nitrocellulose membranes.

4.2.13 Statistical analyses of western hybridization analysis

Densitometry readings corresponded to the relative amount of a specific *G. gracilis* protein under the two treatments in the 24 hour time-course experiment. The aim of these western hybridization experiments was to determine whether PRX2F and serine protease-like protein expression changed significantly as a result of exposure to disease elicitors. Since a control sample was included at each time point along the 24 hour time-course, a student's *t*-test comparing protein expression was deemed appropriate (as applied in Chapter 3, Section 3.2.9 to statistically assess changes in gene expression). In contrast, a one-way ANOVA statistical test was selected to assess levels of the rubisco protein. Rubisco was not expected to change under experimental or control conditions over the 24 hour time-course and was therefore implemented as protein loading control. Data sets that were not normally distributed or that displayed unequal variances were transformed to meet the requirements for parametric statistical testing. Where data could not be adequately transformed, a non-parametric test was used instead. All statistical analyses were performed using SigmaStat for Windows (Version 3.10.0).

4.3 RESULTS

4.3.1 Determination of full-length genes

DNA sequences corresponding to the cDNA clones (originally used to construct the microarray slides in Chapter 1) which harboured the PRX2F and serine protease-like genes were assessed using DNAMAN software (Version 4.13). The open reading frames (ORFs) corresponding to the PRX2F and serine protease-like genes were subjected to BLAST analyses against protein databases in order to align each ORF to full-length sequences of the most homologous proteins in other organisms so that the presence of the 5' and 3' sections in each *G. gracilis* gene could be confirmed.

Several other ORFs were predicted within the cDNA sequence harbouring the PRX2F gene (ORF of 675 nucleotides encoding a 225 amino acid polypeptide of molecular weight 24.3 kDa (Frame +2)) (Figure 4.2). The ORF predicted in reading frame – 1 was approximately 650 nucleotides long and displayed sequence homology to an *A. thaliana* transferase family protein with an insignificant E-value (0.041). As a result, this ORF was disregarded for further analysis. Multiple sequence alignments were set up to compare 10 other complete PRX2F proteins to the *G. gracilis* PRX2F protein (Figure 4.3). Based on the protein alignments, the cDNA clone contained an ORF which represented a full-length PRX2F gene. In terms of size comparison, the *G. gracilis* PRX2F protein (225 amino acids) was similar to the full-length PRX2F proteins of the other organisms tested, which ranged between 211 – 235 amino acids. In addition, the macroalgal PRX2F contained the highly conserved CXXC motif as well as other regions of homology. This result was not unexpected since the kit used to create the original cDNA libraries (Creator^{TM} SMART^{TM} cDNA library construction kit, Clonetech) was designed to preferentially clone full-length genes only. The ORF which corresponded to a *G .gracilis* PRX2F gene was deemed suitable for cloning into the protein expression vector.

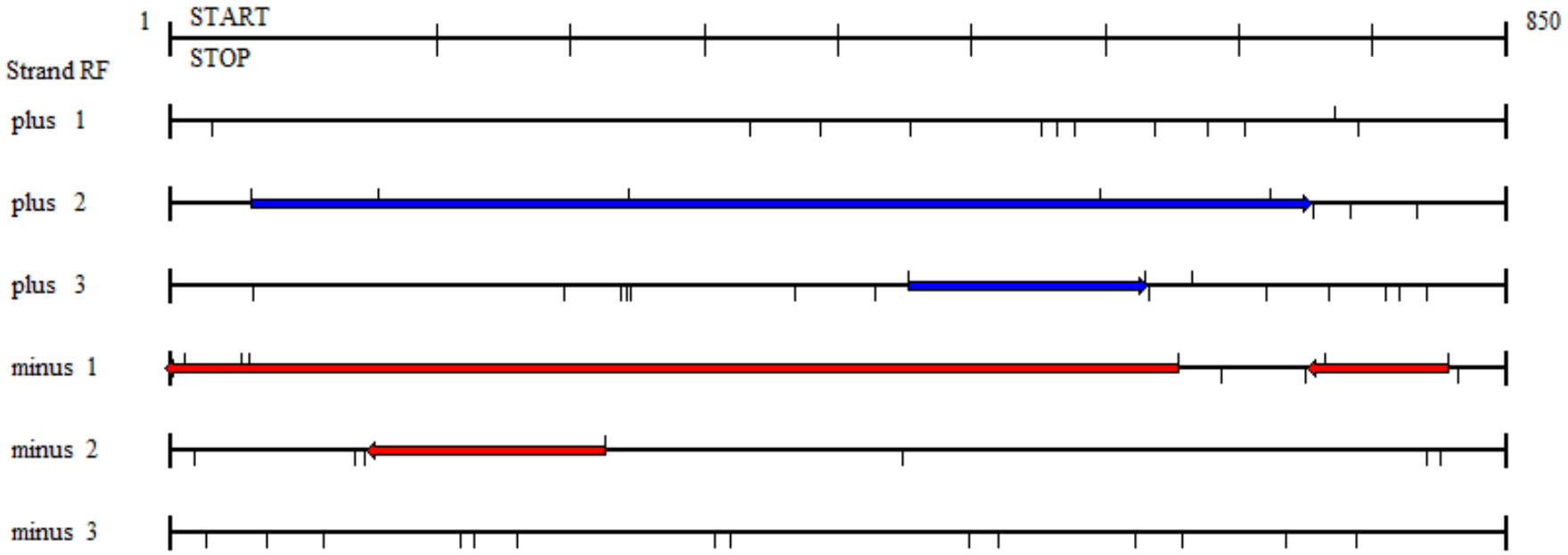


Figure 4.2 Prediction of open reading frames present in the cDNA clone harbouring the *G. gracilis* PRX2F gene sequence as determined using DNAMAN software (version 4.13). The coding region for PRX2F was located in frame 2 (+).

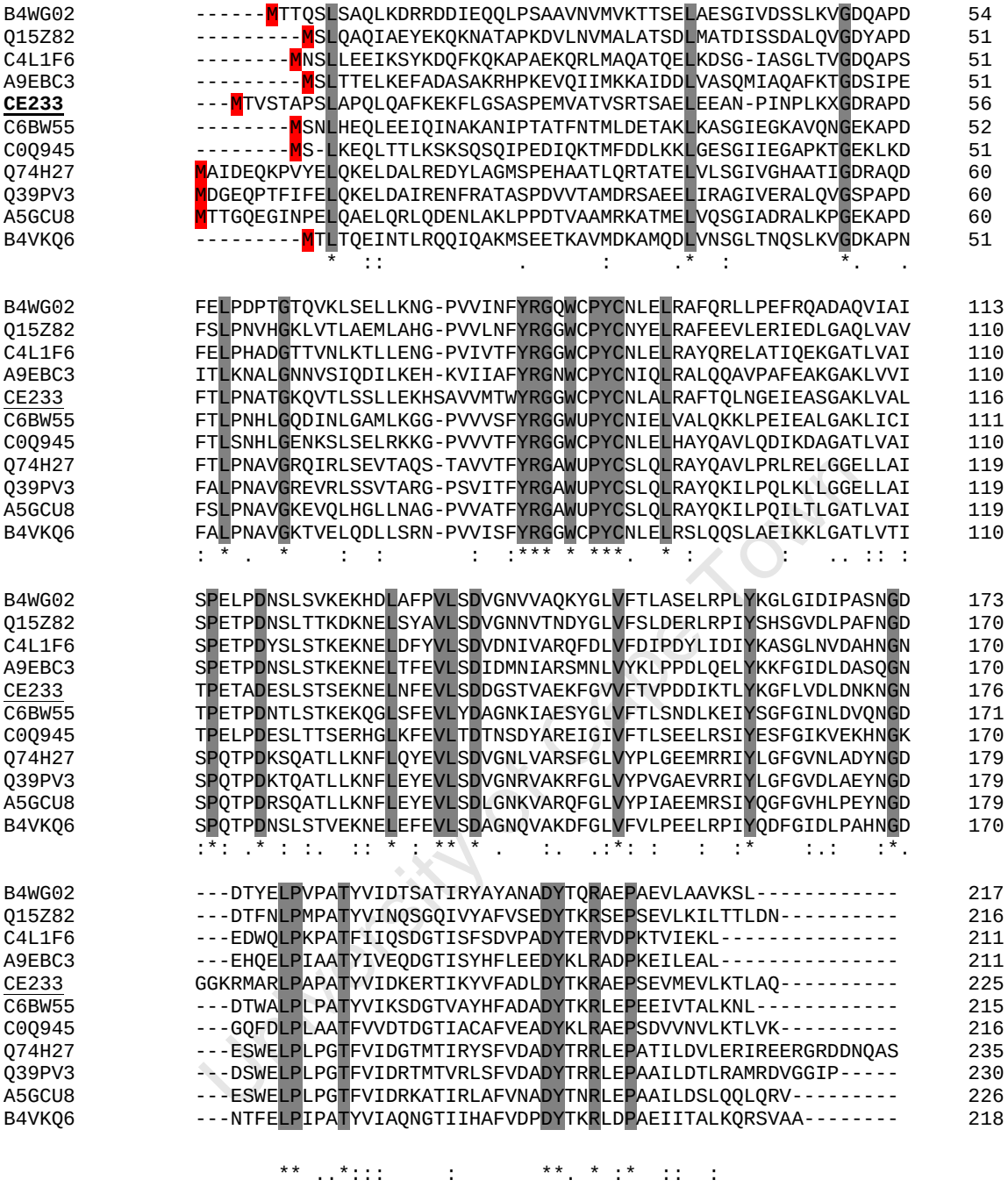


Figure 4.3 A multiple sequence alignment of eleven PRX2F proteins which included the *G. gracilis* PRX2F predicted protein was generated in ClustalW. Uniprot accession numbers are presented to the left of each sequence. The *G. gracilis* PRX2F protein is represented as 'CE233'. Sections highlighted in grey, denoted by "*", represent identical protein residues in each sequence. "." Represents conserved substitutions while ":" represents semi-conserved substitutions (similar shapes). Methionine start codons (M) are highlighted in red.

An ORF consisting of 999 nucleotides (333 amino acids and a molecular weight of 38.4 kDa), corresponding to the *G. gracilis* serine protease-like gene, was predicted in reading frame +3 (Figure 4.4). Several other ORFs detected within the cDNA sequence were small and possessed no significant sequence homology to other protein sequences in the GenBank database. Multiple protein sequence alignments were performed with the *G. gracilis* serine protease-like protein using only 4 other proteins which were specifically annotated as 'serine protease-like'. These included two full-length sequences from *A. thaliana* (Uniprot accession numbers Q8RWB6 and O23555-1) and two partial protein sequences (Q84P70 and BOS8H8) (Figure 4.5). The Q84P70 protein sequence deposited into the database Uniprot database did not contain a methionine start codon (M) but was still included in multiple sequence alignments. Several conserved protein motifs were detected within the *G. gracilis* serine protease-like protein. It appeared as if the 5' portion of the gene was missing from the *G. gracilis* cDNA sequence since the two full-length proteins were almost 100 amino acids larger than the other three proteins in the alignment (including the *G. gracilis* protein). This however was unlikely because the kit used to create the cDNA library should not have cloned partial cDNAs as a consequence of the cDNA amplification strategy it employed. 5' RACE was performed to obtain any 'missing' 5' sequence (data not shown). As expected, when the new DNA sequence data was assembled to the existing *G. gracilis* serine protease-like gene, the full-length gene was already present on the cDNA sequence (data not shown). Thus, the ORF encoding the *G. gracilis* serine protease-like gene was deemed suitable for cloning into the protein expression vector.

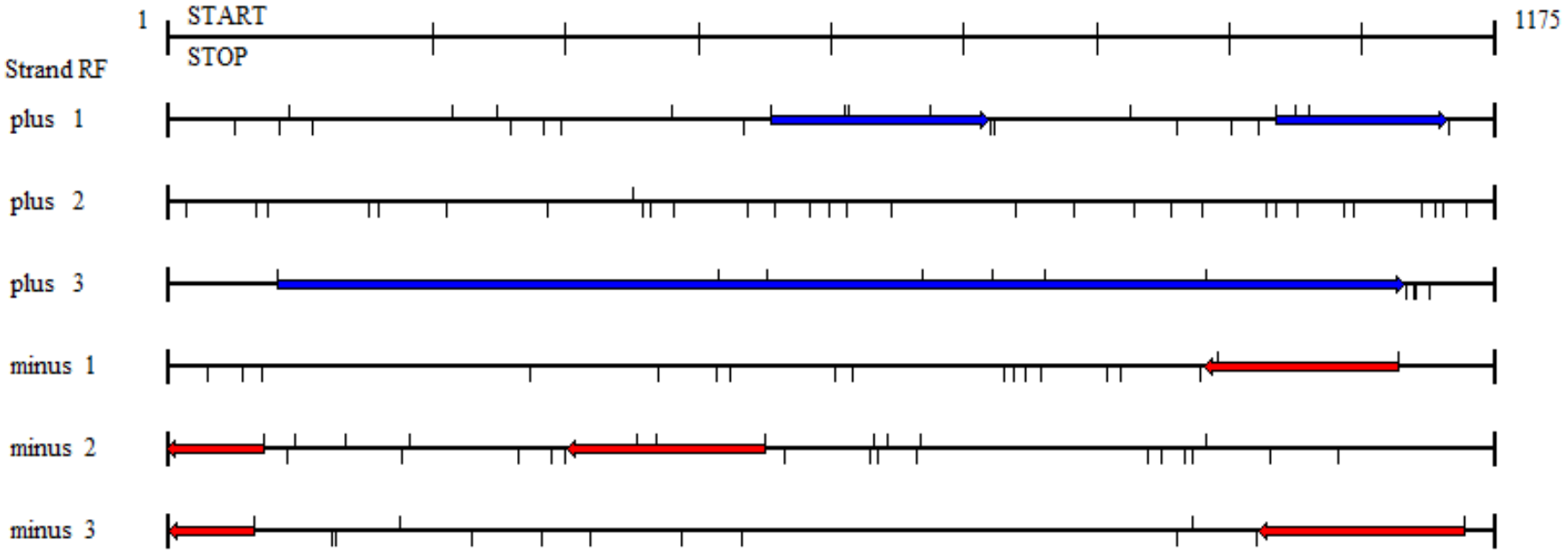


Figure 4.4 Prediction of the open reading frames present in the cDNA clone harbouring the *G .gracilis* serine protease-like gene sequence as determined using DNAMAN software (version 4.13). The coding region for the serine protease-like gene was located in frame 3 (+).

Figure 4.5 A multiple sequence alignment of four serine protease-like protein sequences was generated in ClustalW. Uniprot accession numbers are presented to the left of each sequence. The *G. gracilis* serine protease-like protein sequence is represented as 'CE428'. Sections highlighted in grey, denoted by "*", represent identical protein residues in each sequence. "." Represents conserved substitutions while ":" represents semi-conserved substitutions (similar shapes). Methionine start codons (M) are highlighted in red. The first two protein sequences (highlighted in yellow) correspond to *A. thaliana* proteins characterized as full-length.

4.3.2 PCR amplification of *G. gracilis* ORFs and cloning into pET29a protein expression vectors

A proof-reading DNA polymerase was used to PCR-amplify the two *G. gracilis* ORFs in order to prevent mutations or base-substitutions. The expected sizes of the PRX2F and serine protease-like coding regions were 675 and 999 bp respectively. When template DNA (circular plasmid harbouring cDNA insert) was not linearized, the PCR reactions failed which could not be explained. As observed in Figure 4.6, linearizing the plasmid using a restriction enzyme resulted in successful amplification of PCR products of the expected size with no non-specific amplification.

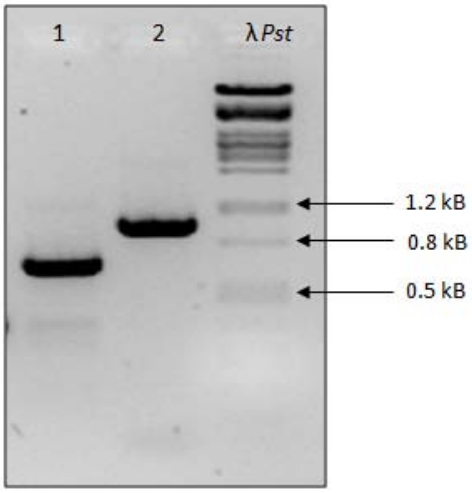


Figure 4.6 PCR amplification of the PRX2F and serine protease-like open reading frames using a proof-reading DNA polymerase. The PCR products were separated on a 1 % (w/v) TAE agarose gel. Lane (1) represented amplification of the PRX2F gene, lane (2) represented amplification of the serine protease-like gene and λ *Pst* represented the DNA molecular weight ladder.

Both PCR amplification products were gel-extracted and digested with the restriction enzymes *Bam*HI and *Hind*III. The protein expression vector pET29a (+) was similarly digested which enabled directional cloning of the amplified ORFs. Successful cloning reactions were assessed by performing a restriction digest using the same restriction enzymes to determine whether an insert of the expected size was excised from the plasmid vector. As

observed in Figure 4.7, inserts of the correct sizes were indeed cloned into the protein expression vectors. Furthermore, DNA sequencing results confirmed that pET29a-PRX2F and pET29a-serine protease-like contained the full-length genes without mutations. Based on DNA sequencing, both the S- and poly-histidine tags were in frame with the *G. gracilis* ORFs. Positive clones were used for subsequent protein expression studies.

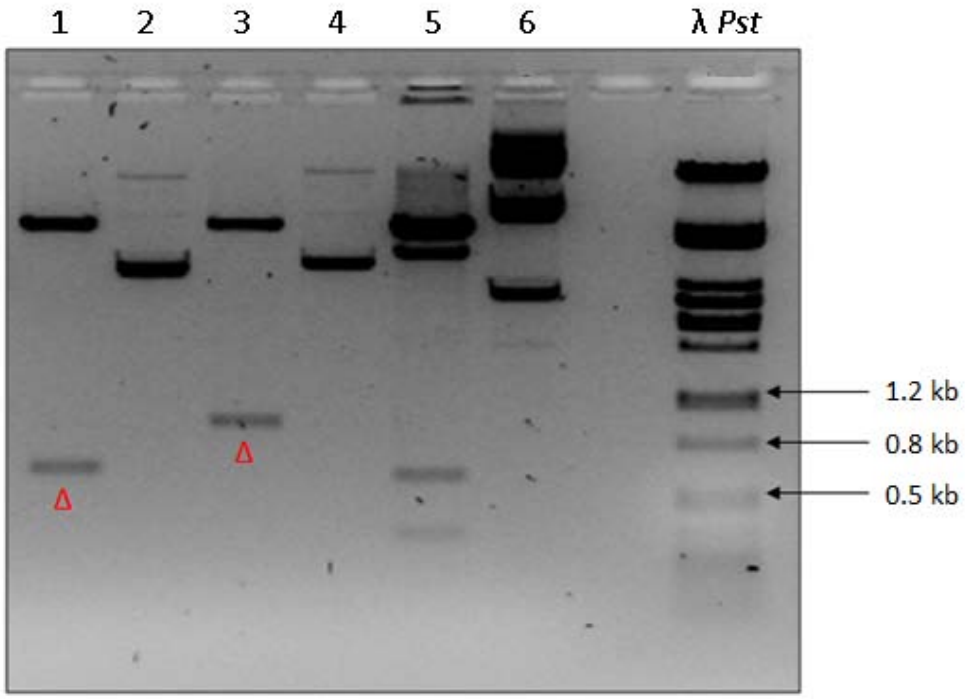


Figure 4.7 Restriction digest (*Bam*HI and *Hind*III) analyses of pET29a (+) vector constructs used for the protein expression studies. Digestion products were electrophoresed on a 1 % (w/v) TAE agarose gel. Lane (1): double restriction digestion of pET29a-PRX2F, lane (3): double restriction digestion of pET29a-serine protease-like gene and lane (5): double restriction digestion of pET29a. The corresponding uncut vector controls for each vector construct were represented in lanes (2), (4) and (6), respectively. λ*Pst*: DNA molecular weight ladder used to size the digestion products. Following restriction digestion with the two enzymes, DNA fragments of the expected sizes were excised from the expression vector (marked by red triangles Δ).

4.3.3 Recombinant protein expression

Foreign proteins that are highly expressed in *E. coli* either remain soluble within the cytoplasm or are deposited in inclusion bodies as insoluble proteins. A soluble recombinant protein was preferred because it would have retained its tertiary structure in addition to simplifying purification. An insoluble recombinant protein on the other hand, was not ideal as it necessitated purification using high concentrations of strong denaturants, such as urea, which would have removed the tertiary structure of the proteins. Considering that the recombinant proteins were required for polyclonal antibody production rather than structural analysis, solubility of the recombinant protein was not a major experimental variable. The predicted solubilities of both recombinant proteins in *E. coli* were based mainly on their amino acid sequences. Predictions of 45.5 % soluble and 93.9 % soluble were determined for the recombinant PRX2F and serine protease-like proteins, respectively.

Small-scale protein expression studies were carried out to confirm whether the recombinant proteins could be successfully expressed in the soluble protein fraction of *E. coli* cells. Equivalent amounts of bacterial cell samples in which protein expression was induced were lysed and the crude protein extracts from the soluble and insoluble fractions were separated by SDS-PAGE. The calculated molecular weights of the recombinant proteins were slightly larger than the native proteins due to incorporation of the two fusion tags and the amino acid residues that comprise the spacer region and protease cleavage site of the pET29a plasmid (Appendix D.3). Consequently, the expected molecular weight of recombinant PRX2F and serine protease-like protein was 30 kDa and 42.6 kDa, respectively. SDS-PAGE revealed no recombinant PRX2F expression in the un-induced samples (negative controls) (Figure 4.7, Lane 1 in Panels A and B) and progressive accumulation by 1 hour after induction of expression (Figure 4.8). At 24 hours post IPTG induction, no recombinant PRX2F protein was detectable in the soluble protein fraction (Figure 4.8; Panel A) which suggested that maximal expression occurred between 3 and 6 hours after IPTG induction. The predicted PRX2F solubility was confirmed since recombinant protein was detected in the soluble fractions. However, recombinant protein was equally expressed in the insoluble protein fractions (Figure 4.8; Panel B).

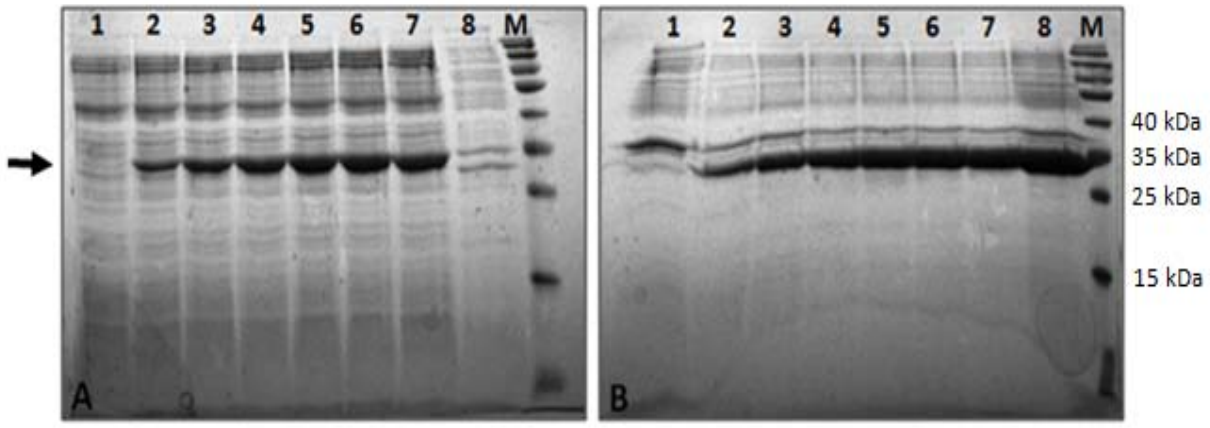


Figure 4.8 Coomassie stained protein gel showing expression of the His-tagged PRX2F protein in *E. coli* in the small scale expression study. *E. coli* cells were grown at 37°C in LB supplemented with 30 µg.ml^{-1} kanamycin until OD_{600} of 0.6 – 0.8 was reached. Cultures were transferred to 30°C and protein expression was induced by adding 1 mM IPTG. Bacterial cells were cultured for a further 24 hours. Aliquots of the broth were collected before induction (lane 1) and at 1, 2, 3, 4, 5, 6 and 24 hours after induction (lanes 2 – 8 respectively). Equal concentrations of the crude protein extracts were mixed with 5 X SDS application buffer, boiled for 5 min and separated on a 10 % SDS-PAGE. An expected 30 kDa recombinant protein (indicated by the arrow) was sized with the protein molecular weight marker (lane M). Panels (A) and (B) represent recombinant PRX2F in the soluble and insoluble fractions of *E. coli*, respectively.

The small-scale expression study of recombinant serine protease-like protein revealed that the protein could not be detected in either the soluble or insoluble protein fractions (data not shown). Various incubation temperatures, length of incubation as well as different IPTG concentrations were tested with no success. The culture media was finally changed to the auto-inducing Overnight Express Media ™ (Novagen). As seen in Figure 4.9, this media was conducive to recombinant protein expression except that expression was only observed in the insoluble protein fraction of *E. coli* (data for soluble protein fraction not shown). The predicted 94% solubility of the recombinant serine protease-like protein was therefore contradicted by its observed insolubility in *E. coli*. Recombinant protein was not detectable in the negative control (pET29a vector only).

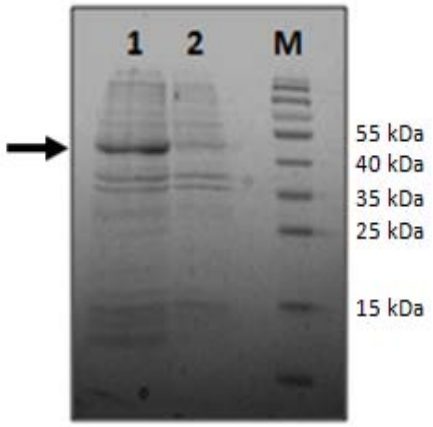


Figure 4.9 Coomassie stained protein gel showing expression of the His-tagged serine protease-like protein in the insoluble protein fraction of *E. coli* in the small scale expression study. *E. coli* cells were grown at 30°C for 26 hours in auto-induction media Overnight Express. Lane 1: recombinant serine protease-like protein expression. Lane 2: protein expression in the negative control. Equal concentrations of crude protein were mixed with 5 X SDS application buffer, boiled for 5 min and separated on a 10 % SDS-PAGE. A protein of the expected size (42.6 kDa) is marked by the arrow. The protein bands were sized according to the protein molecular weight marker (lane M).

The small scale trials proved that both recombinant proteins could be induced and strongly expressed in *E. coli* cells. However, scaling up the culture volume to 500 ml of culture media required further optimization as the established parameters did not yield reproducible results. Re-optimization (outlined in Section 4.2.4) was required to yield maximal amounts of recombinant protein for polyclonal antibody generation. Recombinant PRX2F was strongly induced by IPTG when incubated at 25°C and protein was observed mainly in the soluble protein fraction (Figure 4.10). The lowered temperature appeared to significantly improve solubility as no recombinant protein was detectable in the insoluble protein fraction (data not shown). In contrast, expression of the recombinant serine protease-like protein could not be achieved in the auto-inducing media as per the small scale trial. Instead, optimal expression was induced by IPTG when the bacterial cultures were incubated at a lowered incubation temperature (20°C). However, lowering the incubation temperature did not improve protein solubility since recombinant serine protease-like protein expression was only detected in the insoluble *E. coli* protein fraction (Figure 4.11). Ultimately, large scale expression of the

recombinant proteins was successful and purification using affinity chromatography could be undertaken.

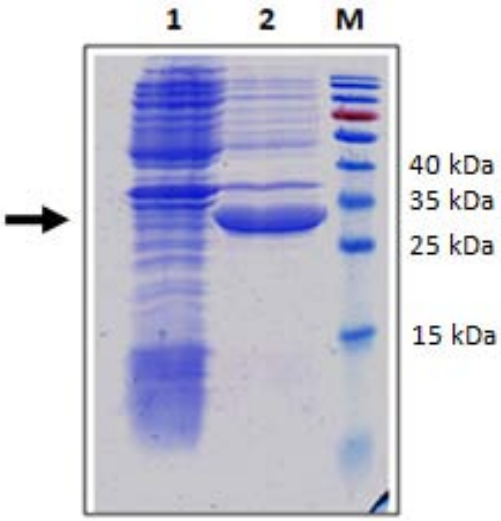


Figure 4.10 Coomassie stained protein gel showing expression of the His-tagged PRX2F protein in *E. coli* in the large scale expression study. *E. coli* cells were grown at 37°C in LB supplemented with LB supplemented with 30 µg.ml^{-1} kanamycin until OD_{600} of 0.6 – 0.8 was reached. Cultures were transferred to 25°C and protein expression was induced by adding 1 mM IPTG. The bacterial cells were cultured for a further 22 hours. Aliquots of the broth were collected before induction (lane 1) and at 22 hours after induction (lane 2). Equal concentrations of the crude protein extracts were mixed with 5 X SDS application buffer, boiled for 5 min and separated on 10% SDS-PAGE. A protein of the expected size (30 kDa), indicated by the arrow, was observed and sized with the protein molecular weight marker (lane M). The SDS-PAGE gel only shows the soluble protein fraction since no recombinant protein was observed in the insoluble *E. coli* protein fraction.

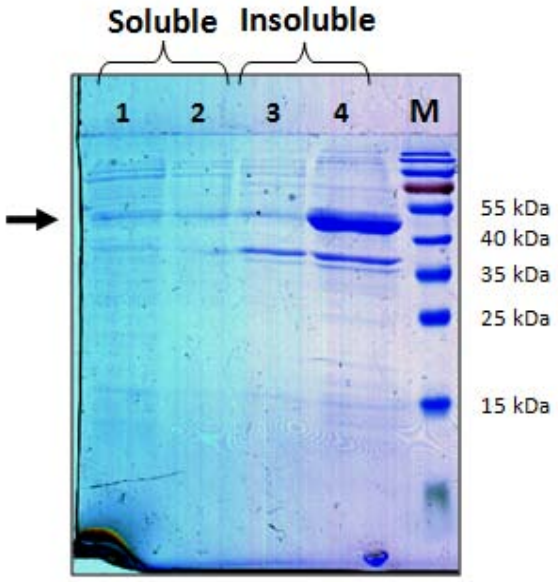


Figure 4.11 Coomassie stained protein gel showing expression of the His-tagged serine protease-like protein in *E. coli* in the large scale expression study. *E. coli* cells were grown at 37°C in LB supplemented with LB supplemented with 30 µg.ml^{-1} kanamycin until OD_{600} of 0.6 – 0.8 was reached. Cultures were transferred to 20°C and protein expression was induced by adding 1 mM IPTG. The bacterial cells were cultured for a further 23.5 hours. Aliquots of the broth were collected before induction (lanes 1 and 3) and at 23.5 hours after induction (lanes 2 and 4). Equal concentrations of the crude protein extracts were mixed with 5 X SDS application buffer, boiled for 5 min and separated on 10% SDS-PAGE. Soluble *E. coli* protein fractions for un-induced and induced samples are represented by lanes 1 and 2, respectively. Insoluble *E. coli* protein fractions un-induced and induced samples are represented by lanes 3 and 4, respectively. A protein of the expected size (42.6 kDa), indicated by the arrow, was observed and sized with the protein molecular weight marker (lane M).

4.3.4 Affinity purification of recombinant proteins

The 6X histidine-tagged recombinant proteins were purified from the *E. coli* cell lysates by immobilization on a HPLC metal-ion affinity-chromatography column, pre-packed with pre-charged Ni^{+} sepharose. Nickel ions have a high affinity for imidazole rings within histidine amino acids. Therefore, only the recombinant protein (and possibly other proteins with high histidine content) would have bound to the column whereas all other proteins would have been washed off the column. Since Ni^{+} ions are flooded in the presence of excess imidazole,

a high concentration of imidazole was used to release the bound histidine-tagged recombinant proteins from the affinity column. An imidazole gradient was employed over a 30 min period which allowed the protein density to be monitored (as a function of mV) as the elution progressed so that the exact fractions containing the highest concentration of eluted histidine-tagged recombinant proteins could be identified. Recombinant PRX2F was successfully purified under native conditions with no other non-specific *E. coli* proteins. Analysis of Figure 4.12 revealed the presence of one major peak which was maximally eluted at approximately 100 mM imidazole. Several protein fractions (represented by red dots in Figure 4.12) comprising the single protein peak were assessed for purity by SDS-PAGE which confirmed that a highly concentrated, pure recombinant PRX2F was obtained (Figure 4.13). Consequently, the eluted proteins in fractions 5 to 15 were pooled in order to maximize the amount of recombinant protein available for generation of polyclonal antibodies in rabbits. A pure recombinant protein was preferred for immunization to minimize non-specific cross-reactivity in later western hybridization analyses.

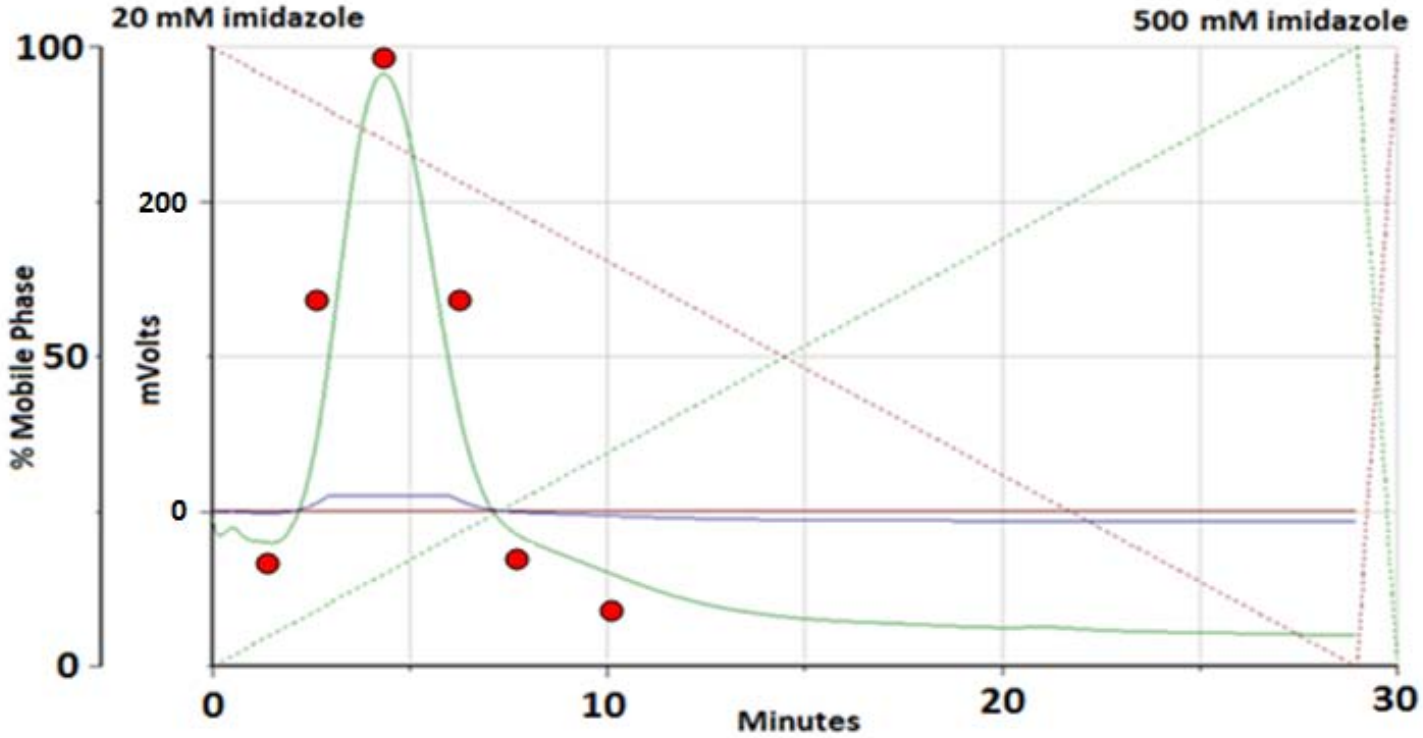


Figure 4.12 The elution profile of recombinant PRX2F following Ni+ affinity chromatography as generated using UniPoint software. The bound protein was washed and eluted over a 30 min period (X – axis) using two buffers which constituted an imidazole gradient (20 – 500 mM). The red dotted line represents the contribution of the 20 mM elution buffer to the imidazole gradient while the green dotted line represents the gradient of the 500 mM elution buffer along the 30 min (second Y-axis, % mobile phase). Voltage (mV) was monitored as elution of the recombinant protein progressed (Y – axis). Red dots from left to right represent elution fraction numbers 5, 7, 9, 11, 13 and 15 respectively which were assessed for purity by SDS-PAGE.

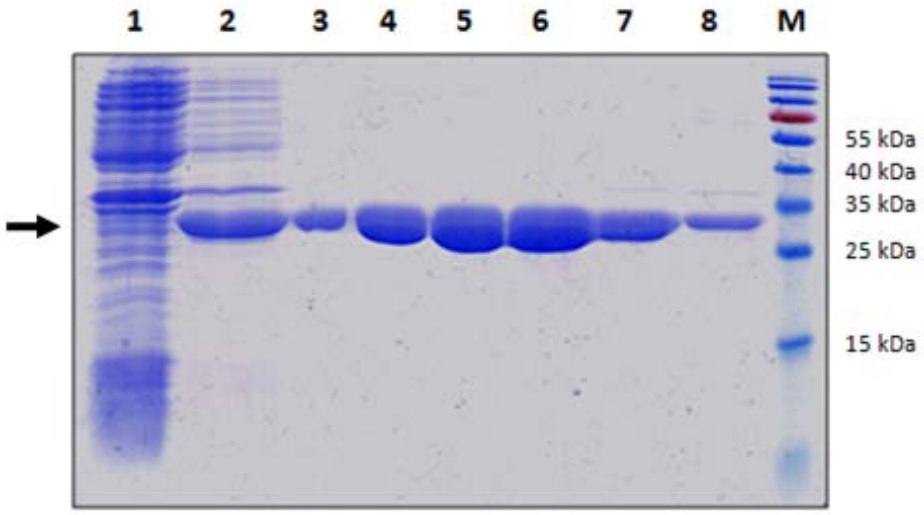


Figure 4.13 Coomassie stained protein gel showing the eluted fractions of His-tagged PRX2F recombinant protein purified under native conditions. Equal volumes of each fraction were mixed with 5 X SDS application buffer, boiled for 5 min and separated on 10% SDS-PAGE. The arrow indicates a protein band of the expected size (30 kDa). Lane 1: total *E. coli* protein without recombinant PRX2F. Lane 2: IPTG-induced expression of recombinant PRX2F in *E. coli*. Lanes 3 – 8: purified recombinant PRX2F from fractions 5, 7, 9, 11, 13 and 15, respectively. Lane (M): protein molecular weight marker.

Recombinant serine protease-like protein was affinity purified under denaturing conditions, using 8 M urea, as it was only expressed in the insoluble protein fraction of *E. coli*. Similarly, a high concentration of imidazole was used to elute recombinant proteins. Analysis of Figure 4.14 revealed the presence of one broad peak between 100 and 250 mM imidazole. This suggested that under denaturing conditions, the bound histidine-tagged proteins were either eluted less efficiently or that non-specific *E. coli* proteins were not properly removed during the wash steps. Nevertheless, several protein fractions (represented by red dots in Figure 4.14) were assessed for purity using SDS-PAGE. As observed in Figure 4.15, a protein with an expected size of 42.6 kDa was eluted from the column. The presence of several contaminating (presumably histidine residue rich) *E. coli* proteins was attributed to insufficient washing of the affinity column. Protein fractions 5 – 23 were pooled to maximize the amount of recombinant serine protease-like protein used to immunize rabbits.

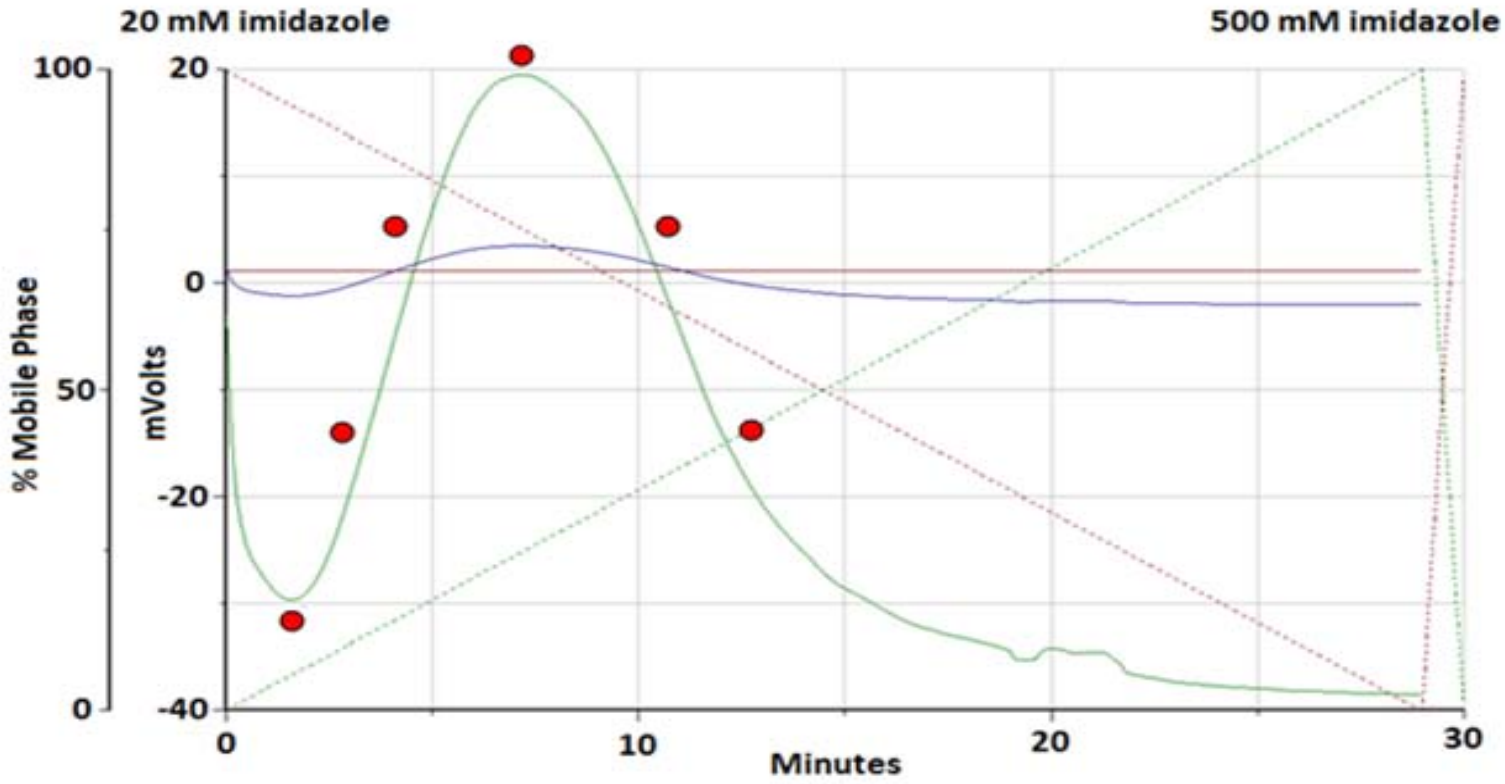


Figure 4.14 The elution profile of recombinant serine protease-like protein following Ni+ affinity chromatography as generated using UniPoint software. The bound protein was washed and eluted (under denaturing conditions) over a 30 min period (X-axis) using two buffers which constituted an imidazole gradient (20 – 500 mM). The red dotted line represents the contribution of the 20 mM elution buffer to the imidazole gradient while the green dotted represents the gradient of the 500 mM elution buffer along the 30 min elution profile (second Y-axis, % mobile phase). Red dots from left to right represent fraction numbers 5, 8, 11, 15, 20 and 23 respectively which were assessed for purity by SDS-PAGE.

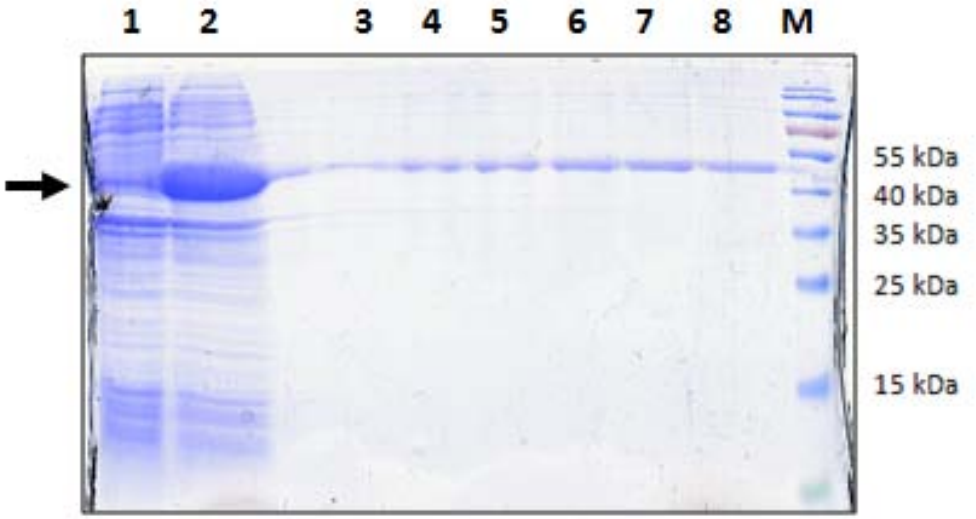


Figure 4.15 Coomassie stained protein gel showing the eluted fractions of His-tagged serine protease-like protein purified under denaturing conditions (8M urea). Equal volumes of each fraction was mixed with 5 X SDS application buffer, boiled for 5 min and separated on 10% SDS-PAGE. The arrow indicates a protein band of the expected size (42.6 kDa). Lane 1: total *E. coli* protein without recombinant serine protease-like protein. Lane 2: IPTG-induced expression of recombinant serine protease-like protein in *E. coli*. Lanes 3 – 8: purified recombinant serine protease-like protein from fractions 5, 8, 11, 15, 20 and 23, respectively. Lane (M): protein molecular weight marker.

Following Ni^{+} affinity chromatography, recombinant protein samples were de-salted, re-suspended in a neutral PBS buffer and quantitated. A western hybridization analysis was performed to further confirm that the histidine-tagged recombinant proteins had been purified. Western hybridizations using a polyclonal anti-histidine antibody revealed that the purity of both recombinant protein preparations was high (Figure 4.16). Purified recombinant PRX2F was completely pure as evidenced by the detection of no other non-specific histidine-containing *E. coli* proteins. As expected, the purified recombinant serine protease-like protein was the pre-dominant protein detected. However, the sample was less pure since two contaminating *E. coli* proteins with molecular weights between 15 and 25 kDa clearly cross-hybridized to the anti-histidine antibody. In addition, proteins of higher molecular weights were also detected by the anti-histidine antibody which confirmed the speculation that histidine residue rich *E. coli* proteins were not efficiently washed off the affinity column when recombinant serine protease-like protein was purified under denaturing conditions.

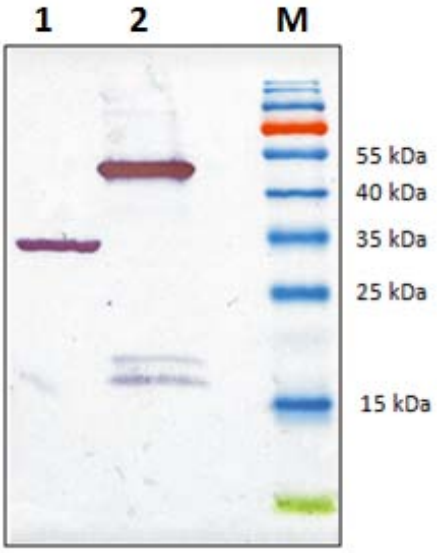


Figure 4.16 Western hybridization immunoblot of the purified Histidine-tagged recombinant proteins PRX2F (lane 1) and serine protease-like protein (lane 2) using a polyclonal anti-histidine antibody. An amount of 2.5 µg of each purified recombinant protein was separated using a 10% SDS-PAGE and transferred to a nitrocellulose membrane. A pre-stained molecular weight marker (lane M) was used to size the detected proteins. Signal development was achieved by monitoring colour development after a peroxidase substrate was added. The colour precipitation was stopped by incubating the membrane in water.

4.3.5 Polyclonal antibody production and determination of the titre and specificity

Affinity purified recombinant PRX2F and serine protease-like proteins were subsequently used to immunize rabbits for the generation of polyclonal antibodies. The rabbits were immunized and appropriately bled according to strict eight week schedule. A preliminary screening of each pre-bleed and four subsequent bleeds was performed to determine which bleed had the greatest titre of anti-PRX2F or anti-serine protease-like antibodies. Serial dilutions of the purified recombinant proteins were separately 'dot-blotted' in duplicate onto nitrocellulose membranes and each membrane was probed with one of the bleeds. Anti-sera from both pre-bleeds were moderately cross-reactive with the purified proteins since faint positive signals were produced (Figures 4.17 and 4.18; Panel A). After immunization, significant positive signals were immediately detectable for both antibodies as early as two weeks post-immunization (corresponding to Bleed 1) even at the lowest antigen dilution

(6.25 ng) (Figures 4.17 and 4.18; Panel B). Bleeds 2 to 4 all produced strong positive signals but they were not significantly higher than the first bleed (Figures 4.17 and 4.18; Panels C - E). Therefore, these observations suggested that immunization with 500 µg of the antigen (recombinant protein) was sufficient to produce a significant antigenic response in the rabbits within two weeks. Since any of the post-immunization bleeds could therefore be used for western hybridization studies, 'Bleed 4' was selected.

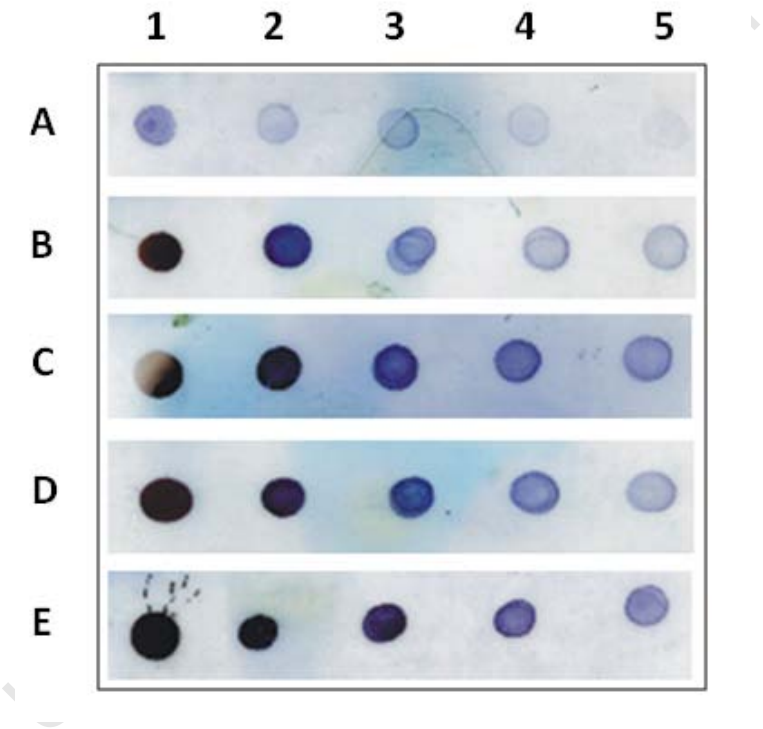


Figure 4.17 A preliminary 'dot-blot' western hybridization used to establish which bleed contained the highest titre of anti-PRX2F antibodies. Panels A – E represented the pre-bleed anti-sera, Bleed 1, Bleed 2, Bleed 3 and Bleed 4, respectively. Position 1 – 5 represented 100, 50, 25, 12.5 and 6.25 ng of purified PRX2F respectively. A chromogenic substrate was used to develop and visualize positive signals.

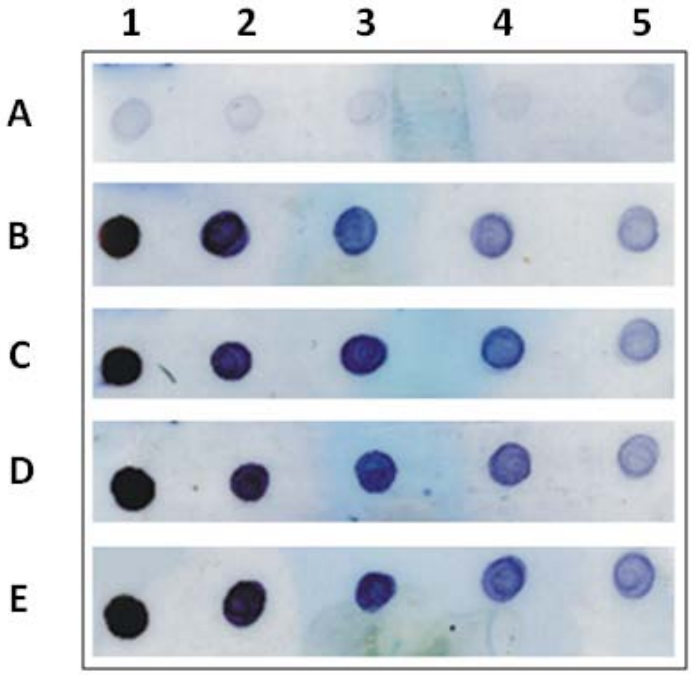


Figure 4.18 A preliminary 'dot-blot' western hybridization used to establish which bleed contained the highest titre of anti-serine protease-like antibodies. Panels A – E represented the pre-bleed anti-sera, Bleed 1, Bleed 2, Bleed 3 and Bleed 4, respectively. Position 1 – 5 represented 100, 50, 25, 12.5 and 6.25 ng of purified PRX2F respectively. A chromogenic substrate was used to develop and visualize positive signals.

The polyclonal antibodies were purified according to the protocol outlined by Polson *et al.* (1964) in order to eliminate antibodies in 'Bleed 4' that were raised against contaminating *E. coli* proteins. Total protein extracts from *G. gracilis* isolated from the 24 hour time-course experiment (Section 4.2.9) were used to test the specificity of the antibodies. Decreasing amounts (50, 40, 30, 20 and 10 µg) of total protein extract from the Time 0 sample was separated using SDS-PAGE and subsequently transferred to nitrocellulose membranes. Western hybridization analyses revealed that the polyclonal anti-PRX2F detected a protein of the expected size (25 kDa) (Figure 4.19; Panel A). A smaller protein was expected since the native *G. gracilis* protein would not have contained the S-tag or histidine-tag present on the recombinant proteins. Positive signal was saturated between 40 and 20 µg of total protein as no decrease in signal intensity was observed. However, a decrease in signal intensity was observed between 20 µg and 10 µg. As a result, an amount of 20 µg *G. gracilis* total protein was used for subsequent western hybridization analyses of PRX2F over the 24 hour time-

course experiment. A strong non-specific positive signal which corresponded to a *G. gracilis* protein with a larger molecular weight was also observed but since the primary antibodies had not been pre-absorbed against proteins from the bacterial protein expression host (*E. coli* BL21), it was not a surprising result. It was concluded that pre-absorption of the antibodies was necessary to eliminate this non-specific cross reactivity of the anti-PRX2F antibody (proven in Section 4.3.6). The anti-serine protease-like polyclonal antibodies were highly specific since the predominant positive signal detected corresponded to a *G. gracilis* protein of the expected size (35 kDa) (Figure 4.19; Panel B). Positive signals were saturated between 50 and 20 µg of total *G. gracilis* protein. However, a decrease in positive signal was observed between 20 and 10 µg total protein. As a result, an amount of 15 µg total *G. gracilis* protein was selected for western hybridization analyses of the serine protease-like protein over the 24 hour time-course experiment. A faint, non-specific positive signal which corresponded to a *G. gracilis* protein with a higher molecular weight was also observed. Similarly, this phenomenon was eliminated by pre-absorbing the primary antibody against proteins from the *E. coli* expression host.

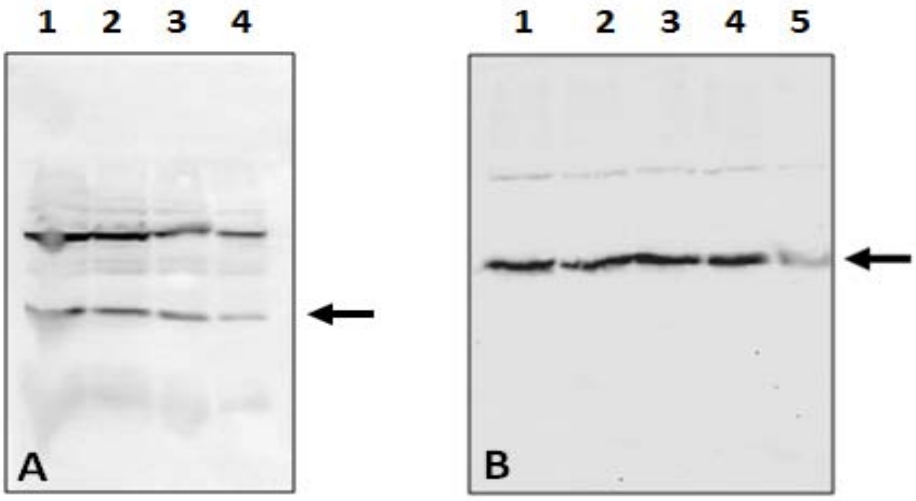


Figure 4.19 Western hybridization analyses to determine the specificity of polyclonal antibodies against *G. gracilis* PRX2F (Panel A) and serine protease-like protein (Panel B). Polyclonal antibodies were purified from 'Bleed 4'. A concentration range of *G. gracilis* total protein extract from the Time 0 sample was separated by 10% SDS-PAGE and transferred to nitrocellulose membranes. Membranes were probed with a 1:1000 dilution of polyclonal antibody and either a 1:1000 (PRX2F) or a 1:50000 dilution (serine protease-like protein) of the secondary antibody (alkaline phosphatase-goat anti-rabbit IgG conjugate). Panel A (Lanes 1- 4) represented 40, 30, 20 and 10 µg total *G. gracilis* protein, respectively. Panel B (Lanes 1- 5) represented 50, 40, 30, 20 and 10 µg total *G. gracilis* protein, respectively. The arrows indicate the putative *G. gracilis* PRX2F and serine protease-like proteins.

4.3.6 Western hybridization analysis of 24 hour time-course experiment

Prior to western hybridization analyses of the PRX2F and serine protease-like proteins, the total protein extracts from all *G. gracilis* samples obtained from the 24 time-course experiment (three biological repeats) were assessed for protein integrity using SDS-PAGE. Isolated protein was of excellent quality with no signs of degradation (Figure 4.20) which ensured that the results observed in subsequent western hybridizations would not be compromised.

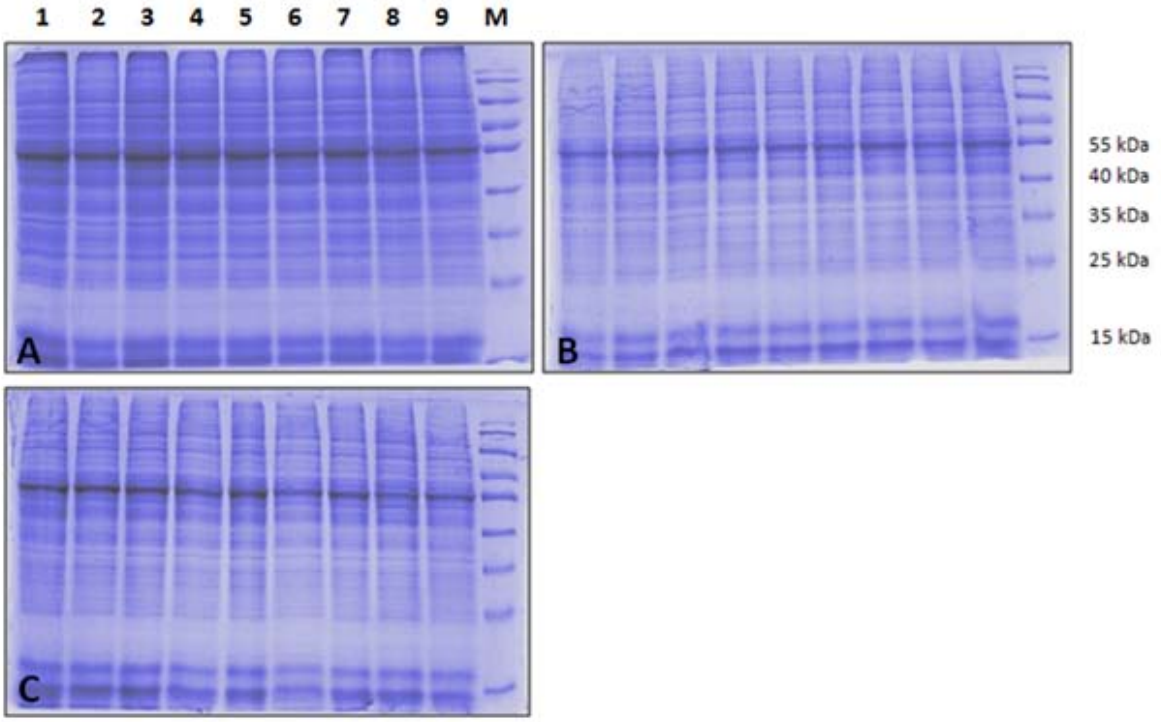


Figure 4.20 Coomassie stained proteins gels to establish the quality and integrity of the *G. gracilis* total protein extracts from the 24 hour time-course experiment. Twenty micrograms of total protein were mixed with 5X SDS application buffer, boiled for 5 min and separated on 10% SDS-PAGE. Biological repeats 1, 2 and 3 are represented by panels A, B and C, respectively. Lanes 1 – 9 represent the loading order of the total protein samples extracted from time 0, 1 hour post exposure to disease elicitors, 1 hour control, 8 hour post exposure to disease elicitors, 8 hour control, 21 hour post exposure to disease elicitors, 12 hour control, 24 hours post exposure to disease elicitors and 24 hour control from the 24 hour time-course experiment, respectively. All three proteins gels were loaded similarly. Proteins were sized with a pre-stained molecular weight ladder (lane M).

Following the protein quality check, the optimal amount of *G. gracilis* protein (Section 4.3.5) from each time-point sample in the 24 hour time-course experiment was separated by SDS-PAGE and subsequently transferred to nitrocellulose membranes for western hybridization analyses of the native *G. gracilis* PRX2F and serine protease-like proteins. Densitometry and statistical analysis was used to determine whether the amount of PRX2F protein changed significantly as a result of exposure to disease elicitors. Between 1 and 8 hours post addition of the disease elicitors, *t*-tests failed to detect significant changes in the levels of PRX2F protein expression in the experimental samples relative to their time-point controls (Figure

4.21; Panel D). In contrast, at 12 and 24 hours, the levels of PRX2F protein in experimental samples were greater (approximately 1.5 fold in both cases) relative to their time-point controls but *t*-tests suggested that the observed changes were not statistically significant ($P > 0.05$). However, in both instances the power of the *t*-tests performed was less than the desired power of 0.800. As previously established in Chapter 3, low power of statistical tests is often as a result of too few replicates or large variability in the data set (Williams *et al*., 1997) and may be indicative of a possible type II error (false negative). Cumming *et al*. (2007) suggests that by increasing the number of biological repeats, narrower inferential error bars with more precise estimates of the true population values can be achieved. Nevertheless, a distinct increase in the expression of PRX2F protein was evident as a consequence of exposure to the disease elicitors. Pre-absorption of the polyclonal anti-PRX2F antibody against *E. coli* proteins proved highly efficient since all the non-specific bands were eliminated. A single positive PRX2F signal was observed in all three biological repeat immunoblots (Figure 4.21; Panels A – C). The nitrocellulose membranes were then stripped and re-probed with anti-rubisco antibodies to confirm equal loading of *G. gracilis* protein. A single positive signal was detected on all three membranes (Figure 4.22; Panels A – C). Densitometric and statistical analysis (one-way ANOVA) of the rubisco bands detected in the *G. gracilis* protein extracts failed to detect statistically significant differences. This finding confirmed equal sample loading which further suggested that the observed changes in PRX2F levels were not due to technical error.

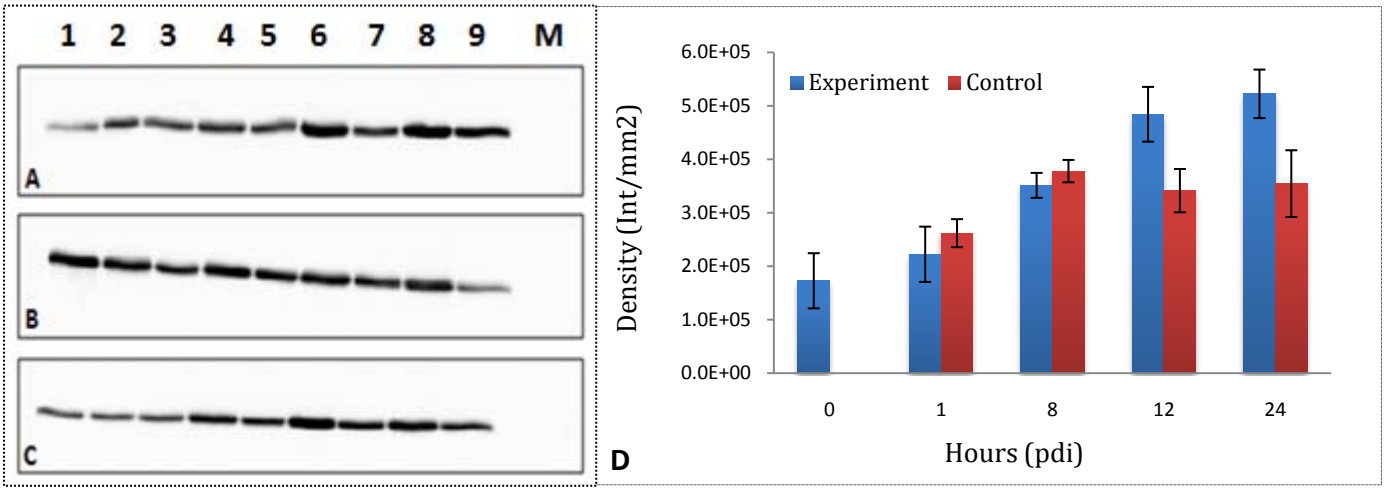


Figure 4.21 Western hybridization analysis of *G. gracilis* PRX2F protein levels over 24 hours following exposure to disease elicitors. Twenty micrograms of protein were mixed with 5X SDS application buffer, boiled for 5 min, separated on 10% SDS-PAGE and transferred to nitrocellulose membranes. A 1:1000 dilution of anti-PRX2F was used to detect the native *G. gracilis* PRX2F proteins. Western hybridization results for biological repeats 1, 2 and 3 are represented by panels A, B and C, respectively. Lanes 1 – 9 represent the loading order of protein samples as 0, 1E, 1C, 8E, 8C, 12E, 12C, 24E and 24C from the 24 hour time-course experiment, respectively. Proteins were sized with a pre-stained molecular weight ladder (lane M). Panel D represents the quantitative analysis of each positive signal which is displayed as a mean densitometry reading ± standard error (intensity.mm^{-2}), n = 3. Parametric *t*-tests failed to detect statistically significant changes in PRX2F expression.

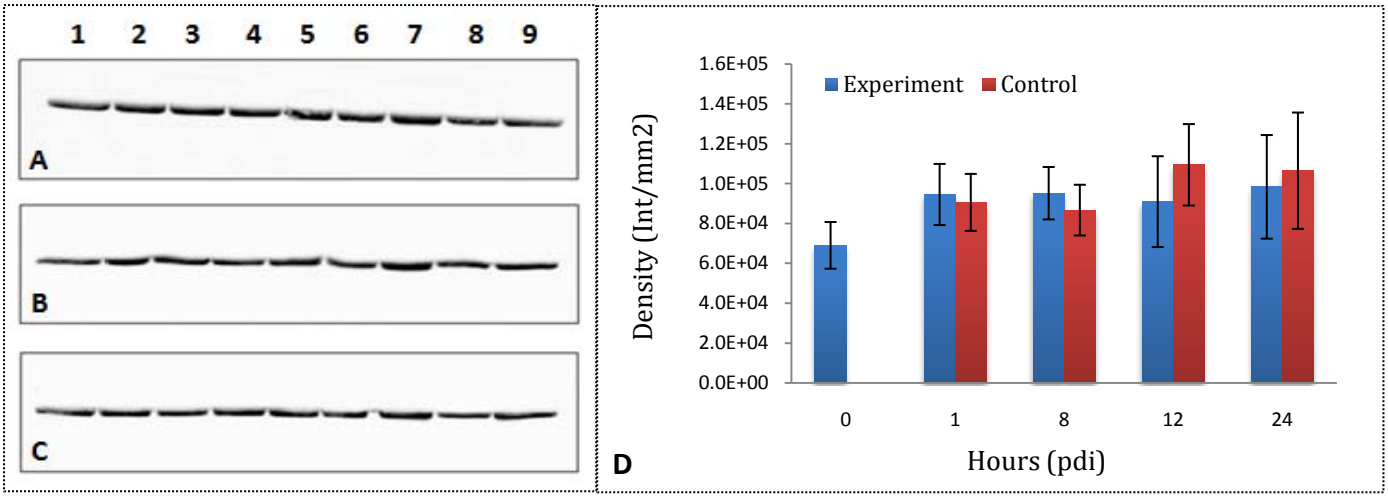


Figure 4.22 Western hybridization analysis of *G. gracilis* Rubisco protein (loading control) over the 24 hour time-course experiment. Nitrocellulose membranes in Figure 4.20 were stripped and re-probed with the anti-rubisco antibody. A 1:1000 dilution of anti-rubisco was used. Biological repeats 1, 2 and 3 are represented by panels A, B and C, respectively. Lanes 1-9 represent the loading order of protein samples as 0, 1E, 1C, 8E, 8C, 12E, 12C, 24E and 24C, respectively from the 24 hour time-course experiment. Proteins were sized with the existing molecular weight ladder already on the blot. Panel D represents the quantitative analysis of each positive signal which is displayed as a mean densitometry reading ± standard error (intensity.mm^{-2}), n = 3. A one-way ANOVA revealed no statistically significant differences between the levels of rubisco protein.

The anti-serine protease polyclonal antibodies had been pre-absorbed against *E. coli* proteins prior to western hybridization to minimize the possibility of non-specific cross-reactivity. This proved highly efficient since only a single positive signal, corresponding to the expected size of the *G .gracilis* serine protease-like protein was observed in all three biological repeat immunoblots (Figure 4.23; Panels A – C). However, western hybridization analysis of the *G. gracilis* serine protease-like protein suggested that exposure to disease elicitors had no clear biological effect on the translational regulation of this protein over a 24 hour time-course (Figure 4.23; Panel D) since the level of serine protease-like protein in the experimental and control seaweed samples appeared approximately equal as determined by densitometric analysis. Although it was possible that the moderate increase in the level of the protein in the 24 hour experimental sample was due to exposure to the disease elicitors (Figure 4.23; Panel D), a parametric *t*-test failed to detect a statistically significant change. Similarly, the power

of the *t*-tests performed was less than 0.800 suggesting that more biological repeats may be required to clarify translational regulation of the serine protease-like protein upon exposure to disease elicitors. The membrane was stripped and probed with anti-rubisco antibodies (Figure 4.24; Panels A – C) to discount unequal loading or unequal protein transfer to the nitrocellulose membranes as the cause of the observed lack of significant changes in the levels of serine protease-like protein. As expected, no statistically significant differences in the level of rubisco in the protein samples were detected by a one-way ANOVA, confirming that equal amounts of *G. gracilis* total protein had been loaded in each well of the polyacrylamide gels.

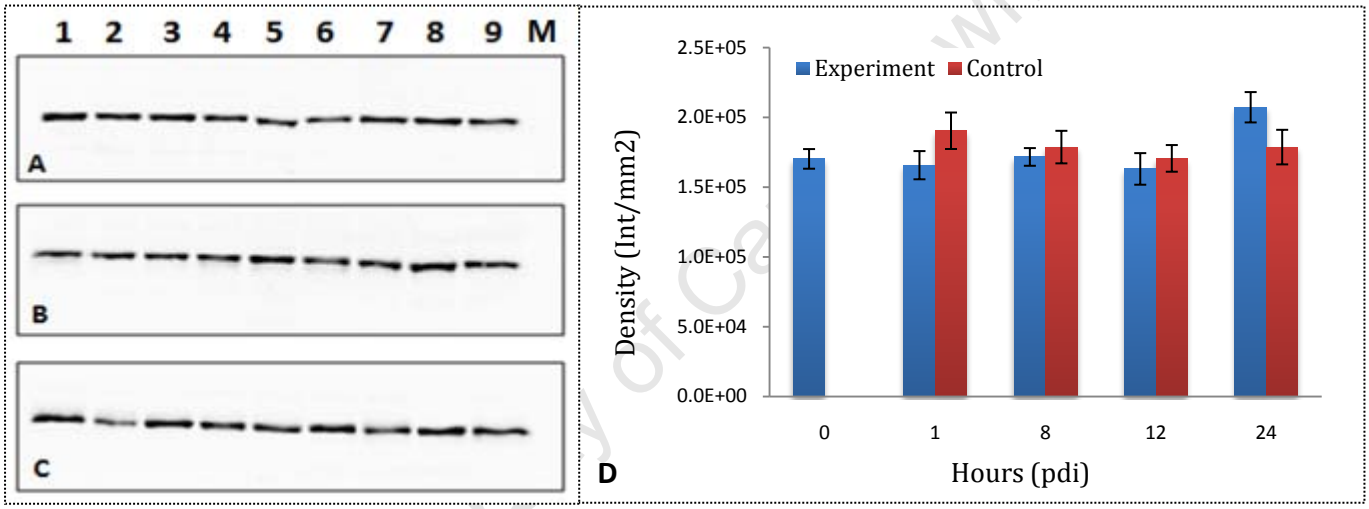


Figure 4.23 Western hybridization analysis of *G. gracilis* serine protease-like protein levels over the 24 hours following exposure to disease elicitors. Fifteen micrograms of protein were mixed with 5X SDS application buffer, boiled for 5 min, separated on 10% SDS-PAGE and transferred to nitrocellulose membranes. A 1:1000 dilution of anti-serine protease-like protein was used to detect the native *G. gracilis* serine protease-like protein. Western hybridization results for biological repeats 1, 2 and 3 are represented by panels A, B and C, respectively. Lanes 1 – 9 represent the loading order of protein samples as 0, 1E, 1C, 8E, 8C, 12E, 12C, 24E and 24C from the 24 hour time-course experiment, respectively. Proteins were sized with a pre-stained molecular weight ladder (lane M). Panel D represents the quantitative analysis of each positive signal which is displayed as a mean densitometry reading ± standard error (intensity.mm^{-2}), n = 3. Parametric *t*-tests failed to detect statistically significant changes in serine protease-like protein expression.

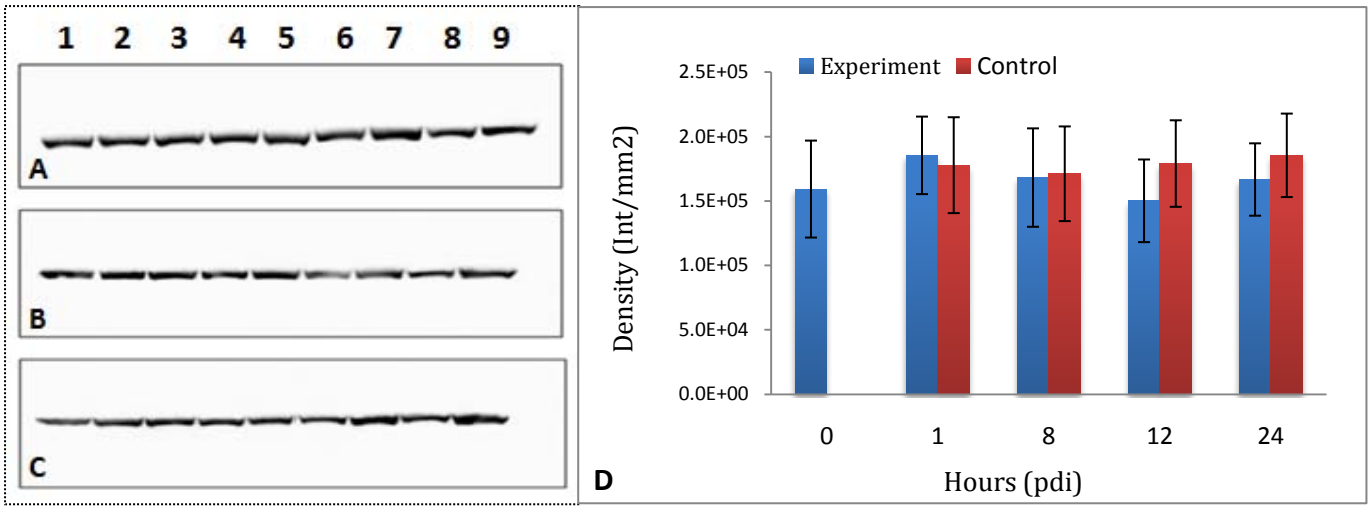


Figure 4.24 Western hybridization analysis of *G. gracilis* Rubisco protein (loading control) over the 24 hour time-course experiment. Nitrocellulose membranes in Figure 4.22 were stripped and re-probed with the anti-rubisco antibody. A 1:1000 dilution of anti-rubisco was used. Biological repeats 1, 2 and 3 are represented by panels A, B and C, respectively. Lanes 1-9 represent the loading order of protein samples as 0, 1E, 1C, 8E, 8C, 12E, 12C, 24E and 24C, respectively from the 24 hour time-course experiment. Proteins were sized with the existing molecular weight ladder already on the blot. Panel D represents the quantitative analysis of each positive signal which is displayed as a mean densitometry reading ± standard error (intensity.mm^{-2}), n = 3. A one-way ANOVA revealed no statistically significant differences between the levels of rubisco protein.

4.4 DISCUSSION

The mechanisms of stress tolerance, including activated defence responses, may be regulated at the genetic and molecular levels. More specifically, regulation of the defence genes in *G. gracilis* may be controlled at four distinct levels, i.e. transcription, post-transcription, translation and post-translation. In addition, the regulation of each defence gene and its encoded protein product may be unique and may involve the coordination of a specific subset of regulatory components to ensure that the required amount of protein product is synthesized at the correct time. In Chapter 3, the transcriptional regulation of several *G. gracilis* (putative) defence genes was assessed in response to exposure to endogenous disease elicitors. Significant changes in the levels of mRNA transcripts were detected in seaweed samples exposed to the elicitors relative to the control seaweed samples. The next logical step based on the observed transcriptional changes was to assess whether corresponding changes in the levels of the functional protein (gene product) could be observed. As stipulated in the introduction to this chapter, the correlation between mRNA and protein levels has previously been shown to be poor and attributable to various experimental and/or biological variables. Nie *et al*. (2006) provided possible reasons for this discrepancy. The authors suggest that measurements of mRNA and protein abundance that have not been carried out using identical samples or using only a small subset of proteins (versus the correlation of total protein samples) may adversely influence the calculated correlation values. In this study, the mRNA : protein correlation of only two proteins was assessed which comprises a very small subset of *G. gracilis* proteins. However, the seaweed samples used in the translational study were identical to those used to characterize transcriptional regulation in the previous chapter.

Multiple protein sequence alignments of *G. gracilis* PRX2F and serine protease-like proteins with their respective homologous protein sequences revealed each ORF encoded a full-length gene. Furthermore, conserved domains and/or semi-conserved amino acid substitutions were detected in both *G. gracilis* protein sequences. ORFs encoding the two genes were cloned separately into protein expression vectors and used to transform *E. coli* cells. Recombinant PRX2F and serine protease-like proteins were successfully expressed, purified using Ni^{+} affinity chromatography and used to generate polyclonal antibodies in rabbits. Polyclonal

antibodies were used in western hybridization analyses to quantitate the levels of native PRX2F and serine protease-like proteins in complex *G. gracilis* protein samples (either exposed on not exposed to disease elicitors) over a 24 hour time-course. At the translational level, expression of the PRX2F protein appeared to increase in *G. gracilis* samples exposed to disease elicitors at the 12 and 24 hour time-points. This observation coincided with the observed induced transcription of the PRX2F gene over the 24 hour time-course (Chapter 3; Section 3.3.6). In both instances (transcription and translation), PRX2F expression generally remained constant in the control samples over the 24 hour time-course presumably due to a lack of oxidative stress in the seaweed samples. Weinberger *et al*. (1999) showed that agar oligosaccharides rapidly induce the release of reactive oxygen species (ROS) in macroalgae, including *G. gracilis*, resulting in a hypersensitive response (HR) and plant cell death (PCD) manifested as thallus bleaching. In this study, the release of ROS, HR and associated PCD was not specifically tested over the 24 hour time-course but the pattern of induced PRX2F gene and protein expression strongly suggested that these biological phenomena may have occurred in *G. gracilis* as a consequence of exposure to disease elicitors. The positive correlation between the observed transcriptional changes of the PRX2F gene and PRX2F protein levels makes biological sense because unlike a mRNA transcript, the antioxidant PRX2F protein would be the functional unit in the cell. In other words, the increased expression of the PRX2F gene in *G. gracilis* samples exposed to disease elicitors had to be accompanied by a corresponding increase in the translation of PRX2F protein in order for the cell to cope with oxidative stress. The translational regulation of PRX2F was not assessed for the 30 min experiment since no significant changes in levels of the protein were detected between the experimental and control samples at the 1 hour time-point. A detailed discussion regarding the possible biological roles of peroxiredoxin proteins in the context of activated defence was provided in Chapters 2 and 3. However, evidence in the literature further supports increased expression of peroxiredoxin-like proteins under various biotic (and abiotic) stresses which are known to cause an oxidative burst. In a study by Ndimba *et al*. (2005), the authors used a proteomics approach (two-dimensional polyacrylamide gel electrophoresis (2D-PAGE)) to identify *A. thaliana* proteins involved in salt and osmotic stress. One protein in particular, with significant homology to a peroxiredoxin-like protein, displayed a 29% increase in expression. Since peroxiredoxins are known to possess antioxidant capabilities, the authors concluded that the salt and osmotic stress caused an oxidative burst in *A. thaliana* which would have been ameliorated through the action of the

antioxidant peroxiredoxin-like protein. Similarly, Aranjuelo *et al*. (2010) used 2D-PAGE to identify protein regulated during a different abiotic stress (drought) on alfalfa (*Medicago sativa*) plants. An alfalfa protein which displayed homology to a peroxiredoxin-like protein was found to be significantly up-regulated. Chen *et al*. (2007) employed 2D-PAGE to identify maize kernel endosperm proteins associated with resistance to aflatoxin contamination by *Aspergillus flavus*. Of the proteins found to be up-regulated in resistant genotypes compared to the susceptible ones, one protein in particular displayed significant sequence homology to an antioxidant peroxiredoxin protein. The authors proposed that induced expression of the peroxiredoxin-like protein in resistant maize was directly linked to maize's ability to resist *A. flavus* infection or to tolerate aflatoxin production more readily. Furthermore, transcriptional regulation of this peroxiredoxin-like gene was also assessed in response to infection with the fungal pathogen. At the end of a 25 day fungal infection time-course study, the level of the peroxiredoxin mRNA transcript was significantly higher in the resistance strain relative to the susceptible one. This observation supports the finding of the current study in which up-regulation of the *G. gracilis* PRX2F gene was positively correlated to up-regulation of the PRX2F protein in seaweed samples exposed to disease elicitors. Mowla *et al*. (2002) also showed a positive correlation between transcription and translation of a peroxiredoxin. Peroxiredoxin mRNA transcript levels were absent in fully hydrated *Xerophyta viscosa*, increased in plants subjected to abiotic stress such as dehydration, heat, high light intensity, abscisic acid treatment and exposure to high salinity. A similar pattern of protein expression was observed in response to the different stress conditions using western hybridization analysis. Zhou *et al*. (2005) used 2D-PAGE to show that infection of wheat (*Triticum aestivum*) with *Fusarium graminearum* resulted in the up-regulation of several defence-related proteins including an antioxidant peroxiredoxin protein. Several other studies also suggest a role for peroxiredoxins in pathogen defence. For example, peroxiredoxin proteins were up-regulated in resistant *Brassica carinata* and remained unaffected in susceptible *B. Carinata* when challenged with a fungal pathogen (Subramanian *et al*., 2005). Similarly, infection of poplar leaves with a fungal pathogen led to a hypersensitive response and resulted in up-regulation of peroxiredoxin proteins over a 72 hour time-course. Interestingly, in disease-resistant poplar, the level of peroxiredoxin protein decreased as the infection proceeded (Rouhier *et al*., 2004).

Western hybridization analysis of *G. gracilis* serine protease-like protein levels in total protein samples suggested that a positive correlation exists between mRNA and protein levels. Although mRNA transcription appeared to be repressed between 1 and 12 hours after *G. gracilis* had been exposed to disease elicitors, *t*-tests failed to detect statistically significant changes between the control and experimental samples at each of the time-points sampled (Chapter 3; Section 3.3.6). Similarly, levels of the serine protease-like protein were approximately equal in all *G. gracilis* samples analysed over the 24 hour time-course. Taken together, these observations suggest that exposure to disease elicitors has no biological effect on the expression and translation of this serine protease-like gene. In contrast, Tornero *et al*. (1996, 1997) showed that expression of proteins with domains homologous to the subtilisin serine protease family were induced and accumulated in tomato (*Lycopersicon esculentum*) upon viral infection. The concentration of these proteins was enhanced by the presence of Ca^{2+} ions and accumulated in the intercellular spaces of leaves from virus-infected plants. Using northern blot analysis, the authors further demonstrated that the increased protein expression was positively correlated with increased gene expression and concluded that the subtilisin serine protease formed part of an activated defence response in tomato. Four different subtilisin-like serine protease (P69 protein family) genes were characterized in tomato (Jordá *et al*., 1999). Although they were closely related (79 – 88% identity), the four genes showed different transcriptional regulation and protein expression patterns during development, pathogen infection or exposure to salicylic acid. Two of the subtilisin-like serine protease genes were constitutively expressed during development whereas the other two were up-regulated during pathogen infection. Jordá *et al*. (2000) identified two more members in the P69 protein family and concluded that this group of proteases was part of a complex gene family of plant serine proteases with different regulation patterns and varied roles in activated defence. Therefore, the observed lack of translational (and transcriptional) regulation of the *G. gracilis* serine protease-like gene identified in this study does not completely exclude its contribution to an activated defence response. It may be involved at the level of elicitor perception, in signalling pathways or for the recycling of amino acids required for the synthesis of defence proteins (Antão and Malcata, 2005; van der Hoorn and Jones, 2004). If this *G. gracilis* serine protease-like protein is in fact constitutively expressed, it may be very biologically relevant. For instance, specific binding of elicitors may activate the protease which may activate downstream signalling components by proteolytic cleavage (Ariki *et al*., 2004). Alternatively, binding of an elicitor to the protease

may inhibit its activity and the elicitor-protease complex or altered proteolytic activity may induce signalling leading to disease resistance (van der Hoorn *et al*., 2002; Jabs, 1999). Constitutively expressed proteases may even execute a defence response by directly degrading proteins belonging to the invading pathogen or activate enzymes from precursor proteins. Proteases that accumulate at the site of an infection may be candidates for this role. Measuring the contribution of each protease may be challenging since they may all be part of an array of cooperative defence responses (van der Hoorn and Jones, 2004). A recent study identified a novel role for subtilisin-like protease from soybean (*Glycine max*) (Pearce *et al*., 2010). A 12 amino acid peptide isolated from soybean was found to be derived from a member of the subtilisin-like protease family. The gene was expressed in all actively growing tissue and was not induced by wounding, methyl jasmonate or salicylic acid. However, the expression of known defence-related genes was induced when a synthetic form of the 12 amino acid peptide was added to soybean cultures. The authors therefore concluded that this 12 amino acid peptide was a unique plant defence signal which was cryptically embedded within a plant protein with an independent physiological role (i.e. protease). In *A. thaliana*, the serine protease family appears to be the largest, most diverse class of proteases in plants with a wide array of potential molecular functions (Schaller, 2004). If the serine protease family of proteins are equally as diverse in *G. gracilis*, this study has merely opened the door for further analysis of other serine proteases in the activated defence responses of macroalgae. Some of the possible roles for plant serine proteases in the context of activated defence have already been alluded to in Chapters 2 and 3.

4.5 CONCLUSION

Characterization of the functions of genes is dependent on deducing the various biological steps between gene transcription and translation into the protein molecule which is ultimately responsible for a physiological effect in the cell. Under a given cellular state – activation of the defence response in this case – elicitors and signalling molecules act as effectors prompting the cell to adapt and survive in the given situation. In previous chapters, cDNA microarrays proved that several *G. gracilis* genes were transcriptionally regulated in response to exposure to disease elicitors. These observations were subsequently tested and validated (2 out of 5 genes) using qPCR. This chapter aimed to establish whether the observed transcriptional changes were positively correlated to altered protein levels. In order to test this hypothesis, western hybridization analyses were used to quantitate the *G. gracilis* proteins (PRX2F and serine protease-like protein) present in total protein preparations obtained from the same seaweed samples used in the qPCR experiments. Open reading frames of the two *G. gracilis* putative defence genes were cloned into protein expression vectors. Recombinant proteins were expressed in *E.coli*, purified using Ni^{+} affinity chromatography and subsequently utilized to immunize rabbits for the generation of polyclonal antibodies. Western hybridization analyses using anti-PRX2F polyclonal antibodies revealed that expression of PRX2F was induced in *G. gracilis* samples exposed to elicitors at 12 and 24 hours. This finding was positively correlated to the transcriptional regulation of the PRX2F gene which was up-regulated in *G. gracilis* samples at 8 and 12 hours post exposure to elicitors. Therefore, based on the fact that PRX2F is annotated with antioxidant activity, these results confirmed that *G. gracilis* responds with an oxidative burst upon exposure to disease elicitors. Increased levels of PRX2F protein implies that *G. gracilis* is therefore able to adapt to and possibly negate the potentially toxic effects of ROS. In contrast, western hybridization analysis revealed that the *G. gracilis* serine protease-like protein was not translationally regulated in response to disease elicitors. However, because serine proteases belong to a diverse family of proteins, the observed lack of transcriptional and translational regulation of this specific *G. gracilis* serine protease-like protein could not rule out a role for it in the activated defence response of *G. gracilis*. Further investigation is thus required to elucidate the specific role of this protein in *G.* gracilis.

CHAPTER 5

General discussion and future work

5.1 General discussion

All eukaryotic organisms may, on occasion, be subjected to an onslaught of pathogen attack be it fungal, nematode, bacterial or viral. It is generally accepted that in response to such an attack, higher eukaryotes activate various defence mechanisms which are initiated by the interactions of innate receptors and disease elicitors. Adaptive receptors exist only in vertebrates, but all metazoans and vascular plants express innate receptors (Nürnberger and Brunner, 2002; Cohn *et al*., 2001). Until recently, it was thought that macroalgal defence mechanisms were constitutive and comprised primarily of chemical-based defence molecules (Weinberger, 2007). However, as established in Chapter 1, published evidence supports the existence of pathogen-induced macroalgal defence mechanisms which function to complement their chemical-based defence repertoire. Furthermore, several molecular components of activated defence in macroalgae appear to be evolutionarily conserved in higher plants, invertebrates and animals since disease resistance proteins and immune signalling pathways harbour common motifs and rely on similar downstream signalling pathways, respectively (Weinberger, 2007). Briefly, the current model of pathogen-induced macroalgal defence complies with an evolutionarily conserved biochemical map in all eukaryotic organisms as follows: disease elicitors (exogenous and endogenous) are recognized by elicitor receptor(s); receptor-elicitor interactions result in the rapid release of reactive oxygen species (ROS) (Lamb and Dixon, 1999; Weinberger *et al*., 1999); ROS potentially confers toxicity directly towards invading pathogens, may activate or halogenate ions of various organic substrates involved in defence signalling pathways and further mediate inter- and intra-cellular signalling necessary for the transcriptional regulation of defence genes. Following transcriptional regulation of defence genes, various pathogenesis related proteins are synthesized which are ultimately required to achieve disease resistance in the host. Furthermore, localized strengthening of the cell wall may take place but as a final measure of pathogen containment, the hypersensitive response (HR) and programmed cell death (PCD) which are both mediated by ROS, may be initiated (Nürnberger and Lipka,

2006; Mysore and Ryu, 2002). The work of Collén *et al.* (2006) demonstrated that, as in higher plants (Schenk *et al.*, 2000), the plant hormone methyl jasmonate mediated transcriptional regulation of several genes (presumably involved in activated defence) in the red macroalga *C. crispus*. However, no such studies have been performed for *G. gracilis* to date.

In Chapter 2, a cDNA microarray was constructed from two previously established stress libraries, i.e. (i) a cDNA library constructed using RNA extracted from *G. gracilis* exposed to endogenous disease elicitors for 24 hours and (ii) a cDNA library constructed using RNA extracted from *G. gracilis* after 18 days of nitrogen limitation. In total, 1620 cDNAs were PCR-amplified, purified and printed to create a microarray. The small size of this cDNA microarray was a limitation in this study because the low density of cDNAs failed to represent the entire nuclear genome of *G. gracilis*. As such, it could not be used as an index of global mRNA expression patterns in *G. gracilis* after 24 hours of exposure to disease elicitors. Nevertheless, the cDNA microarray experiment was successful in identifying transcriptionally regulated genes (within those present on the microarray) and thus satisfied the aims of this study. A total of 51 *G. gracilis* genes were differentially expressed and sequenced. Sequenced cDNA fragments were subjected to various BLAST analyses. *G. gracilis* genes were annotated for biological function by comparing DNA sequence to orthologous genes in other organisms. Significant transcriptional regulation of *G. gracilis* genes after 24 hours of exposure to the disease elicitors appeared to involve genes encoding proteins involved in stress responses (9%), protein processing (3%), cell structure (9%), metabolism (14%), respiration (3%), protein synthesis (6%) and DNA modification (3%). Approximately half of the transcriptionally regulated genes identified encoded proteins with unknown or novel functions. This suggested that, for the genes present on the microarray, *G. gracilis* did indeed display some level of evolutionary conservation in its genetic mechanisms of activated defence. Alternatively, the genes that could not be functionally annotated suggested that several genetic components unique to *G. gracilis* may have been functional as well. A detailed discussion proposing the possible roles of each sequenced gene in the context of activated defence was provided. Despite the observed evolutionary conservation between the molecular mechanisms of activated defence between macroalgae and higher

organisms, insight in this field prior to this study was mainly hindered by a lack of available genetic information for *G. gracilis*.

In Chapter 3, the transcriptional regulation of five *G. gracilis* genes identified in the microarray experiment was validated and further characterized using qPCR analyses. In addition to the quantitation of mRNA transcripts in healthy seaweed (Time 0) and after 24 hours of exposure to disease elicitors, mRNA transcripts for the five genes were quantitated during the first 30 min and at 1, 8 and 12 hours post exposure to the elicitors. The 30 min time-course experiment was prompted by the observations of Weinberger *et al*. (2005, 2000 and 1999) in which *Gracilaria conferta* exposed to disease elicitors released ROS into the algal growth media within minutes. Statistical analysis of the gene expression data for all five putative defence genes was affected by low power of the test which was attributed to the small number of biological repeats (three) or large variation in the data sets (error bars) (Williams *et al*., 1997). As a result, it was possible that a lack of statistically significant changes were in fact type II errors (false positives). Thus, this finding highlighted the importance of performing sufficient biological replication in future qPCR experiments. Narrower inferential error bars with more precise estimates of the true population values (Cumming *et al*., 2007) will be more conducive to making biological inferences in any attempt to elucidate the activated defence response of *G. gracilis*. Consequently, no statistically significant changes in transcription of the five putative defence genes were observed in *G. gracilis* samples exposed to disease elicitors in the 30 min time-course experiment. This observation did not negate the results of Weinberger *et al*. (2005, 2000 and 1999) but rather emphasized the limitations associated with low density cDNA microarrays. It would be interesting to perform the same experiment using cDNA microarrays representative of the entire *G. gracilis* genome. Such an experiment will enable a conclusive result to be drawn with respect to early responsive genes which are transcriptionally regulated. On the other hand, statistically significant changes in transcriptional regulation were however observed in the 24 hour experiment. The most dynamic and significant changes in gene expression were generally observed between 8 and 12 hours after exposure to the elicitors. A gene encoding an antioxidant protein (peroxiredoxin-like) was significantly up-regulated in *G. gracilis* samples exposed to elicitors whereas the control samples displayed no significant changes. Thus, receptor-elicitor interaction not only results

in a rapid oxidative burst in *Gracilaria* species but also implies that ROS may mediate the transcriptional regulation of antioxidant genes (and possibly other defence genes) required for disease resistance. Reactive oxygen mediation of gene transcription during activated defence mechanisms is widely documented in higher eukaryotes (Apel and Hirt, 2004; Laloi *et al*., 2004; Dalton *et al*., 1999; Alvarez *et al*., 1998) which further supports the hypothesis that mechanisms of activated defence may have originated in the oceans. The transcriptional regulation of a phosphoserine phosphatase (PSP) gene suggested that phosphorylation events may potentially be necessary in the activated defence response of *G. gracilis*. However, reversible phosphorylation is complex and further research is required to definitively characterize this specific *G. gracilis* PSP. The transcriptional regulation of phosphotidylserine decarboxylase (PSD) in *G. gracilis* samples exposed to disease elicitors was distinct. At every time-point sampled during the 24 hour time-course, significant repression of PSD expression was observed in seaweed samples exposed to disease elicitors. Although the pattern of PSD transcriptional regulation also implicated phosphorylation events during the activated defence response of *G. gracilis*, further scientific research is required to define its physiological role. Alternatively, transcriptional repression of this *G. gracilis* PSD may coincide with activation of plant cell death (PCD) (O'Brien *et al*., 1997), a widely documented mechanism of activated defence in all eukaryotic organisms (Lam *et al*., 2001; Heath, 2000; Greenberg, 1997; Pennell and Lam, 1997). Although statistically significant changes were not observed in the transcriptional regulation of the *G. gracilis* serine protease-like gene, a trend of gene repression upon exposure to disease elicitors was evident. Serine proteases have been implicated in activated defence mechanisms such as the hypersensitive response (HR), signal transduction pathways and protein processing which are necessary during PCD. Similarly, analysis of the transcriptional regulation of a cytosolic thioredoxin gene did not detect statistically significant changes in gene expression, but did reveal a biologically significant rend of repression in seaweed samples exposed to disease elicitors. Thioredoxins have been implicated in the regulation of several enzymes, modulation of transcription, cellular development, acting as hydrogen donors for other enzymes and for oxidative damage prevention in the cell. Since increased transcriptional expression of a gene encoding an antioxidant protein (peroxiredoxin-like) supported the occurrence of an oxidative burst in *G. gracilis* samples exposed to disease elicitors, the repression of cytosolic thioredoxin gene expression observed in this study eliminated its role as a gene encoding an antioxidant protein. Consequently, further characterization is required

before its role in the activated defence response of *G. gracilis* can be elucidated. It is important to highlight that these five genes were selected for transcriptional analysis because scientific literature linked them specifically to roles in the activated defence mechanisms of other organisms. The high proportion (approximately 50%) of genes identified in the microarray experiment which could not be functionally annotated did not reduce their importance in the activated defence response of *G. gracilis* but excluded them for further characterization by qPCR. Validation of the transcriptional regulation between the microarray and qPCR experiments revealed that two genes (PSP and serine protease-like protein) displayed congruent gene expression patterns whereas the remaining three genes (thioredoxin, phosphoserine phosphatase and phosphatidylserine decarboxylase) displayed conflicting transcriptional regulation patterns. This apparent lack of correlation between transcription and translation for these three genes was mainly attributed to the fact that the mRNA used in both experiments was derived from different biological samples. Nevertheless, the transcriptional regulation of these five genes established through these qPCR experiments enabled their roles in the context of *G. gracilis* to be discussed in more detail.

The observed changes in transcriptional regulation of the five putative defence genes in *G. gracilis* could not be assumed to positively correlate to changes in the levels of the functional unit (protein product). As a result, Chapter 4 sought to establish this relationship for two genes, i.e. peroxiredoxin-like and serine protease-like. The open reading frames (ORFs) of each full-length gene was cloned into a protein expression vector and expressed in an *Escherichia coli* host. Recombinant proteins were expressed and subsequently purified using affinity chromatography. Rabbits were immunized with the purified proteins in order to generate polyclonal antibodies. Polyclonal antibodies were subsequently used to establish the translational regulation of each protein by quantitating their levels via western hybridization analyses. Western hybridization analyses were performed using the same *G. gracilis* samples used in the 24 hour time-course experiment in Chapter 3. Investigation of the translational regulation of the peroxiredoxin-like protein revealed a positive correlation between levels of the mRNA transcript and its protein product in seaweed samples exposed to disease elicitors. Protein levels were significantly increased between 12 and 24 hours post exposure to the elicitors relative to the control samples. This observation expanded on the current model of

activated defence in *G. gracilis* because it proved that *G. gracilis* actively produced an antioxidant protein presumably to detoxify the oxidative burst following elicitor perception. In order to further assess the antioxidant properties of the *G. gracilis* PRX2F protein, an *in vitro* DNA protection assay could be performed (Govender, 2006; Brehelin *et al*., 2003; Klimowski *et al*., 1997). By incubating DNA with purified, recombinant PRX2F protein in the presence of ROS, the level of PRX2F antioxidant activity as well as its ability to protect against DNA damage may be determined. Alternatively, *in vivo* protection assays performed with bacterial cells expressing recombinant PRX2F could be used to determine the level of sensitivity to H_2O_2 (Govender, 2006). Lastly, peroxiredoxin specificity assays as tested by Finkemeier *et al*., (2005) could aid in elucidating the substrate specificities and enzyme kinetics of the *G. gracilis* PRX2F protein.

Investigation of the translational regulation of the *G. gracilis* serine protease-like protein suggested that exposure of the seaweed to disease elicitors had no significant effect on the levels of the protein. Since plant serine proteases appear to be the largest class of proteases (Schaller, 2004), it is possible that *G. gracilis* also possesses many serine protease isoforms. Each may be involved in various physiological roles and could be regulated differently at the level of transcription and translation. More specifically, the *G. gracilis* serine protease-like protein characterized in this study may be required for elicitor perception, involved in signalling pathways or required for the recycling of amino acids needed for the synthesis of other defence proteins (Antão and Malcata, 2005; van der Hoorn and Jones, 2004). Thus the observed lack of transcriptional and translational regulation of this *G. gracilis* serine protease-like protein merely necessitates further characterization if its physiological role in the activated defence response of *G. gracilis* is to be clearly understood. In addition, the role of other possible serine proteases should be researched further in the context of activated defence in macroalgae.

5.2 Future work

Gene expression profiling to elucidate putative defence gene function was approached through a RNA based system biology methodology, i.e. transcriptomics. This technique is powerful, enabling the screening of hundreds of genes simultaneously. Techniques such as

cDNA microarrays, cDNA amplified fragment length polymorphism (AFLP) and serial analysis of gene expression (SAGE), though able to identify novel genes via analyses of transcriptional regulation, are more conducive to model organisms or species with characterized genomes (Carpentier *et al*., 2008). This current study was limited with respect to the density of the cDNA microarray but it was still effective in successfully identifying several transcriptionally regulated putative defence genes. As genomic data for *G. gracilis* expands, future experiments should accurately represent its entire genome on a microarray. Such experiments will provide a broader understanding with respect to which genes are transcriptionally regulated during activated defence. In addition, transcriptomics studies on the genome scale enable clustering of genes with synergistic or antagonistic expression patterns which was not possible in the current study. More comprehensive studies of the *G. gracilis* transcriptome may facilitate testing the hypothesis for evolutionary conservation of activated defence in a more precise manner. Of the putative defence *G. gracilis* identified in the microarray experiment, only five were further transcriptionally characterized due to the high costs of qPCR experiments. Therefore, future studies should apply the experimental approach employed in this study to characterize the remaining differentially expressed *G. gracilis* genes identified in the microarray experiment. Western hybridization analyses were useful for quantitating the levels of two *G. gracilis* proteins but in future, more efficient approaches could be implemented to not only assess translational regulation, but also for the identification of novel defence proteins. The use of a proteomics approach such as 2D SDS-PAGE has already proved successful in this regard (Aranjuelo *et al*., 2010; Chen *et al*., 2007; Ndimba *et al*., 2005; Zhou *et al*., 2005). Separation and analysis of proteins extracted from *G. gracilis* samples either exposed or not exposed to disease elicitors would not only identify all the proteins regulated upon exposure to the elicitors but also yield a relative quantitation for each. Since protein sequences are more conserved, proteins from organisms with unsequenced genomes can be compared to orthologous proteins of well characterized species (Liska and Shevchecnko, 2003; Schevchenko *et al*., 2001). Since proteins are the functional units within the cell, i.e. the actively translated component of the mRNA pool, a proteomics approach may provide greater insight into the defence proteins recruited by *G. gracilis* during an activated defence response. Furthermore, proteomic analysis facilitates the investigation of post-transcriptional and post-translation regulatory events (Ingle *et al*., 2007) which have been documented in activated defence mechanisms of higher eukaryotes.

A separate aspect which arose from the results of this study was the potential use of putative defence genes in genetic transformation studies. The context of this study was that in aquaculture settings, *G. gracilis* farming may be subject to various biotic (and abiotic stresses) which could possibly hinder seaweed production. Therefore, all the genes identified as transcriptionally up-regulated and thus associated with disease resistance are candidates for determining whether their over-expression in *G. gracilis* would confer enhanced disease tolerance. Such studies would involve seaweeds over-expressing various putative defence genes and comparing these to antisense lines as well as control wild-type specimens. Furthermore, future studies involving the analysis of the upstream stress inducible promoter regions and their efficacy to drive stress responsive genes may be required to generate disease resistant seaweed with none of the adverse phenotypic effects that occur with constitutively expressed genes.

In conclusion, this study confirmed the role of agar oligosaccharides as elicitors of defence in the red macroalga *G. gracilis*. The results observed provided an insight into some of the genes transcriptionally regulated in *G. gracilis* after exposure to these elicitors (albeit limited with the incomplete representation of the entire genome on the microarray). Only two genes were characterized at the translational level using western hybridization analyses which highlighted the need for similar studies of all the other putative defence genes. To the best of our knowledge, this is the first study to provide an insight into the genes transcriptionally regulated during the activated defence response (biotic stress) of *G. gracilis*. Further studies are required to obtain a complete understanding of the functions of these genes (and their protein products) in the acquisition of disease resistance. Each defence gene can be characterized individually but ultimately, implementation of a systems biology approach combining proteomics, metabolomics and transcriptomics could offer a comprehensive description of the regulatory networks governing disease resistance in *G. gracilis*.

REFERENCES

Abrahamsson, K., Choo, K. S., Pedersen, M., Johannson, G., & Snoeijs, P. (2003). Effects of temperature on the production of hydrogen peroxide and volatile halocarbons by brackish-water algae. *Phytochemistry*, *64*, 725-734.

Adamska, I. (1997). ELIPs — Light-induced stress proteins. *Physiol. Plant.*, *100*, 794–805.

Aharoni, A., & Vorst, O. (2001). DNA microarrays for functional plant genomics. *Plant Mol Biol*, *48*, 99–118.

Altschul, S. F., Boguski, M. S., Gish, W., & Wootton, J. C. (1994). Issues in searching molecular sequence databases. *Nature Genet.*, *6*, 119-129.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, *25*, 3389–3402.

Altschul, S., Gish, W., Miller, W., Myers, E. W., & Lipmanl, D. J. (1990). Basic Local Alignment Search Tool. *Journal of Moelcular Biology*, *215*, 403-419.

Alvarez, M. E., Pennell, R. I., Meijer, P.-J., Ishikawa, A., Dixon, R. A., & Lamb, C. (1998). Reactive oxygen intermediates mediate a systemic signal network in the establishment of plant immunity. *Cell*, *92*(6), 773-784.

Anderson, L. (1971). Chloroplast and cytoplasmic enzymes. II. Pea leaf triose phosphate isomerases. *Biochim. Biophys. Acta*, *235*, 237-244.

Anderson, R. J., Levitt, G. J., & Share, A. (1996). Experimental investigations for the mariculture of Gracilaria in Saldanha Bay, South Africa. *Journal of Applied Phycology*, *8*, 421-430.

Anderson, R. J., Levitt, G. J., & Simons, R. H. (1993). The role of herbivores in the collapse of the Gracilaria resource at Saldanha Bay, South Africa. *Hydrobiology*, *260/261*, 285-290.

Anderson, R. J., Simons, R. H., & Jarman, N. G. (1989). Commercial seaweeds in Southern Africa: A review of utilization and research. *S. Afr. J. mar. Sci.*, *8*, 277-299.

Anderson, R. J., Bolton, J. J., Molly, F. J., & Rotmann, K. W. G., R. J. (2003). Commercial seaweeds in southern Africa, 1-12. Oxford University Press, Oxford.

Antão, C. M., & Malcata, F. X. (2005). Plant serine proteases : biochemical , physiological and molecular features. *Plant Physiology and Biochemistry*, *43*, 637-650.

Apel, K., & Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annual review of plant biology*, *55*, 373-99.

Aranjuelo, I., Molero, G., Erice, G., Avice, J. C., & Nogues, S. (2010). Plant physiology and proteomics reveals the leaf response to drought in alfalfa (Medicago sativa L.). *Journal of Experimental Botany*, 1-13.

Aravind, L., Galperin, M. Y., & Koonin, E. V. (1998). The HD domain defines a new superfamily of metal-dependent phos-phohydrolases. *Trends Biochem. Sci.*, *23*, 127-129.

Ariki, S., Koori, K., Osaki, T., Motoyama, K., Inamori, K.-I., & Kawabata, S.-I. (2004). A serine protease zymogen functions as a pattern-recognition receptor for lipopolysaccharides. *Proc. Natl. Acad. Sci. USA*, *101*, 953-958.

Armbrust, V. E., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., et al. (2004). The genome of the diatom Thalassiosira pseudonana: Ecology, evolution and metabolism. *Science*, *306*, 79-86.

Armisen, R. (1995). World-wide use and importance of Gracilaria. *Journal of Applied Phycology*, *7*, 231-243.

Arnér, E. S. J., & Holmgren, A. (2000). Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem.*, *267*, 6102–6109.

Aro, E. M., McCaffery, S., & Anderson, J. M. (1993). Photoinhibition and D1 protein degradation in peas acclimated to different growth irradiances. *Plant Physiology*, *103*(3), 835-843.

Ask, E. I., & Azanza, R. V. (2002). Advances in cultivation technology of commercial eucheumatoid species: a review with suggestions for future research. *Aquaculture*, *206*, 257-277.

Atkins, C. A., Pate, J. S., & Sharcey, P. J. (1975). Asparagine metabolism-key to the nitrogen nutrition of developing legume seeds. *Plant Physiol.*, *56*, 807-812.

Audenaert, K., Pattery, T., Cornelis, P., & Höfte, M. (2002). Induction of systemic resistance to Botrytis cinerea in tomato by Pseudomonas aeruginosa 7NSK2: role of salicylic acid, pyochelin, and pyocyanin. *Molecular Plant-Microbe Interactions*, *15*(11), 1147-56.

Azevedo, H., Lino-neto, T., & Tavares, R. M. (2003). An Improved Method for High-Quality RNA Isolation From Needles of Adult Maritime Pine Trees. *Plant Molecular Biology Reporter*, *21*(December), 333-338.

Bachem, C. W., Der Hoeven, R. S. van, Bruijn, S. M. de, Vreugdenhil, D., Zabeau, M., & Visser, R. G. (1996). Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant J.*, *9*, 745-753.

Baldan, B., Andolfo, P., Navazio, L., Tolomio, C., & Mariani, P. (2001). Cellulose in algal cell wall : an " in situ " localization. *Eur. J. Histochem., 45*, 51-56.

Baldi, P., & Long, A. D. (2001). A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. *Bioinformatics, 17*, 509-519.

Baldy, P., Jacquot, J.-P., Lavergne, D., & Champigny, M. L. (1989). Corn phosphoglycolate phosphatase: Modulation of activity by pyridine nucleotides and adenylate energy charge. *Photosynthesis Research, 22*(147-155).

Barbier, G., Oesterhelt, C., Larson, M. D., Halgren, R. G., Wilkerson, C., Garavito, R. M., et al. (2005). Comparitive genomics of two closley related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbo. *Plant Physiol., 137*, 460-474.

Barnaby, N. G., He, F., Liu, X., Wilson, K. A., & Tan-Wilson, A. (2004). Light-responsive subtilisin-related protease in soybean seedling leaves. *Plant Physiol. Biochem., 42*, 125-134.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., et al. (2002). The Pfam protein families database. *Nucleic Acids Res., 30*(1), 276-280.

Baumann, U., & Juttner, J. (2002). Plant thioredoxins: the multiplicity conundrum. *Cell. Mol. Life Sci., 59*, 1042-1057.

Beers, E. P., Woffenden, B. J., & Zhao, C. (2000). Plant proteolytic enzymes: possible roles during programmed cell death. *Plant Mol. Biol., 44*, 399–415.

Begley, T. P., Xi, J., Kinsland, C., Taylor, S., & McLafferty, F. (1999). The enzymology of sulfur activation during thiamine and biotin biosynthesis. *Current Opinion in Chemical Biology, 3*, 623-629.

Belanger, F. C., & Ogren, W. L. (1987). Phosphoglycolate phosphatase : purification and preparation of antibodies. *Photosynthesis Research, 14*, 3-13.

Bellanger, F., Verdus, M. C., Henocq, V., & Christiaen, D. (1990). Determination of the composition of the fibrillar part of *Gracilaria verrucosa* (Gracilariales, Rhodophyta) cell wall in order to prepare protoplasts. *Hydrobiologia, 204-205*(1), 527-531.

Benes, V., & Muckenthaler, M. (2003). Standardization of protocols in cDNA microarray analysis. *Trends in Biochemical Sciences, 28*(5), 244-249.

Bernhard, W. R., Thoma, S., Botella, J., & Somerville, C. R. (1991). Isolation of a cDNA Clone for spinach lipid transfer protein and evidence that the protein is synthesized by the secretory pathway. *Plant Physiol., 95*, 164-170.

Berninsone, P. M., & Hirschberg, C. B. (2000). Nucleotide sugar transporters of the Golgi apparatus. *Current Opinion in Structural Biology*, *10*, 542-547.

Besse, I., & Buchanan, B. B. (1997). Thioredoxin-linked plant and animal processes: the new generation. *Bot. Bull. Academia Sinica (Taipei).*, *38*, 1-11.

Besse, I., Wong, J. H., Kobrehel, K., & Buchanan, B. B. (1996). Thiocalsin: a thioredoxin-linked, substrate-specific protease dependent on calcium. *Proc. Natl. Acad. Sci. USA*, *93*, 3169-3175.

Bhagwat, A. S. (2004). DNA-cytosine deaminases : from antibody maturation to antiviral defense. *DNA Repair*, *3*, 85-89.

Bird, C. J., Rice, E. L., Murphy, C. A., Lui, Q. Y., & Ragan, M. A. (1990). Nucleotide sequences of 18S ribosomal RNA genes from the red algae *Gracilaria tikvahiae* McLachlan, *Gracilaria verrucosa* (Hudson) Papenfuss and Gracilariopsis sp. *Nucleic Acids Research*, *18*, 4023-4024.

Bird, K. T., & Benson, P. H. (1987). *Seaweed cultivation for renewable resources.* (pp. 241-276). Amsterdam, New York: Elsevier.

Bishop, A. C., Xu, J., Johnson, R. C., Schimmel, P., & Crécy-Lagard, V. de. (2002). Identification of the tRNA-Dihydrouridine synthase family. *Journal of Biological Chemistry*, *277*, 25090-25095.

Blein, J.-P., Coutos-Thévenot, P., Marion, D., & Ponchet, M. (2002). From elicitins to lipid-transfer proteins : a new insight in cell signalling involved in plant defence mechanisms. *Trends in Plant Science*, *7*(7), 293-296.

Blomberg, A., & Adler, L. (1989). Roles of Glycerol and Glycerol-3-Phosphate Dehydrogenase (NAD +) in Acquired Osmotolerance of Saccharomyces cerevisiae. *Journal of Bacteriology*, *171*(2), 1087-1092.

Boesch, D. F. (1999). The role of science in ocean governance. *Ecological Economics*, *31*, 189-198.

Bork, P., & Gibson, T. J. (1996). Applying motif and profile searches. *Methods Enzymol.*, *266*, 162-184.

Borsani, G., Ballabio, A., & Banfi, S. (1998). A practical guide to orient yourself in the labyrinth of genome databases. *Human Mol. Genet.*, *7*(10), 1641-1648.

Bowler, C., & Chua, N.-H. (1994). Emerging themes of plant signal transduction. *Cell*, *6*, 1529-1541.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, *72*, 248-254.

Branch, G., & Branch, M. (1981). *The Living Shores of Southern Africa* (pp. 251-261). Cape Town: C. Struik.

Brazma, A., & Vilo, J. (2000). Gene expression data analysis. *FEBS Lett.*, *480*, 17-24.

Brehelin, C., Meyer, E. H., Souris, J. P. de, Bonnard, G., & Meyer, Y. (2003). Resemblance and dissemblance of Arabidopsis type II peroxiredoxins: similar sequences for divergent gene expression, protein localization, and activity. *Plant Physiol .*, *132*, 2045-2057.

Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D. H., Johnson, D., et al. (2000). Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.*, *18*, 630-634.

Brown, A. D. (1978). Compatible solutes and extreme water stress in eukaryotic micro-organisms. In H. Rose & J. G. Morris (Eds.), *Advances in microbial physiology* (17th ed.). New York: Academic Press, Inc.

Brown, P. O., & Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. *Nature Genetics Supplement*, (21), 33–37.

Buckner, B., Johal, G. S., & Janick-buckner, D. (2000). Cell death in maize. *Physiologia Plantarum*, *108*, 231-239.

Budde, R. J. A., & Randall, D. D. (1990). Protein kinases in higher plants. In D. J. Moore & W. F. Boss (Eds.), *Inositol Metabolism in Plants* (pp. 351-367). New York: Liss.

Burger, K. N. J. (2000). Greasing membrane fusion and fission machineries. *Traffic*, *1*, 605-613.

Bustin, S. A. (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology*, *29*, 23-39.

Bustin, S. A., & Nolan, T. (2004). Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal of Biomolecular Techniques*, *15*, 155-166.

Cammue, B. P. A., Thevissen, K., Hendriks, M., Eggermont, K., Goderis, I. J., Proost, P., et al. (1996). A potent antimicrobial protein from onion seeds showing sequence homology to plant lipid transfer proteins. *Plant Physiology*, *109*, 445-455.

Campillo, E. del. (1999). Multiple endo-1,4-beta-D-glucanase (cellulase) genes in Arabidopsis. *Current Topics in Developmental Biology*, *46*, 39-61.

Carlton, J. T., & Cohen, A. N. (2003). Episodic global dispersal in shallow water marine organisms : the case history of the European shore crabs Carcinus maenas and C . aestuarii. *Journal of Biogeography*, *30*, 1809-1820.

Carlton, J. T., & Geller, J. B. (1993). Ecological Roulette: The global transport of nonindigenous marine organisms. *Science, 261*(5117), 78-82.

Carpentier, S. C., Panis, B., Vertommen, A., Swennen, R., Sergeant, K., Renaut, J., et al. (2008). Proteome analysis of non-model plants: a challenging but powerful approach. *Mass Spectrometry Reviews, 27,* 354-377.

Causton, H. C., Quakenbush, J., & Brazma, A. (2003). *Microarray Gene Expression Data Analysis. A Beginner's Guide.* (H. C. Causton, J. Quakenbush, & A. Brazma, Eds.) (p. 160). Blackwell Publishing.

Chae, H. Z., Chung, S. J., & Rhee, S. G. (1994). Thioredoxin-dependent peroxide reductase from yeast. *Journal of Biological Chemistry, 269*, 27670-27678.

Chae, H. Z., Kang, S. W., & Rhee, S. G. (1999). Isoforms of mammalian peroxiredoxin that reduce peroxides in presence of thioredoxin. *Methods Enzymol., 300*, 219-226.

Chae, H. Z., Robison, K., Poole, L. B., Church, G., Storz, G., & Rhee, S. G. (1994). Cloning and sequencing of a thiol-specific antioxidant from mammalian brain: alkyl hydroperoxide reductase and thiol-specific antioxidant define a large family of antioxidant enzymes. *Proc. Natl. Acad. Sci. USA, 91*, 7017–7021.

Chae, H. Z., Robison, K., Poole, L. B., Church, G., Storz, G., & Rhee, S. G. (1994). Cloning and sequencing of thiol-specific antioxidant from mammalian brain: alkyl hydroperoxide reductase and thiol-specific antioxidant define a large family of antioxidant enzymes. *Proc. Natl. Acad. Sci. USA, 91*, 7017–7021.

Chan, C. X., Teo, S. S., Ho, C. L., Othman, R. Y., & Phang, S. M. (2004). Optimization of RNA extraction from *Gracilaria changii* (Gracilariales, Rhodophyta). *Journal of Applied Phycology, 16*, 297-301.

Chan, C., Teo, S., Ho, C., Othman, R. Y., & Phang, S. (2004). Optimisation of RNA extraction from *Gracilaria changii* (Gracilariales, Rhodophyta). *Journal of Applied Phycology, 16*, 297-301.

Chan, C. X., Ho, C. L. & Phang, S. W. (2006) Trends in seaweed research. *TRENDS in Plant Science, 11 (4),* 165-166.

Chandra, S., & Low, P. S. (1995). Role of phosphorylation in elicitation of the oxidative burst in cultured soybean cells. *Proc. Natl. Acad. Sci. USA, 92*, 4120-4123.

Chapman, V. J. (1950). *Seaweeds and their uses* (p. 287). London: Methuen & Co.

Chen, G., Gharib, T. G., Huang, C. C., Taylor, J. M., Misek, D. E., Kardia, S. L., et al. (2002). Discordent protein and m RNA expression in lung adenocarcinomas. *Mol. Cell Proteomics, 1*, 304-313.

Chen, Y., Dougherty, E. R., & Bittner, M. L. (1997). Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomed. Optics*, *2*, 364–374.

Chen, Z.-Y., Brown, R. L., Damann, K. E., & Cleveland, T. E. (2007). Identification of Maize Kernel Endosperm Proteins Associated with Resistance to Aflatoxin Contamination by Aspergillus flavus. *Phytopathology*, *97*(9), 1094-103.

Cheong, J.-J., & Choi, Y. D. (2003). Methyl jasmonate as a vital substance in plants. *Trends in Genetics*, *19*, 409-413.

Chiang, Y. M. (1981). Cultivation of Gracilaria (Rhodophyta, Gigartinales) in Taiwan. *Proc. int. Seaweed Symposium*, *10*, 569-574.

Choi, J.-Y., Wu, W.-I., & Voelker, D. R. (2005). Phosphatidylserine decarboxylases as genetic and biochemical tools for studying phospholipid traffic. *Analytical Biochemistry*, *347*, 165-175.

Chopin M., T. and S., & Steele Turekian, K. K. and Thorpe, S. A., J. H. (2009). Seaweeds and their mariculture (Second., pp. 4477-4487). Elsevier.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O., et al. (1998). The transcriptional program of sporulation in budding yeast. *Science*, *282*, 699–705.

Chuaqui, R. F., Bonner, R. F., Best, C. J. M., Gillespie, J. W., Flaig, M. J., Hewitt, M. S., et al. (2002). Post-analysis follow-up and validation of microarray experiments. *Nature Genetics Supplement*, *32*, 509-514.

Claverie, J. M. (1999). Computational methods for the identification of differential and coordinated gene expression. *Hum Mol Genet*, *8*, 1821-1832.

Claverie, J.-M. (1999). Computational methods for identi cation of differential and coordinated gene expression. *Human Mol. Genet.*, *8*, 1821–1832.

Coffeen, W. C., & Wolpert, T. J. (2004). Purification and characterization of serine proteases that exhibit caspase-like activity and are associated with programmed cell death in Avena sativa. *Plant Cell*, *16*, 857–873.

Cohen, P. (1989). Structure and regulation of protein phosphatases. *Annu. Rev. Biochem.*, *58*, 453-508.

Cohn, J., Sessa, G., & Martin, G. B. (2001). Innate immunity in plants. *Current Opinion in Immunology*, *13*(1), 55-62.

Cole, R. G., & Babcock, R. C. (1996). Mass mortality of a dominant kelp (Laminariales) at Goat Island, North-Eastern New Zealand. *Mar. Freshwater Res.*, *47*, 907–911.

Collet, J.-F., Stroobant, V., Pirard, M., Delpierre, G., & Van Schaftingen, E. (1998). A new class of phosphotransferases phosphorylated on an aspartate residue in an amino-terminal DXDX(T/V) motif. *Journal of Biological Chemistry, 273*, 14107–14112.

Collén, J., Guisle-marsollier, I., Léger, J. J., & Boyen, C. (2007). Response of the transcriptome of the intertidal red seaweed *Chondrus crispus* to controlled and natural stresses. *New Phytologist*, 1-11.

Collén, J., Herve, C., Guisle-Marsollier, J., Leger, J., & Boyen, C. (2006). Expression profiling of *Chondrus crispus* (Rhodophyta) after exposure to methyl jasmonate. *J. Exp. Bot., 57*, 3869-3881.

Corona, D. F., Langst, G., Clapier, C. R., Bonte, E. J., Ferrari, S., Tamkun, J. W., et al. (1999). ISWI is an ATP-dependent nucleosome remodeling factor. *Molecular Cell., 3*, 239-245.

Correa, J. A., & Craigie, J. S. (1991). Algal Pathology. In M. Garcia-Reina, G. And Pedersen (Ed.), *Seaweed Cellular Biotechnology, Physiology and Intensive Cultivation* (pp. 67-82). COST 48, Universidad de Las Palma de Gran Canaria.

Cosgrove, D. J. (2000). Expansive growth of plant cell walls. *Plant Physiol. Biochem., 38*, 109-124.

Cosse, A., Leblanc, C., & Potin, P. (2008). Dynamic defense of marine macroalgae against pathogens: from early activated to gene-regulated responses. *Advances in Botanical Research, 46*, 221-266.

Costanza, R., Andrade, F., Antunes, P., Den Belt, M. van, Boersma, D., Boesch, D. F., et al. (1998). Principles for Sustainable Governance of the Oceans. *Science, 281*(5374), 198-199.

Cox, B., Kislinger, T., & Emili, A. (2005). Integrating gene and protein expression data: pattern analysis and profile mining. *Methods, 35*, 303–314.

Cox, G. W., Beaudet, M. P., Agnew, J. Y., & Ruth, J. L. (2004). Possible sources of dye-related correlation bias in two-colour DNA microarray assays. *Analytical Biochemistry, 331*, 243-254.

Critchley, A. (1993). Gracilaria (Rhodophyta, Gracilariales): an economically important agarophyte. In M. Ohno & A. Critchley (Eds.), *Seaweed cultivation and marine ranching* (pp. 89-112). Japan International Cooperation Agency (JICA).

Critchley, C., & Russell, A. W. (1994). Photoinhibition of photo- synthesis in vivo: the role of protein turnover in photosystem II. *Physiologia Plantarum, 92*, 188-196.

Cui, X., & Churchill, G. A. (2003). Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology, 4*(4), Article 210.

Cumming, G., Fidler, F., & Vaux, D. L. (2007). Error bars in experimental biology. *The Journal of Cell Biology*, *177*(1), 7-11.

Cummings, C. A., & Relman, D. A. (2000). Using DNA microarrays to study host-microbe interactions. *Emerging Infectious Diseases*, *6*(5 (September–October)), 513-525.

Dagert, M., & Ehrlich, S. D. (1979). Prolonged incubation in calcium chloride improves competence of Escherichia coli cells. *Gene*, *6*, 23-28.

Dallas, P. B., Gottardo, N. G., Firth, M. J., Beesley, A. H., Hoffmann, K., Terry, P. A., et al. (2005). Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR: how well do they correlate? *BMC Genomics*, *6*, 59.

Dalton, T. P., Shertzer, H. G., & Puga, A. (1999). Regulation of gene expression by reactive oxygen. *Annual review of pharmacology and toxicology*, *39*, 67-101.

Dayer, R., Fischer, B. B., Eggen, R. I. L., & Lemaire, S. D. (2008). The peroxiredoxin and glutathione peroxidase families in *Chlamydomonas reinhardtii*. *Genetics*, *179*, 41-57.

De Oliveira, E. C., & Plastino, E. M. (1994). Gracilariaceae. In I. Akatsuka (Ed.), *Biology of Economic Algae* (pp. 185-226). The Hague, Netherlands: SPB Academic Publishing.

Degrelle, S. A., Hennequet-Antier, C., Chiapello, H., Piot-Kaminski, K., Piumi, F., Robin, S., et al. (2008). Amplification biases: possible differences among deviating gene expressions. *BMC genomics*, *9*(i), 46.

Delaney, T. P., Uknes, S., Vernooij, B., Weymann, K., Negrotto, D., Gaffney, T., et al. (1994). A central role of salicylic acid in plant disease resistance. *Science*, *266*(5188), 1247-1250.

Delledonne, M., Zeier, J., Marocco, A., & Lamb, C. (2001). Signal interactions between nitric oxide and reactive oxygen intermediates in the plant hypersensitive disease resistance response. *Proc. Natl. Acad. Sci. USA*, *98*, 13454–13459.

Demmig-Adams, B., & Adams, W. W. (1992). Photoprotection and Other Responses of Plants to High Light Stress. *Annual Review of Plant Physiology and Plant Molecular Biology*, *43*, 599-626.

Der Hoorn, R. A. L. van, De Wit, P. J. G. M., & Joosten, M. H. A. J. (2002). Balancing selection favors guarding resistance proteins. *Trends Plant Sci.*, *7*, 67-71.

Der Hoorn, R. A. L. van, & Jones, J. D. G. (2004). The plant proteolytic machinery and its role in defence. *Current opinion in plant biology*, *7*, 400-7.

Derelle, E., Ferraz, C., Rombauts, S., Rouze, P., Worden, A. Z., Robbens, S., et al. (2006). Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci. USA*, *103*, 11647-116652.

DeRisi, J., Penland, L., Brown, P. O., Bittner, M. A., & Et, A. (1996). Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nature Genetics*, *14*, 457–460.

Desikan, R., A, S., Hancock, J. T., & Neill, S. J. (2001). Regulation of the Arabidopsis transcriptome by oxidative stress. *Plant Physiol.*, *127*(September), 159-172.

Devoto, A., Muskett, P. R., & Shirasu, K. (2003). Role of ubiquitination in the regulation of plant defence against pathogens. *Current Opinion in Plant Biology*, *6*, 307-311.

Dietz, K.-J. (2003). Plant peroxiredoxins. *Annual Review of Plant Biology*, *54*, 93-107.

Donson, J., Fang, Y., Espiritu-Santo, G., Xing, W., Salazar, A., Miyamoto, S., et al. (2002). Comprehensive gene expression analysis by transcript profiling. *Plant Mol. Biol.*, *48*, 75–97.

Dopazo, J., Zanders, E., Dragoni, I., Amphlett, G., & Falciani, F. (2001). Methods and approaches in the analysis of gene expression data. *Journal of Immunological Methods*, *250*, 93-112.

Dubreil, L., Gaborit, T., Bouchet, B., Gallant, D. J., Broekaert, W. F., Quillien, L., et al. (1998). Spatial and temporal distribution of the major isoforms of puroindolines (puroindoline-a and puroindoline-b) and non specific lipid transfer protein (ns-LTP1e1) of Triticum aestivum seeds. Relationships with their in vitro antifungal properties. *Plant Science*, *138*, 121-135.

Dudoit, S., Yang, Y. H., Callow, M. J., & Speed, T. P. (2000). *Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments*. *Statistica Sinica* (Vol. 12, pp. 111-140). Stanford, CA.

Duggan, D. J., Bittner, M., Chen, Yidong, Meltzer, P., & Trent, Jeffrey M. (1999). Expression profiling using cDNA microarrays. *Nature Genetics Supplement*, *21*(January), 10-14.

Dunlop, P. C., & Roon, R. J. (1975). L-Asparaginase of Saccharomyces cerevisiae : an extracellular enzyme. *Journal of Bacteriology*, *122*(3), 1017-1024.

Dunn, B. M. (2001). Determination of protease mechanism. In R. Beynon & J. S. Bond (Eds.), *Plant proteolytic enzymes - a practical approach* (p. 77–79). New York: Oxford University Press.

Durnford, D. G., & Falkowski, P. G. (1997). Chloroplast redox regulation of nuclear gene transcription during photoacclimation. *Photosynthesis Research, 53*(2-3), 229-241.

Eckardt, N. A. (2004). Aminotransferases Confer "Enzymatic Resistance" to Downy. *The Plant Cell, 15*, 1-4.

Edgley, M., & Brown, A. D. (1983). Yeast water relations: physiological changes induced by solute stress in Saccharomyces cerevisiae and Saccharomyces rouxii. *J. Gen. Microbiology, 129*, 3453-3463.

Eisen, M. B., Spellman, P. T., Brown, P. O., & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA, 95*, 14863-14868.

Eisen, M., & Brown, P. O. (1999). DNA arrays for analysis of gene expression. *Methods in Enzymology, 303,* 179-205.

Ellis, C., Turner, J. G., & Devoto, A. (2002). Protein complexes mediate signalling in plant responses to hormones, light, sucrose and pathogens. *Plant Mol Biol, 50*, 971-980.

Elstner, E. F., Oswald, W., & Schneider, I. (1996). *Phytopathologie* (p. 328). Germany: Spektrum Akademischer Verlag.

Engel, S., Jensen, P. R., & Fenical, W., S. (2002). Chemical ecology of marine microbial defense. *Journal of Chemical Ecology, 28*, 1971-1985.

Epstein, P. R., Sherman, K., Spanger-Siegfried, E., Langston, A., Prasad, S., & McKay, B. (1998). *Marine ecosystems: emerging diseases as indicator of change.* Washington, DC.

Errede, B., Cade, R. M., Yashar, B. M., Kamada, Y., Levin, D. E., Irie, K., et al. (n.d.). Dynamics and organization of MAP kinase signal pathways. *Mol. Reprod. Dev., 42*, 477-485.

Farley, C. A. (1992). Dispersal of pathogens, parasites, pests, predators and competitors. In A. Rosenfield & R. Mann (Eds.), *Dispersal of Living Organisms into Aquatic Ecosystems* (p. 139–154). College Park, MD: Maryland Sea Grant Ford SE.

Fiehn, O., Kloska, S., & Altmann, T. (2001). Integrated studies on plant biology using multiparallel techniques. *Curr. Opin. Biotechnol., 12*, 82-86.

Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R. N., & Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. *Nat. Biotechnol., 18*, 1157-1161.

Finkemeier, I., Goodman, M., Lamkemeyer, P., Kandlbinder, A., Sweetlove, L. J., & Dietz, K. J. (2005). The mitochondrial type II peroxiredoxin F is essential for

redox homeostasis and root growth of *Arabidopsis thaliana* under stress. *Journal of Biological Chemistry, 280*, 12168-12180.

Fontanini, D., & Jones, B. L. (2002). SEP-1—a subtilisin-like serine endopeptidase from germinated seeds of Hordeum vulgare L. cv,Morex. *Planta*, *215*, 885–893.

Foyer, C. H., & Noctor, G. (2000). Tansley review no. 112, oxygen processing in photosynthesis: regulation and signaling. *New Phytol.*, *146*, 359–388.

Freeman, W. M., Walker, S. J., & Vrana, K. E. (1999). Quantitative RT-PCR: pitfalls and potential. *Biotechniques*, *26*, 112-125.

Frei, E., & Preston, R.D. (1964). Non-cellulosic structural polysaccharides in algal cell wall. II. Association of xylan and mannan in *Porphyra umbilicalis*. *Proc. R. Soc. London Ser. B, 160*, 314-327.

Friedlander, M. A., & Gunkel, W. (1992). Factors leading to thallus disintegration and the control of these factors in Gracilaria sp. *Proceedings of the 4th German-Israeli Status Seminar, EAS specia*, 221-243.

Friedlander, M., & Levy, I. (1995). Cultivation of Gracilaria in outdoor tanks and ponds. *Journal of Applied Phycology, 7*, 315-324.

Fujita, M., Fujita, Y., Noutoshi, Y., Takahashi, F., Narusaka, Y., Yamaguchi-Shinozaki, K., et al. (2006). Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Current opinion in plant biology*, *9*(4), 436-42.

Fujita, Y. (1973). Maceration of laver frond by enzymes of a marine bacterium that produces red-spots on the culture bed of makonbu *Laminaria japonica* cultivation. *Bull. Jpn. Soc. Sci. Fish., 54*, 665–672.

Fujita, Y., Zenitani, B., Nakao, Y., & Matsubara, T. (1972). Bacterial studies on diseases of cultured laver. II. Bacteria associated with diseased laver. *Bulletin of the Japanese Society of Scientific Fisheries*, *38*, 565-569.

Fujiwara, S., Tanaka, N., Kaneda, T., Takayama, S., Isogai, A., & Che, F. (2004). Rice cDNA microarray-based gene expression profiling of the response to flagellin perception in cultured rice cells. *Molecular Plant-Microbe Interactions, 17*(9), 986-998.

Gabaldón, T., Snel, B., Zimmeren, F. van, Hemrika, W., Tabak, H., & Huynen, M. A. (2006). Origin and evolution of the peroxisomal proteome. *Biology Direct, 1*, 8.

Ganzhorn, J., Rohovec, J. S., & Fryer, J. L. (1992). Dissemination of microbial pathogens through introductions and transfer of finfish. In A. Rosenfeld & R. Mann (Eds.), *Dispersal of living organisms into aquatic ecosystems.* (pp. 175-192). University of Maryland Press, College Park.

Garcia-Olmedo, F., Molina, A., Segura, A., & Moreno, M. (1995). The defensive role of nonspecific proteins in plants. *Trends in Microbiology*, *3*(2), 72-74.

Gautier, M. F., Lullien-Pellerin, V., Lamotte-Guery, F. de, Guirao, A., & Joudrier, P. (1998). Characterization of wheat thioredoxin h cDNA and production of an active Triticum aestivum protein in Escherichia coli. *Eur. J. Biochem.*, *252*, 314–324.

Gebrekiros, S. T. (2003). Characterization of nitrogen stress response genes of the marine alga *Gracilaria gracilis*. Cape Town: University of Cape Town.

Gelhaye, E., Rouhier, N., Navrot, N., & Jacquot, J. P. (2005). The plant thioredoxin system. *Cellular and Molecular Life Sciences*, *62*, 24- 35.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., et al. (2004). Bioconductor : open software development for computational biology and bioinformatics. *Genome Biology*, *5*, R80.

Gibson, U. E., Heid, C. A., & Williams, P. M. (1996). A novel method for real time quantitative RT-PCR. *Genome Res.*, *6*, 995-1001.

Glickman, M. H., & Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.*, *82*, 373-428.

Good, A. G., & Muench, D. G. (1992). Purification and characterization of an anaerobically induced alanine aminotransferase from barley roots. *Plant Physiology*, *99*, 1520-1525.

Gopalakrishna, R., & Janken, S. (2000). Protein kinase C signaling and oxidative stress. *Free Radical Biology and Medicine*, *28*(9), 1349-1361.

Goryachev, A. B., Macgregor, P. F., & Edwards, A. M. (2001). Unfolding of Microarray Data. *Journal of Computational Biology*, *8*(4), 443-461.

Govender, K. (2006). *Characterisation of XvPrx2, a type II peroxiredoxin isolated from the resurrection plant Xerophyta viscosa (Baker).*

Greenberg, J. T. (1997). Programmed cell death in plant-pathogen interactions. *Annual Review of Plant Physiology and Plant Molecular Biology*, *48*, 525-545.

Gribskov, M., McLachlan, M., & Eisenberg, D. (1987). Profile analysis: detection of distantly related proteins. *Proc. Natl. Acad. Sci. USA*, *84*, 4355–4358.

Guerbette, F., Grosbois, M., Jolliot-Croquin, A., Kader, J-C., & Zachowski, A. (1999). Comparison of lipid binding and transfer properties of two lipid transfer proteins from plants. *Biochemistry*, *38*, 14131-14137.

Guo, A., Remiers, P. J., & Leach, J. E. (1993). Effect of light on incompatible interactions between Xanthomonas oryzae pv. oryzae and rice. *Physiol Mol Plant Pathol.*, *42*, 413-425.

Hadfield, K. A., & Bennett, A. B. (1998). Update on Cell Walls Polygalacturonases : Many Genes in Search of a Function 1. *Plant Physiology*, *117*, 337-343.

Halliwell, B., & Gutteridge, J. M. C. (1999). *Free Radicals in Biology and Medicine.* (pp. 1-104). New York: Oxford University Press Inc.

Hamilton, J. M. U., Simpson, D. J., Hyman, S. C., Ndimba, B. G., & Slabas, A. R. (2003). Ara12 subtilisin-like protease from *Arabidopsis thaliana*: purification, substrate specificity and tissue localization. *Biochem. J.*, *370*, 57-67.

Hammond-Kosack, K. E., & Jones, J. D. G. (1996). Resistance Gene-Dependent Plant Defense Responses. *The Plant Cell*, *8*(October), 1773-1791.

Hanisak, M. D. (1981). Methane production from the red seaweed *Gracilaria tikvahiae*. *Proc. int. Seaweed Symposium*, *10*, 681-692.

Harvell, C. D., Kim, K., Burkholder, J. M., Colwell, R. R., Epstein, P. R., Grimes, D. J., et al. (2009). Emerging Marine Diseases — Climate links and anthropogenic factors. *Science*, *285*, 1505-1510.

He, Q., Dolganov, N., Bjo, O., & Grossman, A. R. (2001). The High Light-inducible Polypeptides in Synechocystis PCC6803. *The Journal of Biological Chemistry*, *276*(1), 306 -314.

Heath, M. C. (2000). Hypersensitive response-related death. *Plant Mol. Biol.*, *44*, 321-334.

Hedge, P., Qi, R., Abernathy, K., Gay, C., Dharap, S., Gaspard, R., et al. (2000). A concise guide to cDNA microarray analysis. *Biotechniques*, *29*, 548-562.

Higuchi, R., Fockler, C., Dollinger, G., & Watson, R. (1993). Kinetic PCR analysis - real time monitoring of DNA amplification. *Bio/Technology*, *11*, 1026-1030.

Hirota, K., Murata, M., Sachi, Y., Nakamura, H., Takeuchi, J., Mori, K., et al. (1999). Distinct roles of thioredoxin in the cytoplasm and in the nucleus – a two-step mechanism of redox regulation of transcription factor NF-kappa B. *J. Biol. Chem.*, *274*, 27891–27897.

Hirschberg, C. B., Robbins, P. W., & Abeijon, C. (1998). Transporters of nucleotide sugars, atp, and nucleotide sulfate in the endoplasmic reticulum and golgi apparatus. *Annual Review of Biochemistry*, *67*, 49-69.

Ho, C-L., Teoh, S., Teo, S-S., Rahim, R. A. & Phang, S-W. (2009). Profiling of the Transcriptome of *Gracilaria changii* in Response to Light Depravation. *Mar. Biotechnol., 11*, 513 - 519.

Hofmann, B., J., H. H., & Flohé, L. (2002). Peroxiredoxins. *Biological Chemistry*, *383*(3-4), 347-346.

Hoheisel, J. D. (2006). Microarray technology: beyond transcript profiling and genotype analysis. *Nature Reviews. Genetics*, *7*(3), 200-10.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Statist.*, *6*, 65-70.

Holmgren, A. (1989). Thioredoxin and glutaredoxin systems. *J. Biol. Chem.*, *264*, 13963–13966.

Holtorf, H., Guitton, M. C., & Reski, R. (2002). Plant functional genomics. *Naturwissenschaften*, *89*, 235-249.

Horling, F., Lamkemeyer, P., Ko, J., Finkemeier, I., Kandlbinder, A., Baier, M., et al. (2003). Divergent Light-, Ascorbate-, and Oxidative Stress-Dependent Regulation of Expression of the Peroxiredoxin Gene Family in Arabidopsis 1. *Plant Physiology*, *131*(January), 317-325.

Hu, L., Wang, J., Baggerly, K., Wang, H., Fuller, G. N., Hamilton, S. R., et al. (2002). Obtaining reliable information from minute amounts of RNA using cDNA microarrays. *BMC Genomics*, *3*, 16.

Huber, S. C., Huber, J. L., & McMichael, R. W. (1994). Control of plant enzyme activity by reversible protein phosphorylation. *Int. Rev. Cytol.*, *149*, 47-98.

Hurtado, H. Q., Critchely, A. T., Trespoey, A., & Bleicher-L'Honneur, G. (2006). Occurrence of Polysiphonia epiphytes in Kappaphycus farms at Calagas Is., Camrines Norte, Philippines. *Journal of Applied Phycology*, *18*, 301-306.

Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. *J. Comput. Graph. Statist.*, *5*, 299-314.

Immenscuh, S., & Baumgart-Vogt, E. (2005). Peroxiredoxins. oxidative stress, and cell proliferation. *Antioxidants and Redox Signalling*, *7*(5 & 6), 768-777.

Ingle, R. A., Schmidt, U. G., Farrant, J. M., Thomson, J. A., & Mundree, S. G. (2007). Proteomic analysis of leaf proteins during dehydration of the resurrection plant Xerophyta viscosa. *Plant, Cell and Environment*, *30*, 435–446.

Isaac, W. E. (1956). The ecology of *Gracilaria confervoides* (L.) Grev. in South Africa with special reference to its ecology in the Saldanha-Langebaan Lagoon. In T. Braarud & N. A. Sorensen (Eds.), *Second Proceedings of the International Seaweed Symposium, Trondheim* (pp. 173-185). Pergamon, London.

Ishikawa, Y., & Saga, N. (n.d.). The diseases of economically valuable seaweeds and pathology in Japan. In S. Miyachi, I. Karube, & Y. Ishda (Eds.), *Current Topics in Marine Biotechnology*. Tokyo: Fuji Technology Press Ltd.

Iyer, R. (2002). *Systematics of the Gracilariaceae (Rhodophyta) of southern Africa.*

Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C. F., et al. (1999). The transcriptional program in the response of humānbroblasts to serum. *Science, 283*, 83–87.

Jabs, T. (1999). Reactive oxygen intermediates as mediators of programmed cell death in plants and animals. *Biochem. Pharmacol., 57*, 231-245.

Jaffray, A. E., Anderson, R. J., & Coyne, V. E. (1997). Investigation of bacterial epiphytes of the agar-producing red seaweed *Gracilaria gracilis* (Stackhouse) Steentoft, Irvine et Farnham from Saldanha Bay, South Africa and Luderitz, Namibia. *Bot. Mar.*, (40), 569-576.

Jaffray, A. E., & Coyne, V. E. (1996). Development of an in situ assay to detect bacterial pathogens of the red alga *Gracilaria gracilis* (Stackhouse) Steentoft, Irvine et Farnham. *Journal of Applied Phycology, 8*, 409-414.

Jordá, L., Coego, A., Conejero, V., & Vera, P. A. (1999). A genomic cluster containing four differentially regulated subtilisin-like processing protease genes is in tomato plants. *J. Biol. Chem., 274*, 2360-2365.

Jordá, L., Conejero, V., & Vera, P. (2000). Characterization of P69E and P69F, two differentially regulated genes encoding new members of the subtilisin-like proteinase family from tomato plants. *Plant physiology, 122*(1), 67-74.

Kader, J-C. (1996). Lipid-transfer proteins in plants. *Annual Review of Plant Physiology and Plant Molecular Biology, 47*, 627-654.

Kader, J-C. (1997). Lipid-transfer proteins: A puzzling family of plant proteins. *Trends in Plant Science, 2*, 66-70.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., et al. (2002). The Human Genome Browser at UCSC. *Genome Res., 12*, 996-1006.

Kikuchi, H., Hirose, S., Toki, S., Kazuhito, A., & Takaiwa, F. (1999). Molecular characterization of a gene for alanine aminotransferase from rice (Oryza sativa). *Plant Mol. Biol., 39*, 149-159.

Kim, K., Kim, I. H., Lee, K. Y., Rhee, S. G., & Stadtman, E. R. (1988). The isolation and purification of a specific protector protein which inhibits enzyme inactivation by a thiol/Fe(III)/O2 mixed-function oxidation system. *Journal of Biological Chemistry, 263*, 4704-4711.

Kingston, R. E., & Narlikar, G. J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes and Development, 13*, 2339-2352.

Klimowski, L., Chandrashekar, R., & Tripp, C. A. (1997). Molecular cloning, expression and enzymatic activity of a thioredoxin peroxidase from Dirofilaria immitis. *Mol Biochem Parasitol., 90*, 297-306.

Knutsen, S. H., Myslabodski, D. E., Larsen, B., & Usov, A. I. (1994). A modified system of nomenclature for red algal galactans. *Botanica Marina, 37*, 163-169.

Kobayashi, E., Sato, S., Hotta, Y., Miyajima, N., Tanaka, A., & Tabata, S. (1994). Characterization of cDNAs induced in meiotic prophase in lily microsporocytes. *DNA Research, 1*, 15-26.

Kobrehel, K., Wong, J. H., Balogh, A., Kiss, F., Yee, B. C., & Buchanan, B. B. (1992). Specific reduction of wheat storage proteins by thioredoxin h. *Plant Physiology, 99*, 919-924.

Krebs, S., Fischaleck, M., & Blum, H. (2009). A simple and loss-free method to remove TRIzol contaminations from minute RNA samples. *Analytical Biochemistry, 387*, 136-138. Elsevier Inc.

Kuhn, E. (2001). From Library Screening to Microarray Technology: Strategies to Determine Gene Expression Profiles and to Identify Differentially Regulated Genes in Plants. *Annals of Botany*, *87*(2), 139-155.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, 227*(August), 680-685.

Lafferty, K. D., Porter, J. W., & Ford, S. E. (2004). Are diseases increasing in the ocean? *Annu. Rev. Ecol. Sysr., 35*, 31-54.

Laloi, C., Apel, K., & Danon, A. (2004a). Reactive oxygen signalling: the latest news. *Current opinion in plant biology, 7*(3), 323-8.

Laloi, C., Mestres-ortega, D., Marco, Y., Meyer, Y., & Reichheld, J.-P. (2004b). The Arabidopsis Cytosolic Thioredoxin h5 Gene Induction by Oxidative Stress and Its W-Box-Mediated Response to Pathogen Elicitor 1. *Plant Physiology*, *134*(March), 1006-1016.

Lam, E., Kato, N., & Lawton, M. (2001). Programmed cell death, mitochondria and the plant hypersensitive response. *Nature, 411*, 848-853.

Lam, E., & Pozo, O. del. (2004). Caspase-like protease involvement in the control of plant cell death. *Plant Mol. Biol., 44*, 417–428.

Lamb, C., & Dixon, R. A. (1997). The Oxidative Burst in Plant Disease Resistance. *Annual review of plant physiology and plant molecular biology, 48*, 251-275.

Laturnus, F., Svensson, T., Wiencke, C., & Oberg, G. (2004). Ultraviolet radiation affects emission of ozone-depleting substances by marine macroalgae: results from a laboratory incubation study. *Environ. Sci. Technol., 38*, 6605-6609.

Lavilla-Pitigo, C. R. (1992). Agar-digesting bacteria associated with “rotten thallus syndrome” of Gracilaria sp. *Aquaculture, 102*, 1-7.

Laxalt, A. M., & Munnik, T. (2002). Phospholipid signalling in plant defence. *Current Opinion in Plant Biology*, *5*(4), 332-338.

Lebi, T. (2006). Identification of differentially expressed genes in the commercially important agarophyte, *Gracilaria gracilis*, following nitrogen deprivation. MSc. Thesis, University of Cape Town.

LeBlanc, C., Colin, C., Cosse, A., Delage, L., Le Barre, S. L., Morin, P., et al. (2006). Iodine transfers in the coastal marine environment: the key role of brown algae and of their vanadium-dependent haloperoxidases. *Biochimie*, *88*, 1773-1785.

Lee, H., Lee, H. K., An, G., & Lee, Y. K. (2007). Analysis of expressed sequence tags from the red alga *Griffithsia okiensis*. *The Journal of Microbiology*, *45*(6), 541-546.

Lee, Y. H., Huang, G. M., Cameron, R. A., Graham, G., Davidson, E. H., Hood, L., et al. (n.d.). EST analysis of gene expression in early cleavage-stage sea urchin embryos. *Development*, *126*, 3857-3867.

Lee, Y., Huang, G. M., Cameron, R. A., Graham, G., Davidson, E. H., Hood, L., et al. (1999). EST analysis of gene expression in early cleavage-stage sea urchin embryos*. *Development*, *126*, 3857-3867.

Leung, Y., & Cavalieri, D. (2003). Fundamentals of cDNA microarray data analysis. *Trends in Genetics*, *19*(11), 649-659.

Levring, T., Hopper, H. A., & Schmid, O. J. (1969). *Marine Algae. A survey of research and utilization* (p. 421). Hamburg: Cram, De Gruyter and Co., Hamburg.

Li, Q., Chen, F., Sun, L., Zhang, Z., Yang, Y., & He, Z. (2006). Expression profiling of rice genes in early defense responses to blast and bacterial blight pathogens using cDNA microarray. *Physiological and Molecular Plant Pathology*, *68*, 51-60.

Liepman, A. H., & Olsen, L. J. (2003). Alanine aminotransferase homologs catalyze the glutamate : glyoxylate aminotransferase reaction in peroxisomes of Aarabidopsis. *Plant Physiology*, *131*(January), 215-227.

Lipshutz, R. J., Fodo, S. P. A., Gingeras, T. R., & Lockhart, D. J. (1999). High density synthetic oligonucleotide arrays. *Nature Genet.*, *21*, 20–24.

Liska, A. J., & Shevchenko, A. (2003). Expanding the organismal scope of proteomics: Cross species protein identification by mass spectrometry and its implications. *Proteomics*, *3*, 19-28.

Littler, M. M., & Littler, D. S. 1. (1998). An undescribed fungal pathogen of reef-forming crustose corraline algae discovered in American Samoa. *Coral Reefs*, *17*(2), 144.

Liu, W., & Saint, D. A. (2002). A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. *Anal. Biochem., 302*, 52-59.

Lluisima, A. O., & Ragan, M. A. (1997). Expressed sequence tags (ESTs) from the marine red alga *Gracilaria gracilis*. *Journal of Applied Phycology, 9*(3), 287-293.

Loomis, W. D. (1974). Overcoming problems of phenolics and quinones in the isolation of plant enzymes and organelles. *Methods in Enzymology, 31*, 528-545.

Luan, S. (1998). Protein phosphatases and signaling cascades in higher plants. *Trends in Plant Science, 3*(7), 271-275.

Luo, L., Salunga, R. C., Guo, H., Bittner, A., Joy, K. C., Galindo, J. E., et al. (1999). Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat. Med., 5*, 117-122.

Lusser, A., & Kadonaga, J. T. (2003). Chromatin remodeling by ATP-dependent molecular machines. *BioEssays, 25*, 1192-200.

Mackey, D., & McFall, A. J. (2006). MAMPs and MIMPs: proposed classifications for inducers of innate immunity. *Mol. Microbiol., 61*, 1365-1371.

Mackie, W., & Preston, R.D. (1974). Cell wall and intercellular region polysaccharides. In W. D. P. Stewart (Ed.), *Algal Physiology and Biochemistry* (pp. 40-85). Oxford: Blackwell Scientific Publishers.

Mackintosh, C., Lyon, G. D., & Mackintosh, R. W. (1994). Protein phosphatase inhibitors activate anti-fungal defence responses of soybean cotyledons and cell cultures. *Plant J., 5*, 137-147.

Maeda, K., Finnie, C., Ostergaard, O., & Svensson, B. (2003). Identification, cloning and characterization of two thioredoxin h isoforms, HvTrxh1 and HvTrxh2, from the barley seed proteome. *European Journal of Biochemistry, 270*(12), 2633-2643.

Mahalingam, R., Gomez-Buitrago, A., Eckardt, N., Shah, N., Guevara-garcia, A., Day, P., et al. (2003). Characterizing the stress / defense transcriptome of Arabidopsis. *Genome Biology, 4*(R20), 1-14.

Maier, T., Güell, M., & Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Letters, 583*, 3966-3973. Federation of European Biochemical Societies.

Maldonado, A. M., Doerner, P., Dixonk, R. A., Lamb, C. J., & Cameron, R. K. (2002). A putative lipid transfer protein involved in systemic resistance signalling in Arabidopsis. *Nature, 419*, 399-403.

Mamedov, T. G., Suzuki, K., Miura, K., Kucho, K.-I., & Fukuzawa, H. (2001). Characteristics and sequence of Phosphoglycolate Phosphatase from a eukaryotic green alga *Chlamydomonas reinhardtii*. *Biological Chemistry, 276*(49), 1-15.

Mantri, N. L., Ford, R., Coram, T. E., & Pang, E. C. K. (2010). Evidence of unique and shared responses to major biotic and abiotic stresses in chickpea. *Environmental and Experimental Botany*, *69*(3), 286-292. Elsevier B.V.

Marchler-Bauer, A., Anderson, J. B., Chitsaz, F., & Et, A. (2009). CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Research, 37(D)*, 205-210.

Maredza, A. T. (2007). *Isolation of the aldose reductase gene (XvAld1) from the resurrection plant Xerophyta viscosa, and characterisation of the gene product and transgenic plants expressing the gene.*

Marek, L. F., & Spalding, M. H. (1991). Changes in photorespiratory enzyme activity in response to limiting CO2 in *Chlamydomonas reinhardtii*. *Plant Physiol. (Bethesda), 97*, 420-425.

Marfella, C. G. A., & Imbalzano, A. N. (2007). The Chd family of chromatin remodelers. *Mutation Research, 618*, 30-40.

Markovic, O., & Janecek, S. (2001). Pectin degrading glycoside hydrolases of family 28: sequence-structural features, specificities and evolution. *Protein Eng.*, *14*, 615-631.

Marquet, A. (2001). Enzymology of carbon-sulfur bond formation. *Current Opinion in Chemical Biology, 5*, 541–549.

Marrion, O., Fleurence, J., Schwertz, A., Guéant, J.-L., Mamelouk, L., Ksouri, J., et al. (2005). Evaluation of protein in vitro digestibility of Palmaria palmata and *Gracilaria verrucosa*. *Journal of Applied Phycology, 17*, 99-102.

Marty, I., Brugidou, C., Chartier, Y., & Meyer, Y. (1993). Growth-related gene expression in Nicotiana tabacum mesophyll protoplasts. *Plant J.*, *4*, 265-278.

Mathieu, Y., Kurdijan, A., Xia, H., Guera, J., Koller, A., Spiro, M. D., et al. (1991). Membrane responses induced by oligogalacturonides in suspension-cultured tobacco cells. *Plant J.*, *1*, 333-343.

Matsuzaki, M., Misumi, O., Shin, I. T., Maruyama, S., Takahara, M., Miyagishima, S. Y., et al. (2004). Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature, 428*, 653-657.

McCandless, E. L. (1981). Polysaccharides of the seaweeds. In C. S. Lobban & M. J. Wynne (Eds.), *The biology of the seaweeds, Bot. Monogr. 17* (pp. 559-588). Oxford: Blackwell Scientific Publishers.

Mccarthy, D. J., & Smyth, G. K. (2009). Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics (Oxford, England)*, *25*(6), 765-71.

McDonald, J. H. (2008). *Handbook of Biological Statistics.* (p. 293).

McFadden, G. I., & Dooren, G. G. van. (2004). Evolution: red algal genome affirms a common origin of all plastids. *Current Biology*, *14*(13), 514-516.

McLachlan, J. (1985). Macroalgae (seaweeds): Industrial resources and their utilization. *Plant and Soil*, *89*, 137-157.

Meijer, E. A., Vries, S. C. de, Sterk, P., Gadella, D. W., Wirtz, K. W. A., & Hendriks, T. (1993). Characterization of the non-specific lipid transfer protein EP2 from carrot (Daucus carota L.). *Molecular and Cellular Biochemistry.*, *123*, 159-166.

Melis, A. (1999). Photosystem-II damage and repair cycle in chloroplasts: what modulates the rate of photodamage in vivo? *Trends in Plant Science*, *4*(4), 130-135.

Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., et al. (2007). The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science*, *318*(5848), 245-251.

Meyer, Y., Vignols, F., & Reichheld, J. P. (2002). Classification of plant thioredoxins by sequence similarity and intron position. *Methods Enzymol.*, *347*, 394-402.

Mihara, H., & Esaki, N. (2002). Bacterial cysteine desulfurases: their function and mechanisms. *Applied microbiology and biotechnology*, *60*, 12-23.

Miyoshi, E., Moriwaki, K., & Nakagawa, T. (2008). Biological function of fucosylation in cancer biology. *The Japanese Biochemical Society, Journal of*(143), 725-729.

Mizoguchi, T., Ichimura, K., & Shinozaki, K. (1997). Environmental stress response in plants: the role of mitogen-activated protein kinases. *Trends in Biotechnology*, *15*, 15-19.

Molina, A., Segura, A., & Garcia-Olmedo, F. (1993). Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant pathogens. *FEBS Lett.*, *316*, 119-122.

Montane, M. H., Petzold, B., & Kloppstech, K. (1999). Formation of early-light-inducible-protein complexes and status of xanthophyll levels under high light and cold stress in barley. *Planta*, *208*(519-527).

Montane, M. H., Dreyer, S., Triantaphylides, C., & Kloppstech, K. (1997). Early light-inducible proteins during long-term acclimation of barley to photooxidative stress caused by light and cold: high level of accumulation by posttranscriptional regulation. *Planta*, *202*, 293-302.

Morey, J. S., Ryan, J. C., & Van Dolah, F. M. (2006). Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biological procedures online*, *8*(1), 175-93.

Moriwaki, K., Noda, K., Asahi, M., Yoshihara, H., Taniguchi, N., Hayashi, N., et al. (2007). A high expression of GDP-fucose transporter in hepatocellular carcinoma is a key factor for increases in fucosylation. *Glycobiology*, *17*(12), 1311-1320.

Mouaheb, N., Thomas, D., Verdoucq, L., Monfort, P., & Meyer, Y. (1998). In vivo functional dicrimination between plant thioredoxins by heterologous expression in the yeast Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA*, *95*, 3312-3317.

Mowla, S. B. (2005). Molecular Characterization of XvPer1, a novel antioxidant enzyme from the resurrection plant Xerophyta viscosa, and AC3, a LEA-like protein from *Arabidopsis thaliana*. PhD Thesis, University of Cape Town.

Mowla, S. B., Thomson, J. A., Farrant, J. M., & Mundree, S. G. (2002). A novel stress-inducible antioxidant enzyme identified from the resurrection plant Xerophyta viscosa Baker. *Planta*, *215*, 716-726.

Mtolera, M. S., Collen, J., Pedersem, M., & Semesi, A. K. (1995). Destructive hydrogen peroxide production in Eucheuma denticulatum (Rhodophyta) during stress caused by elevated pH, high light intensities and competition with other species. *Eur. J. Phycol.*, *30*, 289-297.

Mtolera, M., Collen, J., Pedersen, M., Ekdahl, A., Abrahamsson, K., & Semisi, A. (1996). Stress induced production of volatile halogenated organic compounds in Eucheuma denticulatum (Rhodophyta) caused by elevated pH and high light intensity. *European Journal of Phycology*, *31*, 89-95.

Muench, D. G., Christopher, M. E., & Good, A. G. (1998). Cloning and expression of a hypoxic and nitrogen inducible maize alanine aminotransferase gene. *Physiol. Plant.*, *103*, 503-512.

Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., et al. (2007). New developments in the InterPro database. *Nucleic Acids Res.*, *35*(Database Issue), D224-228.

Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., et al. (2005). InterPro, progress and status in 2005. *Nucleic Acids Res.*, *33*(Database Issue), D201 - 5.

Murano, E. (1995a). Chemical structure and quality of agars from Gracilaria. *Journal of Applied Phycology*, *7*, 245–254.

Murano, E. (1995b). Chemical structure and quality of agars from Gracilaria. *Journal of Applied Phycology*, *7*, 245-254.

Mutch, D. M., Berger, A., Mansourian, R., Rytz, A., & Roberts, M. (2002). The limit fold change model : A practical approach for selecting differentially expressed genes from microarray data. *BMC Bioinformatics*, *3*(17), 1-11.

Muzi-Falconi, M., Brown, G. W., & Kelly, T. J. (1996). Controlling initiation during the cell cycle. DNA replication. *Curr. Biol.*, *6*, 229-233.

Mysore, K. S., & Ryu, C. M. (2004). Nonhost resistance: how much do we know? *Trends Plant Sci.*, *9*, 97-104.

Müntz, K., Belozersky, M. A., Dunaevsky, Y. E., Schlereth, A., & Tiedemann, J. (2001). Stored proteases and the initiation of storage protein mobilization in seeds during germination and seedling growth. *Journal of Experimental Botany*, *50*, 1741-1752.

Ndimba, B. K., Chivasa, S., Simon, W. J., & Slabas, A. R. (2005). Identification of Arabidopsis salt and osmotic stress responsive proteins using two-dimensional difference gel electrophoresis and mass spectrometry. *Proteomics*, *5*(16), 4185-96.

Neill, S. J., Desikan, R., Clarke, A., Hurst, R. D., & Hancock, J. T. (2002). Hydrogen peroxide and nitric oxide as signalling molecules in plants. *Journal of Experimental Botany*, *53*(372), 1237-1247.

Nie, L., Wu, G., & Zhang, W. (2006). Correlation between mRNA and protein abundance in Desulfovibrio vulgaris: A multiple regression to identify sources of variations. *Biochemical and Biophysical Research Communications*, *339*, 603-610.

Nightingale, P. D., Malin, G., & Liss, P. S. (1995). Production of chloroform and other low molecular weight halocarbons by some species of macroalgae. *Limnol Oceanogr.*, *40*, 680-689.

Nikaido, I., Asamizu, E., Nakajima, M., Nakamura, Y., Saga, N., & Tabata, S. (2000). Generation of 10,154 expressed sequence tags from a leafy gametophyte of a marine red alga, *Porphyra yezoensis*. *DNA Research*, *7*, 223-227.

Nishizuka, Y. (1986). Studies and perspectives of protein kinase c. *Science*, *233*, 305-312.

Niyogi, K. K. (1999). Photoprotection revisited: genetic and molecular approaches. *Annual Review of Plant Physiology and Plant Molecular Biology*, *50*, 333-359.

Nolan, T., Hands, R. E., & Bustin, S. A. (2006). Quantification of mRNA using real-time RT-PCR. *Nature protocols*, *1*(3), 1559-82.

Norman, E. G., & Colman, B. (1991). Purification and Characterization of Phosphoglycolate Phosphatase from the Cyanobacterium Coccochloris peniocystis. *Plant Physiol. (Bethesda)*, *95*(693-698).

Nozaki, H., Takano, H., Misumi, O., Terasawa, K., Matsuzaki, M., Maruyama, S., et al. (2007). A 100% complete sequence reveals unusually simple genomic features in the hot-spring red alga *Cyanidioschyzon merolae*. *BMC Biology*, *5*(28), 1-8.

Nurnberger, T., & Lipka, V. (2005a). Non-host resistance in plants : new insights into an old phenomenon. *Molecular Plant Pathology*, *6*(3), 335-345.

Nurnberger, T., & Lipka, V. (2005b). Non-host resistance in plants: new insights into an old phenomenon. *Mol. Plant Pathol.*, *6*, 335–345.

Nürnberger, T., & Brunner, F. (2002). Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns. *Current Opinion in Plant Biology*, *5*(4), 318-324.

Ohta, N., Matsuzaki, M., Misumi, O., Miyagishima, S., Nozaki, H., Tanaka, K., et al. (2003). Complete sequence and analysis of the plsatid genome of the unicellular red alga *Cyanidioschyzon merolae*. *DNA Research*, *10*(2), 67-77.

Ohyama, C., Smith, P. L., Angata, K., Fukuda, M. N., Lowe, J. B., & Fukuda, M. (1998). Molecular cloning and expression of GDP-D-mannose-4,6-dehydratase, a key enzyme for fucose metabolism defective in Lec13 cells. *Journal of Biological Chemistry*, *273*, 14582-14587.

Oliveira, E. C., Alveal, K., & Anderson, R. J. (2000). Mariculture of the Agar-Producing Gracilarioid Red Algae. *Reviews in Fisheries Science*, *8*(4), 345-377.

Osterhaus, A. D. M. E., Swart, R. L. de, Vos, H. W., Ross, P. S., Kenter, M. J. H., & Barrettc, T. (1995). Morbillivirus infections of aquatic mammals: newly identified members of the genus. *Veterinary Microbiology*, *44*(2-4), 219-227.

Otter, T., Penther, J. M., & Mohr, H. (1992). Control of the appearance of alanine aminotransferase in the scots pine (Pinus sylvestris L.) seedling. *Planta*, *188*, 376-383.

O'Brien, I. E. W., Reutelingsperger, C. P. M., & Holdaway, K. M. (1997). Annexin-V and TUNEL use in monitoring the progression of apoptosis in plants. *Cytometry*, *29*, 28-33.

Pal, D., & Eisenberg, D. (2005). Inference of protein function from protein structure. *Structure*, *13*, 121-130.

Parsons, G. G., & Spencer, C. A. (1997). Mitotic repression of RNA polymerase II transcription is accompanied by release of transcription elongation complexes. *Mol. Cell Biol.*, *17*, 5791–5802.

Patrzykat, A., & Douglas, S. E. (2003). Gone gene fishing: How to catch novel marine antimicrobials21. *Trends in Biotechnology*, *21*, 362-369.

Paul, J. H., & Cooksey, K. E. (1981). Regulation of Asparaginase , Glutamine Synthetase, and Glutamate Dehydrogenase in response to medium nitrogen concentrations in a euryhaline Chlamydomonas species. *Plant Physiol.*, *68*, 1364-1368.

Paul, N. A., Nys, R. de, & Steinberg, P. D. (2006). Chemical defence against bacteria in the red alga Asparagopsis armata (Bonnemaisoniaceae). *Mar Ecol Prog Ser*, *306*, 87-101.

Pearce, G., Yamaguchi, Y., Barona, G., & Ryan, C. A. (2010). A subtilisin-like protein from soybean contains an embedded, cryptic signal that activates defense-related genes. *Proc. Natl. Acad. Sci. USA*, *107*(33), 14921-14925.

Pedersen, M., Collen, J., Abrahamsson, K., Mtolera, M., Semesi, A., & Garcia Reina, G. (1996). The ice-ice disease and oxidative stress of marine algae. In M. Bjork, A. Semesi, M. Pedersen, & B. Bergman (Eds.), *Proceedings of the 1995 symposium on the biology of microalgae, macroalgae and seagrasses in the western Indian ocean.* (pp. 11-24). Ord & Vetande AB, Uppsala, Sweden.

Pennell, R. I., & Lamb, C. (1997). Programmed Cell Death in Plants. *The Plant cell*, *9*(7), 1157-1168.

Penther, J. M. (1991). Analysis of alanine and aspartate aminotransferase isoforms in mustard (Sinapis alba L.) cotyledons. *J Chromatogr.*, *587*, 101-108.

Peterson, C. H., & Lubchenco, J. (1997). Marine ecosystem services. In G. C. Daily (Ed.), *Nature's Services: Societal dependence on natural ecosystems* (pp. 174-194). Island Press, Washington D. C.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research*, *29*(9), 16-21.

Phelan, M. L., Schnitzler, G. R., & Kingston, R. E. (2000). Octamer transfer and creation of stably remodeled nucleosomes by human SWI-SNF and its isolated ATPases. *Mol. Cell. Biol.*, *20*, 6380-6389.

Phillips, J., & Eberwine, J. H. (1996). Antisense RNA amplification: a linear amplification method for analyzing the mRNA population from single living cells. *Methods*, *10*, 283-288.

Phillips, M. (2009). Maricutlure Overview, 3294-3917.

Pitzschke, A., Forzani, C., & Hirt, H. (2006). Reactive oxygen species signalling in plants. *Antioxid. Redox Signaling*, *8*, 1757–1764.

Plastino, E. M. (1985). *As especies de Gracilaria (Rhodophyta, Gigartinales) da praia Dura, Ubatuba, SP - aspectos biologiocos e fenologia.*

Polson, A., Potgieter, G. M., Largier, J. F., Mears, G. E., & Joubert, F. J. (1964). The fractionation of protein mixtures by linear polymers of high molecular weight. *Biochim. Biophys. Acta*, *82*, 463-475.

Ponhert, G. (2004). Chemical defense strategies of marine organisms. *Topics in Current Chemistry*, *239*, 179-219.

Potin, P. (2008). Oxidative burst and related responses in biotic interaciton of algae. In C. D. Amsler (Ed.), *Algal Chemical Ecology* (pp. 245-271). Berlin Heidenlberg: Springer-Verlang.

Potin, P., Bouarab, K., Salaun, J.-P., Pohnert, G., & Kloareg, B. (2002). Biotic interactions of marine algae. *Current Opinion in Plant Biology*, *5*, 308-317.

Provasoli, L. (1968). Media and prospects for the cultivation of marine algae.,.,. In A. Watanabe & A. Hattori (Eds.), *Cultures and collections of algae.* (pp. 63-75). Hakone: Proc. U.S.-Japan Conf.

Ptashne, M., & Gann, A. (2002). *Genes & Signals.* New York: Cold Spring Harbor Laboratory Press.

Qin, S., Jiang, P., & Tseng, C. (2005). Transforming kelp into a marine bioreactor. *Trends in Biotechnology*, *23*(5), 264-268.

Quackenbush, J. (2002). Microarray data normalization and transformation, *32*, 496-501.

Quackenbush, J. (2001). Computational analysis of microarray data. *Nat. Rev. Gent.*, *2*, 418-427.

Rajeevan, M. S., Ranamukhaarachchi, D. G., Vernon, S. D., & Unger, E. R. (2001a). Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies. *Methods*, *25*, 443-451.

Rajeevan, M. S., Vernon, S. D., Taysavang, N., & Unger, E. R. (2001b). Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *Journal of Molecular Diagnostics*, *3*(1), 26-31.

Ralph, S. G., Yueh, H., Friedmann, M., Aeschliman, D., Zeznik, J. A., Nelson, C. C., et al. (2006). Conifer defence against insects : microarray gene expression profiling of Sitka spruce (Picea sitchensis) induced by mechanical wounding or feeding by spruce budworms (Choristoneura occidentalis) or white pine weevils (Pissodes strobi) reveals large. *Plant, Cell and Environment*, *29*, 1545-1570.

Raskin, I. (1992). Role of salicylic acid in plants. *Annual Review of Plant Physiology and Plant Molecular BiologyMolecular Biology*, *43*, 439-463.

Rawling, N. D., & Barret, A. J. (2004). Families of serine peptidases. *Methods in Enzymology*, *244*, 19-61.

Rebello, J., Ohno, M., Ukeda, H., & Sawamura, M. (1997). Agar quality of commercial agarophytes from different geographical origins: 1. Physical and rheological properties . *Journal of Applied Phycology*, *8*, 517-521.

Rehfield, D. W., & Tolbert, N. E. (1972). Aminotransferases in peroxisomes from spinach leaves. *Journal of Biological Chemistry*, *247*, 4803-4811.

Rep, M., Albertyn, J., Thevelein, J. M., & Prior, B. A. (1999). Different signalling pathways contribute to the control of GPD1 gene expression by osmotic stress in Saccharomyces cerevisiae. *Microbiology*, *145*, 715-727.

Reymond, P. (2001). DNA microarrays and plant defence. *Plant Physiol. Biochem.*, *39*, 313-321.

Richmond, C. S., Glasner, J. D., R., M., Jin, H., & Blattner, F. R. (1999). Genome-wide expression profiling in Escherichia coli K-12. *Nucleic Acids Res.*, *27*, 3821–3835.

Richmond, T., & Somerville, S. (2000). Chasing the dream : plant EST microarrays. *Current Opinion in Plant Biology*, *3*, 108-116.

Rivera-Madrid, R., Mestres, D., Marinho, P., Jacquot, J. P., Decottignies, P., Miginiac-Maslow, M., et al. (1995). Evidence for five divergent thioredoxin h sequences in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA*, *92*, 5620–5624.

Roberts, J. A., Elliott, K. A., & Gonzalez-Carranza, Z. H. (2002). Abscission, dehiscence, and other cell separation processes. *Annual Review of Plant Biology*, *53*, 131-158.

Rockett, J. C., & Hellmann, G. M. (2004). Confirming microarray data - is it really necessary? *Genomics*, *83*, 541 - 549.

Rodriguez, M., Matulewicz, M., Noseda, M., Ducatti, D., & Leonardi, P. (2009). Agar from *Gracilaria gracilis* (Gracilariales, Rhodophyta) of the Patagonic coast of Argentina – Content, structure and physical properties. *Bioresource Technology*, *100*, 1435-1441. Elsevier Ltd.

Roeder, V., Collén, J., Rousvoal, S., Corre, E., Leblanc, C., & Boyen, C. (2005). Identification of stress genes from *Laminaria digitata* (Phaeophyceae) protoplast cultures by expressed sequence tag analysis. *Journal of Phycology*, *41*, 1227-1235.

Rohringer, R., Stahmann, M. A., & Walker, J. C. (1958). Effect of Fusarium oxysporum f. sp. lycopersici and its metabolites on leaf constituents of susceptible and resistant tomatoes. *Journal of Agricultural and Food Chemistry*, *6*, 838-843.

Rose, J. K. C., & Bennett, A. B. (1999). Cooperative disassembly of the cellulose-xyloglucan network of plant cell walls: parallels between cell expansion and fruit ripening. *Trends in Plant Science*, *4*, 176-183.

Rotari, V. I., He, R., & Gallois, P. (2005). Death by proteases in plants: whodunit. *Physiologia Plantarum*, *123*, 376-385.

Rothman, M. D., Anderson, R. J., Boothroyd, C. J. T., Kemp, F. A., & Bolton, J. J. (2009). The gracilarioids in South Africa: a long-term monitoring of a declining resource. *Journal of Applied Phycology*, *21*, 47-53.

Rotmann, K. W. G. (1990). Saldanha Bay, South Africa: recovery of *Gracilaria verrucosa* (Gracilariales, Rhodophyta). *Hydrobiology*, *204/205*, 325-330.

Rouhier, N., Gelhaye, E., Gualberto, J. M., Jordy, M.-N., Fay, E. D., Hirasawa, M., et al. (2004). Poplar Peroxiredoxin Q . A thioredoxin-linked chloroplast antioxidant functional in pathogen defense. *Plant Physiology*, *134*(March), 1027-1038.

Rouhier, N., & Jacquot, J-P. (2002). Plant peroxiredoxins: alternative hydroperoxide scavenging enzymes. *Photosynthesis research*, *74*, 259-68.

Rybicki, E. P. (1979). *The serology of Bromoviruses.* Cape Town.

Ryther, J. H., DeBoer, J. A., & Lapointe, B. E. (1979). Cultivation of seaweeds for hydrocolloids, waste treatment and biomass for energy conversion. *Proc. int. Seaweed Symposium*, *9*, 1-16.

Salcedo, G., Sánchez-monge, R., Barber, D., & Díaz-Perales, A. (2007). Plant non-specific lipid transfer proteins : An interface between plant defence and human allergy. *Biochimica et Biophysica Acta*, *1771*, 781 - 791.

Salem, K., & Waasbergen, L. G. van. (2004). Light Control of hliA transcription and transcript stability in the Cyanobacterium Synechococcus elongates strain PCC 7942. *Journal of Bacteriology*, *186*, 1729-1736.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning. A Laboratory Manual.* (Second.). New York: Cold Spring Harbour Laboratory Press 2.

Santelices, B., & Doty, M. S. (1989). A review of Gracilaria farming. *Aquaculture*, *78*, 95-133.

Sarre, T. F. (1989). The phosphorylation of eukaryotic initiation factor 2: a principle of translational control in mammalian cells. *Biosystems*, *22*, 311-325.

Scala, S., Carels, N., Falciatore, A., Chiusano, M. L., & Bowler, C. (2002). Genome properties of the diatom *Phaeodactylum tricornutum*. *Plant Physiol.*, *129*, 993-1002.

Schaller, A. (2004). A cut above the rest: the regulatory function of plant proteases. *Planta*, *220*, 183–197.

Schena, M., Shalon, D., Davis, R. W., & Brownt, P. (1995). Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science*, *270*(October), 467-470.

Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P. O., & Davis, R. W. (1996). Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. USA*, *93*, 10614–10619.

Schenk, H., Klein, M., Erdbrugger, W., Droge, W., & Schulze-Osthoff, K. (1994). Distinct effects of thioredoxin and antioxidants on the activation of transcription factors NF-kappa-B and AP-1. *Proc. Natl. Acad. Sci. USA*, *91*, Schenk H., Klein M., Erdbrugger W., Droge W. and S.

Schenk, P. M., Kazan, K., Wilson, I., Anderson, J. P., Richmond, T., Somerville, S. C., et al. (2000). Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. *Proc. Natl. Acad. Sci. USA*, *97*(21), 11655-11660.

Schindler, H., Wiese, A., Auer, J., & Burtscher, H. (2005). cRNA target preparation for microarrays : Comparison of gene expression profles generated with diVerent amplifcation procedures. *Analytical Biochemistry*, *344*, 92-101.

Schnell, J. D., & Hicke, L. (2003). Non-traditional functions of ubiquitin and ubiquitin-binding proteins. *Journal of Biological Chemistry, 278*, 35857–35860.

Schroeder, D. C., Jaffer, M. A., & Coyne, V. E. (2003). Investigation of the role of a B(1-4) agarase produced by Pseudoalteromonas gracilis B9 in eliciting disease symptoms in the red alga *Gracilaria gracilis*. *Microbiology*, *149*, 2919-2929.

Schuchhardt, J., Beule, D., Malik, A., Wolski, E., Eickhoff, H., Lehrach, H., et al. (2000). Normalization strategies for cDNA microarrays. *Nucleic Acids Res.*, *28*, e47.

Schulze, A., & Downward, J. (2001). Navigating gene expression using microarrays--a technology review. *Nature Cell Biology*, *3*(8), E190-5.

Schweighofer, A., Hirt, H., & Meskiene, I. (2004). Plant PP2C phosphatases : emerging functions in stress signaling. *Trends in Plant Science*, *9*(5).

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., et al. (2002). Monitoring the expression profilrd of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *The Plant Journal*, *31*(3), 279-292.

Serrato, A. J., Crespo, J. L., Florencio, F. J., & Cejudo, F. J. (2001). Characterization of two thioredoxins h with predominant localization in the nucleus of aleurone and scutellum cells of germinating wheat seeds. *Plant Mol. Biol.*, *46*, 361–371.

Shaffer, J. P. (1995). Multiple hypothesis testing. *Annu. Rev. Psychol.*, *46*, 561-584.

Shalon, D., Smith, S.J. And Brown, P. O. (1996). A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.*, *6*, 639-645.

Sharma, A. D., Gill, P. K., & Singh, P. (2003). RNA isolation from plant tissues rich in polysaccharides. *Analytical Biochemistry*, *314*, 319-321.

Shenolikar, S. (1994). Protein serine/threonine phosphatases: new avenues for cell regulation. *Annu. Rev. Cell Biol.*, *10*, 55-86.

Shevchenko, A., Sunyaev, S., Loboda, A., Bork, P., Ens, W., & Standing, K. G. (2001). Charting the proteomes of organisms with unsequenced geneomes by MALDI-quadrupole time of flight mass spectrometry and BLAST homology searching. *Anal Chem*, *73*, 1917-1926.

Siebert, P. D. (1997). RT-PCR methods and applications. In U. Reischi & N. J. Otowa (Eds.), *Methods in Molecular Medicine, Vol 13: Molecular Diagnosis of Infectious Diseases.* (pp. 29-53). Humana Press, Inc.

Smalle, J., & Vierstra, R. D. (2004). The ubiquitin 26S proteasome proteolytic pathway. *Annual Review of Plant Biology*, *55*, 555–590.

Smith, R. D., & Walker, J. C. (1996). Plant protein phosphatases. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, *47*, 101-125.

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, *3*(1, Article 3), 1-27.

Smyth, G., Thorne, N., & Wettenhall, J. (2005). *limma : Linear Models for Microarray Data User ' s Guide*. www.bioconductor.org.

Somerville, S., Nishimura, M., Hughes, D., Wilson, I., & Vogel, J. (1998). Alternate methods of gene discovery - the candidate EST approach and DNA micoarrays. In F. Lo Schiavo, R. L. Last, G. Morelli, & N. Raikhel (Eds.), *Cellular Integration of Signalling Pathways in Plant Development*, *NATOASI Series* (p. 297–309). Berlin: Springer-Verlag.

Son, D., Jo, J., & Sugiyama, T. (1991). Purification and characterization of alanine aminotransferase from Panicum miliaceum leaves. *Arch Biochem Biophys.*, *289*, 262-266.

Son, D., & Sugiyama, T. (1992). Molecular cloning of an alanine aminotransferase from NAD-malic enzyme type C4 plant Panicum miliaceum. *Plant Mol. Biol.*, *20*, 705-713.

Song, J. T., Lu, H., & Greenberg, J. T. (2004). Divergent Roles in *Arabidopsis thaliana* Development and Defense of Two Homologous Genes , ABERRANT GROWTH AND DEATH2 and AGD2-LIKE DEFENSE RESPONSE PROTEIN1 , Encoding Novel Aminotransferases. *The Plant Cell*, *16*, 353-366.

Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., et al. (1998). Comprehensive identification of cell cycle-regulated genes of the

yeast Saccharomyces cerevisiae by microarray hybridization. *Mol. Biol. Cell.*, *9*, 3273–3297.

Stanley, M. S., Perry, R. M., & Callow, J. A. (2005). Analysis of expressed sequence tags from the green algal *Ulva linza* (Chlorophyta). *Journal of Phycology*, *41*, 1219-1226.

Sterk, P., Booij, H., Scellekens, G. A., Kammen, A. van, & Vries, S. G. de. (1991). Cell-specific expression of the carrot EP2 lipid transfer protein gene. *Plant Cell*, *3*, 907-921.

Stone, J. M., & Walker, J. C. (1995). Protein kinase families and signal transduction. *Plant Physiology*, *108*, 451-457.

Storz, G., Jacobson, F. S., Tartaglia, L. A., Morgan, R. W., Silveira, L. A., & Ames, B. N. (1989). An alkyl hydroperoxide reductase induced by oxidative stress in Salmonella typhimurium and Escherichia coli: genetic characterization and cloning of ahp. *Journal of Bacteriology*, *171*, 2049–2055.

Subramanian, B., Bansal, V. K., & Kav, N. N. (2005). Proteome-level investigation of Brassica carinata-derived resistance to Leptosphaeria maculans. *J. Agric. Food Chem.*, *53*, 313-324.

Sullivan, F. X., Kumar, R., Kriz, R., Stahl, M., Xu, G. Y., Rouse, J., et al. (1998). Molecular cloning of human GDPmannose 4,6-dehydratase and reconstitution of GDP-fucose biosynthesis in vitro. *Journal of Biological Chemistry*, *273*, 8193-8202.

Suto, S., Saito, Y., Akiyama, K., & Umebayashi, O. (1972). Textbook of diseases and their symptoms in *Porphyra* (p. 37). Japan: Tokai Regional Fish. Res. Lab.

Suzuki, K. (1995). Phosphoglycolate phosphatase-deficient mutants of *Chlamydomonas reinhardtii* capable of growth under air. *Plant Cell Physiol.*, *36*, 95-100.

Suzuki, K. E. A. (1999). Slow and prolonged activation of the p47 protein kinase during hypersensitive cell death in a culture of tobacco cells. *Plant Physiology*, *119*, 1465–1472.

Suzuki, K., Mamedov, T. G., & Ikawa, T. (1999). A mutant of *Chlamydomonas reinhardtii* with reduced rate of photorespiration. *Plant and Cell Physiology.*, *40*, 792-799.

Suzuki, K., Marek, L. F., & Spalding, M. H. (1990). A photorespiratory mutant of *Chlamydomonas reinhardtii*. *Plant Physiology*, *93*, 231-237.

Suzuki, K., & Shinshi, H. (1995). Transient activation and tyrosine phosphorylation of a protein kinase in tobacco cells treated with a fungal elicitor. *Plant Cell*, *7*(639-647).

Taki, N., Sasaki-Sekimoto, Y., & Obayashi, T. E. A. (2005). 12-Oxophytodienoic acid triggers expression of a distinct set of genes and plays a role in wound-induced gene expression in Arabidopsis. *Plant Physiology, 139*, 1268-1283.

Taler, D., Galperin, M., Benjamin, I., Cohen, Y., & Kenigsbuch, D. (2004). Plant eR Genes That Encode Photorespiratory Enzymes Confer Resistance against Disease The Plant Cell. *The Plant Cell, 16*, 172-184.

Tanaka, T. S., Jaradat, S. A., Lim, M. K., Kargul, G. J., Wang, X., Grahovac, M. J., et al. (2000). Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc. Natl. Acad. Sci. USA, 97*, 9127–9132.

Tchang, F., This, P., Stiefel, V., Arondel, V., March, M.-D., Pages, M., et al. (1988). Phospholipid transfer protein: full-length cDNA and amino acid sequence in maize. Amino acid sequence homologies between plant phospholipid transfer proteins. *Journal of Biological Chemistry*, *263*(32), 16849-16855.

Teo, S., Ho, C.-L., Teoh, S., Lee, W.-W., Tee, J.-M., Rahim, R. A., et al. (2007). Analyses of expressed sequence tags from an agarophyte, *Gracilaria changii* (Gracilariales, Rhodophyta). *European Journal of Phycology*, *41*(1), 41-46.

Teo, S-S., Ho, C-L., Teoh, S. & Rahim, R. A. (2009). Transcriptomic analysis of *Gracilaria changii* (Rhodophyta) in response to hyper- and hypoosmotic stresses. *Journal of Phycology, 45,* 1093-1099.

Terras, F. R. G., Goderis, I. J., Leuvren, F. van, Vanderleyden, J., Cammue, B. P. A., & Broekaert, W. F. (n.d.). In vitro antifungal activity of a radish (Raphanus sativus L.) seed protein homologous to nonspecific lipid transfer proteins. *Plant Physiology*, *100*, 1055-1058.

Thomas, J. G., Olson, J. M., Tapscott, S. J., & Zhao, L. P. (2001). An efficient and robust statistical modeling approach to discover differentially expressed genes using genomic expression profiles. *Genome Res.*, *11*, 1227-1236.

Tiwari, B. S., Belenghi, B., & Levine, A. (2002). Oxidative stress increased respiration and generation of reactive oxygen species, resulting in ATP depletion, opening of mitochondrial permeability transition, and programmed cell death. *Plant Physiology, 128*, 1271-1281.

Tonetti, M., Sturla, L., Bisso, A., Benatti, U., & De Flora, A. (1996). Synthesis of GDP-L-fucose by the human FX protein. *Journal of Biological Chemistry, 271*, 27274-27279.

Torki, M., Mandaron, P., Mache, R., & Falconet, D. (2000). Characterization of a ubiquitous expressed gene family encoding polygalacturonase in *Arabidopsis thaliana*. *Gene, 242*, 427-436.

Tornero, P., Conejero, V., & Vera, P. (1996). Primary structure and expression of a pathogen-induced protease (PR-P69) in tomato plants: similarity of functional

domains to subtilisin-like endoproteases. *Proc. Natl. Acad. Sci. USA*, *93*, 6332–6337.

Tornero, P., Conejero, V., & Vera, P. (1997). Identification of a new pathogen induced member of the subtilisin-like processing protease family from plants. *J. Biol. Chem.*, *272*, 14412–14419.

Tseng, G. C., Oh, M. K., Rohlin, L., Liao, J. C., & Wong, W. H. (2001). Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res.*, *29*(12), 2549-57.

Tsuboi, S., Osafune, T., Nishimura, M., & Yamada, M. (1992). Nonspecific lipid transfer protein in castor bean cotyledon cells: subcellular localization and a possible role in lipid metabolism. *Journal of Biochemistry*, *111*, 500-508.

Uchikoba, T., Yonezawa, H., & Kaneda, M. (1998). Cucumisin-like protease from the sarcocarp of Benincasa hispida var, Ryukyu. *Phytochemistry*, *49*, 2215-2219.

Uritani, I., & Stahmann, M. A. (1961). Changes in nitrogen metabolism in sweet potato with black rot. *Plant Physiology*, *36*, 770-782.

Van Breusegem, F., & Dat, J. F. (2006). Reactive oxygen species in plant death. *Plant Physiology*, *141*, 384–390.

Van Gelder, R. N., Zastrow, M. E. von, Yool, A., Dement, W. C., Barchas, J. D., & Eberwine, J. H. (1990). Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci. USA*, *87*, 1663-1667.

Van Grundy, S. D., & Walker, J. C. (1957). Relation of temperature and host nutrition to angular leaf spot of cucumber. *Phytopathology*, *47*, 615-619.

Vance, J. E., & Steenbergen, R. (2005). Metabolism and functions of phosphatidylserine. *Progress in Lipid Research*, *44*, 207-234.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, *3*, 34.1-34.11.

Velculescu, V. E., Zhang, L., Vogelstein, B., & Kinzler, K. W. (1995). Serial analysis of gene expression. *Science*, *270*, 484-487.

Verdoucq, L., Vignols, F., Jacquot, J-P., Chartier, Y., & Meyer, Y. (1999). In vivo characterization of a thioredoxin h target protein defines a new peroxiredoxin family. *The Journal of Biological Chemistry*, *274*(28), 19714-22.

Vermeulen, J., Derveaux, S., Lefever, S., De Smet, E., De Preter, K., Yigit, N., et al. (2009). RNA pre-amplification enables large-scale RT-qPCR gene-expression studies on limiting sample amounts. *BMC research notes*, *2*, 235.

Vignols, F., Lung, G., Pammi, S., Tremousaygue, D., Grellet, F., Kader, J.-C., et al. (1994). Characterization of a rice gene coding for a lipid transfer protein. *Gene, 142*, 265-270.

Voelker, D. R. (1997). Phosphatidylserine decarboxylase. *Biochimica et Biophysica Acta, 1348*, 236-244.

von Rad, U., Meuller, M. & Durne, R. (2005). Evaluation of natural and synthetic immunity by microarray technology. *New Phytologist, 165*, 191-202.

Wakibia, J. G., Anderson, R. J., & Keats, D. W. (2001). Growth rates and agar properties of three Gracilarioids in suspended open-water cultivation in St. Helena Bay, South Africa. *Journal of Applied Phycology, 13*, 195-207.

Walker, N. (2002). A technique whose time has come. *Science, 296*, 557-558.

Wang, E., Miller, L. D., Ohnmacht, G. A., Liu, E. T., & Marincola, F. M. (2000). High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol, 18*, 457-459.

Wang, G.-G., W. Lin, L.-J. Zhang, X.-J. Yan, A. D.-L. D. (2004). Programmed cell death in *Laminaria japonica* (Phaeophyta) tissues infected with alginic acid decomposing bacterium. *Prog. Nat. Sci., 14*, 1064-1068.

Wang, H.-T., Rahaim, P., Robbins, P., & Yocum, R. R. (1994). Cloning, sequence, and disruption of the Saccharomyces diastaticus DAR1 gene encoding a Glycerol-3-Phosphate Dehydrogenase. *Journal of Bacteriology, 176*(22), 7091-7095.

Wang, W., Cho, H. S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H. H., et al. (2002). Structural Characterization of the Reaction Pathway in Phosphoserine Phosphatase: Crystallographic "snapshots" of Intermediate States. *Journal of Molecular Biology, 319*, 421-431.

Wang, W., Kim, R., Jancarik, J., Yokota, H., & Kim, S.-H. (2001). Crystal Structure of Phosphoserine Phosphatase from Methanococcus jannaschii, a Hyperthermophile, at a Resolution of 1.8 A. *Structure, 9*, 65-71.

Weinberger, F. (2007). Pathogen-induced defense and innate immunity in macroalgae. *Biol. Bull., 213*, 290-302.

Weinberger, F., & Freidlander, M. (2000). Endogenous and exogenous elicitors of a hypersensitive response in *Gracilaria conferta* (Rhodophyta). *Journal of Applied Phycology, 12*, 139-145.

Weinberger, F., & Friedlander, M. (2000). Response of *Gracilaria conferta* (Rhodophyta) to oligoagars results in defense against agar-degrading epiphytes. *Journal of Phycology, 36*, 1079-1086.

Weinberger, F., Friedlander, M., & Gunkel, W. (1994). A bacterial facultative parasite of *Gracilaria conferta*. *Dis. Aquat. Org.*, *18*, 135-141.

Weinberger, F., Friedlander, M., & Hoppe, G. (1999). Oligoagars elicit a physiological response in *Gracilaria conferta* (Rhodophyta). *Journal of Phycology, 35*, 747-755.

Weinberger, F., Guillemin, M.-L., Destombe, C., Valero, M., Faugeron, S., Correa, J. A., et al. (2010). Defense Evolution in the Gracilariaceae (Rhodophyta): Substrate-Regulated Oxidation of Agar Oligosaccharides Is More Ancient Than the Oligoagar-Activated Oxidative Burst. *Journal of Phycology*, *46*(5), 958-968.

Weinberger, F., Hoppe, H.-G., & Friedlander, M. (1997). Bacterial induction and inhibition of a fast necrotic response in *Gracilaria conferta* (Rhodophyta). *Journal of Applied Phycology*, *9*, 277-285.

Weinberger, F., Leonardi, P., Miravalles, A., Correa, J. A., Lion, U., Kloareg, B., et al. (2005). Dissection of two distinct defense-related responses to agar oligosaccharides in *Gracilaria chilensis* (Rhodophyta) and *Gracilaria conferta* (Rhodophyta). *Journal of Phycology, 41*, 863-873.

Welchman, R. L., Gordon, C., & Mayer, R. J. (2005). Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Mol. Cell. Biol.*, 599-609.

Wen, X., Fuhrman, S., Michaels, G. S., Carr, D. B., Smith, S., Barker, J. L., et al. (n.d.). Large-scale temporal gene expression mapping of central nervous system development. *Proc. Natl. Acad. Sci. USA, 95*, 334-339.

Wheeler, W. N., M., N., & Woessner, J. W. (1979). Marine agriculture: progress and problems. *Cellular and Molecular Life Sciences*, *35*(4), 433-435.

White, K. P., Rifkin, S. A., Hurban, P., & Hogness, D. S. (1999). Microarray analysis of Drosophila development during metamorphosis. *Science*, *286*, 2179–2184.

Wildsmith, S. E., Archer, G. E., Winkley, A. J., Lane, P. W., & Bugelski, P. J. (2001). Maximization of signal derived from cDNA microarrays. *Biotechniques, 30*(1), 202-206.

Wilkins, T. A., & Smart, L. W. (1996). Isolation of RNA from plant tissue. In P. A. Krieg (Ed.), *A laboratory guide to RNA: isolation, analysis and synthesis.* (pp. 21-42). New York: Wiley-Liss, Inc.

Wilkinson, M. G., & Millar, J. B. (2000). Control of the eukaryotic cell cycle by MAP kinase signaling pathways. *FASEB J.*, *14*, 2147-2157.

Williams, E., & Bunkley-Williams, L. (1990). The worldwide coral reef bleaching cycle and related sources of coral mortality. *Atoll Res. Bull.*, *335*(1), 1-73.

Williams, J. L., Hathaway, C. A., Kloster, K. L., & Layne, B. H. (1997). Low power, type II errors, and other statistical problems in recent cardiovascular research. *Am J Physiol Heart Circ Physiol., 273*, H487-H493.

Wingler, A., Lea, P. J., Quick, W. P., & Leegood, R. C. (2000). Photorespiration: metabolic pathways and their role in stress protection. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 355*, 1517-1529.

Winzeler, E. A., Schena, M., & Davis, R. W. (1999). Fluorescence based expression monitoring using microarrays. *Methods in Enzymology, 306*, 3-19.

Wittwer, C. T., Herrmann, M. G., Moss, A. A., & Rasmussen, R. P. (1997). Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques, 22*, 130-138.

Wodicka, L., Dong, H., Mittmann, M., Ho, M. H., & Lockhart, D. J. (1997). Genomewide expression monitoring in *Saccharomyces cerevisiae*. *Nat Biotechnol, 15*, 1359-1367.

Wolfenden, R. (1970). Binding of substrate and transition state analogs to triosephosphate isomerase. *Biochemistry, 9*, 3404-3407.

Woltering, E. J., Der Bent, A. van, & Hoeberichts, F. A. (2002). Do plant caspases exist? *Plant Physiology, 130*, 1764–1769.

Wong, J. H., Besse, I., Lozano, R. M., Jiao, J. A., Yee, B. C., Peters, A., et al. (1996). New evidence for a role for thioredoxin h in germination and seedling development. *Plant Physiology, 111*, 103.

Wong, M. L., & Medrano, J. F. (2005). Real-time PCR for mRNA quantitation. *Biotechniques, 39*, 1-11.

Wood, Z. A., Poole, L. B., & Karplus, P. A. (2003). Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science, 300*, 650–653.

Wood, Z. A., Schro, E., Harris, J. R., & Poole, L. B. (2003). Structure, mechanism and regulation of peroxiredoxins. *Trends in biochemical sciences, 28*(1), 32-40.

Woolf, P. J., & Wang, Y. (2000). A fuzzy logic approach to analyzing gene expression data. *Physiol Genomics, 3*, 9-15.

Worley, J., Bechtol, K., Penn, S., Roach, D., & Hanzel, M. (2000). A system approach to fabricating and analysing DNA microarrays. In M. Schena (Ed.), *Microarray Chip Technology* (pp. 65-86). Eaton Publishing.

Wu, S.-hsing, Ramonell, K., Gollub, J., & Somerville, Shauna. (2001). Plant gene expression profiling with DNA microarrays. *Plant Physiol. Biochem., 39*, 917-926.

Wurmbach, E., Yuen, T., & Sealfon, S. C. (2003). Focused microarray analysis. *Methods*, *31*, 306-316.

Xu, Q., Fu, H.-H., Gupta, R., & Luan, S. (1998). Molecular characterization of a Tyrosine-specific protein phosphatase encoded by a stress-responsive gene in Arabidopsis. *Plant Cell*, *10*, 849-858.

Yamada, M. (1992). Lipid transfer proteins in plants and microorganisms. *Plant and Cell Physiology.*, *33*, 1-6.

Yamazaki, D., Motohashi, K., Kasama, T., Hara, Y., & Hisabori, T. (2004). Target proteins of the cytosolic thioredoxins in *Arabidopsis thaliana*. *Plant and Cell Physiology*, *45*(1), 18-27.

Yang, Y., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., et al. (2002). Normalization for cDNA microarray data : a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.*, *30*(4e15), 1-10.

Yano, A., Suzuki, K., & Shinsh, H. (1999). A signaling pathway, independent of the oxidative burst, that leads to hypersensitive cell death in cultured tobacco cells includes a serine protease. *Plant J.*, *18*, 105-109.

Yokoya, N., & Oliveira, E. C. (1993). Effects of salinty on the growth rate, morphology and water content of some Brazilian red algae of economic importance. *Ciencias Marinas (Mexico)*, *13*, 47-56.

Zaneveld, J. S. (1959). The utilization of marine algae in tropical south and east Asia. *Econ. Bot*, *1*, 89-131.

Zeidler, D., Zahringer, U., Gerber, I., Dubery, I., Hartung, T., Bors, W., et al. (2004). Innate immunity in *Arabidopsis thaliana*: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. *Proc. Natl. Acad. Sci. USA*, *101*, 15811–15816.

Zhang, S. E. A. (2000). Multiple levels of tobacco WIPK activation during the induction of cell death by fungal elicitins. *Plant J.*, *23*, 339–347.

Zhang, S., & Klessig, D. F. (1998). Resistance gene N-mediated de novo synthesis and activation of a tobacco mitogen-activated protein kinase by tobacco mosaic virus infection. *Proc. Natl. Acad. Sci. USA*, *95*, 7433–7438.

Zhang, S., & Klessig, D. F. (2001). MAPK cascades in plant defense signaling. *Trends in Plant Science*, *6*(11), 520-527.

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, *7*, 203-214.

Zhao, C., Johnson, B. J., Kositsup, B., & Beers, E. P. (2000). Exploiting secondary growth in Arabidopsis: construction of xylem and bark cDNA libraries and cloning of three xylem endopeptidases. *Plant Physiol., 123*, 1185-1196.

Zhao, H., Hastie, T., Whitfield, M. L., & Jeffrey, S. S. (2002). Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis. *BMC Genomics*, *15*, 1-15.

Zhou, W., Kolb, F. L., & Riechers, D. E. (2005). Identification of proteins induced or upregulated by Fusarium head blight infection in the spikes of hexaploid wheat (Triticum aestivum). *Genome*, *48*, 770-780.

Zhou, Y.-H., & Ragan, M. A. (1995). Characterization of the polyubiquitin gene in the red marine alga *Gracilaria verrucosa*. *Biochimica et Biophysica Acta*, *1261*, 215-222.

Zinser, E., Sperka-Gottlieb, C. D. M., Fasch, E. V., Kohlwein, S. D., Paltauf, F., & Daum, G. (1991). Phospholipid synthesis and lipid composition of subcellular membranes in the unicellular eukaryote Saccharomyces cerevisiae. *Journal of Bacteriology*, *173*, 2026-2034.

Zou, J., Rodriguez-Zas, S., Aldea, M., Li, M., Zhu, J., Gonzalez, D. O., et al. (2005). Expression Profiling Soybean Response to Pseudomonas syringae Reveals New Defense-Related Genes and Rapid HR-Specific Downregulation of Photosynthesis. *Molecular plant-microbe interactions : MPMI*, *18*(11), 1161-74.

APPENDIX A

MEDIA, BUFFERS AND SOLUTIONS

All media were autoclaved at 121°C prior to use, unless otherwise specified.

Water used for making all solutions, media and diluting buffers was purified using a Milli-RO Plus (Millipore) water purification system.

Ultrapure water used was obtained by further purification of the above water using a Milli-Q Plus (Millipore) water purification system.

A.1 MEDIA

A.1.1 Luria Agar

Tryptone (Biolab)	10 g
Yeast Extract (Biolab)	5 g
NaCl (Saarchem)	10 g
Agar (Biolab)	15 g
Water to	1000 ml

Autoclave

A.1.2 Luria Broth

Tryptone (Biolab)	10 g
Yeast Extract (Biolab)	5 g
NaCl (Saarchem)	10 g
Water to	1000 ml

Autoclave

A.1.3 Artificial Sea Water (ASW)

NaCl (Saarchem)	24.7 g
$MgCl_2.6H_2O$ (Saarchem)	4.7 g
KCl (Saarchem)	0.66 g
$CaCl_2.2H_2O$	1.9 g
$MgSO_4.7H_2O$ (Saarchem)	6.3 g
$NaHCO_3$	0.18 g
Water to	1000 ml

Autoclave

A.1.3.1 Fe-Solution

$Fe(NH_4)_2(SO_4)_2.6H_2O$	702 mg
Na_2 EDTA	600 mg
Water to	1000 ml

A.1.3.2 PII metal solution

Na_2 EDTA	100 mg
H_3BO_3	114 mg
$FeCl_3.6H_2O$	4.9 mg
$MnSO_4$	16.4 mg
$ZnSO_4.7H_2O$	2.2 mg
$CoSO_4.7H_2O$	0.48 mg
Water to	100 ml

A.1.3.3 PES-enriched seawater medium (1/3 strength) (Provasoli, 1968)

$NaNO_3$	350 mg
Na_2glycerophosphate 5H2O	50 mg
Fe solution	25 ml
PII	25 ml
Vitamin B_{12}	10 µg
Thiamine	0.5 mg
Biotin	5 µg
Tris buffer (Sigma Co.)	500 mg
Water to	100 ml

Adjust pH to 7.8 prior to autoclaving
Store at 4°C
Add 6.6 ml per 1 L ASW

A.2 GENERAL BUFFERS

A.2.1 Tris-Acetate-EDTA (TAE) Buffer (50 X stock)

Tris base (Roche)	242 g
Glacial acetic acid (Saarchem)	57.1 ml
EDTA (0.5M, pH 8) (Saarchem)	100 ml
Water to	1000 ml

Use a 1 X solution as the running buffer for DNA agarose gel electrophoresis

A.2.2 Tris-EDTA (TE) buffer (pH 8) (1 M stock)

Tris-Cl (1 M, pH 8)	1 ml
EDTA (0.5 M, pH 8)	200 µl
Water to	100 ml

A.2.3 DNA gel tracking dye (6 X)

Bromophenol blue (Saarchem)	62.5 mg
Sucrose (Saarchem)	10 g
Water	15 ml
EDTA (0.5 M, pH 8)	1 ml

Adjust final volume to 25 ml with water
Autoclave

A.2.4 Phosphate buffered saline [PBS] (10 X)

NaCl	80.0 g
KCl	2.0 g
$Na_2HPO_4.7H_2O$	17.8 g
KH_2PO_4	2.4 g
Water to	1000 ml

Adjust pH to 7.4 with NaOH
Autoclave

A.3 SOLUTIONS

A.3.1 Solutions for Plasmid Isolations

A.3.1.1 Chloramphenicol (Cm) (30 mg/ml)

Chloramphenicol (Roche)	30 mg
Ethanol	10 ml

Store in 1 ml aliquots at -20°C

A.3.1.2 Kanamycin (Kan) (30 mg/ml)

Kanamycin (Roche)	30 mg
Water	10 ml

Store in 1 ml aliquots at -20°C

A.3.1.3 Solution 1 (10 X Stock)

Tris-Cl (1 M, pH 8)	25 ml
Glucose (20 % w/v) (Saarchem)	45.5 ml
EDTA (0.5 M, pH 8)	20 ml
Water	9.5 ml

Solution 2

NaOH (10 N) (Saarchem)	2 ml
Sodium Dodecyl Sulphate (25 %, w/v) (Saarchem)	4 ml
Water	94 ml

Prepare this solution weekly

Solution 3

K-acetate (Saarchem)	147 g
Water	250 ml

Adjust pH to 4.8 – 5.0 with acetic acid
Adjust final volume to 500 ml with water

A.3.2 Solutions for RNA isolations

A.3.2.1 Diethylpyrocarbonate water (DEPC-dH_2O)

DEPC (Sigma)	1 ml
Distilled water	1000 ml

Shake solution vigorously until no DEPC droplets remain
Leave solution in the fumehood overnight
Autoclave

A.3.2.2 RNA extraction buffer (Azvedo *et al.*, 2003)

Tris HCl (1 M, pH 8)	1.5 ml
CTAB (10 %, w/v) (USB)	3.0 ml
EDTA (0.5 M, pH 8)	0.9 ml
NaCl (5 M)	6.0 ml
Spermidine (10 %, w/v) (Sigma)	75 ul
Polyvinylpolypyrrolidinone (PVPP) (Sigma)	0.3 g
β-Mercaptoethanol	0.3 ml

Add PVPP and β-Mercaptoethanol just before use
Add Proteinase K (10 mg/ml) (Sigma) to a final concentration of 1.5 mg/ml

A.3.2.3 10 M LiCl

LiCl (Saarchem)	42.39 g
Water	100 ml

Autoclave

A.3.2.4 2 M LiCl

LiCl (Saarchem)	8.48 g
Water	100 ml

Autoclave

A.3.2.5 1.2% RNA Agarose Formaldehyde gel

Agarose	0.72 g
DEPC-dH_2O	43.92 ml

Boil solution to dissolve agar
Allow to cool before adding the following:

MOPS (10 X, pH 7)	6 ml
Formaldehyde	10.08 ml

Pour gel and allow to polymerize

A.3.2.6 10X MOPS

MOPS (Sigma)	20 g
Na Acetate (Saarchem)	1 g
EDTA (0.5 M, pH 8)	10 ml
DEPC- dH_2O	470 ml

Adjust pH to 7 with NaOH
Filter sterilize through a 0.22 µm Millipore filter using a 60 cc syringe
Store in the dark at 4 ºC
Use 1X solution as running buffer for RNA agarose gels

A.3.2.6 RNA sample application buffer

MOPS (10 X, pH 7)	300 µl
Formaldehyde (37 %) (Saarchem)	80 µl
Formamide (Saarchem)	900 µl
Ethidium Bromide (10 mg/ml) (Sigma)	2 µl
Dye	220 µl

Prepare dye as follows:

Xylene cyanol	50 mg
Bromophenol blue	50 mg
DEPC-dH_2O	1 ml

Add RNA sample buffer at a ratio of 1:2 for RNA sample : sample buffer
Heat samples at 65°C for 15 min
Snap cooling on ice before loading into gel wells

A.3.3 Solutions for Microarray Labelling and Hybridizations

A.3.3.1 Coupling buffer

Carbonate buffer changes composition over time
Store as 100 µl aliquots at -20 °C to allow longer term storage or at room temperature for a maximum of two weeks

a) 0.1 M Sodium carbonate (anhydrous)

Na_2CO_3 (Saarchem)	1.06 g
MilliQ dH_2O	100 ml
Autoclave	

b) 0.1 M Sodium bicarbonate (sodium hydrogen carbonate)

$NaHCO_3$ (Saarchem)	0.84 g
MilliQ dH_2O	100 ml
Autocale	
Autoclave	

Combine 4 ml 0.1M carbonate solution, 46 ml 0.1M bicarbonate solution and 150ml MilliQ dH_2O
pH to 9.2 using concentrated 1M HCl

A.3.3.2 Cy-dye esters

Cy3 and Cy5 NHS ester dyes (AmershamPharmacia) are provided as a dried powder
Re-suspend a tube in 12 µl DMSO
Dry completely using a speedivac and store completely desiccated at 4 °C
Wrap in foil and keep covered in order to prevent photo-bleaching of the dyes

A.3.3.3 4 M Hydroxylamine

NH_2OH (Sigma)	0.05 g
MilliQ dH_2O	0.5 ml
Use immediately or store at -20 °C	

A.3.3.4 Target hybridization mixture

COT1-DNA (Life Technologies) (20 µg.µl^{-1})	1.0 µl
Poly(A)-DNA (Sigma) (20 µg.µl^{-1})	1.0 µl
Cy3 target	5.0 µg
Cy5 target	5.0 µg
MilliQ dH_2O	to 30 µl
2 X Hybridization buffer	30 µl

Heat mixture at 90 °C for 3 minutes before injecting onto microarray slide

A.3.3.5 2 X Hybridization buffer

SSC (Sigma) (20 X)	100 µl
SDS (Life Technologies) (20 %)	2 µl
Formamide (Life Technologies)	100 µl

A.3.3.6 20 X SSC

NaCl (Saarchem)	175 g
$C_6H_5Na_3O_7.2H_2O$ (Saarchem)	88.2 g
Water to	90 ml

Adjust pH to 7.4 before making the volume up to 1 L
Add 2 ml of DEPC-dH_2O before autoclaving

A.3.3.7 20 % SDS

SDS (Life Technologies)	20 g
Water to	100 ml

A.3.3.8 10 % BSA

BSA (Sigma)	10 g
Water to	100 ml

Filter-sterilize and store at -20 °C

A.3.3.9 Pre-hybridization buffer

SSC (Sigma) (20 X)	50 µl
SDS (Life Technologies) (20 %, w/v)	1 µl
BSA (10 %)	20 µl
Water to	129 µl

A.3.3.10 Low stringency wash buffer

SSC (Sigma) (20 X)	40 ml
SDS (Life Technologies) (25 %, w/v)	10 ml
Water to	350 ml

A.3.3.11 Medium stringency wash buffer

SSC (Sigma) (20 X)	10 ml
Water to	390 ml

A.3.3.12 High stringency wash buffer

SSC (Sigma) (20 X)	0.5 ml
Water to	199.5 ml

A.3.4 Solutions for Recombinant Protein Purification

A.3.4.1 Solutions for recombinant protein purification under native conditions

A.3.4.1.1 Lysis Buffer

NaH_2PO_4 [0.05M] (Saarchem)	6.9 g
NaCl [0.3 M] (Saarchem)	17.54 g
Imidazole [0.01 M] (Sigma)	0.34 g
Water to	1000 ml

Adjust pH to 8.0 using NaOH
Autoclave
De-gas by filtering solution through 0.22 µM pore-size filter under vacuum

A.3.4.1.2 Wash / Equilibration Buffer

NaH_2PO_4 [0.05M] (Saarchem)	6.9 g
NaCl [0.3 M] (Saarchem)	17.54 g
Imidazole [0.02 M] (Sigma)	0.68 g
Water to	1000 ml

Adjust pH to 8.0 using NaOH
Autoclave and de-gas by filtering solution through 0.22 µM pore-size filter under vacuum

A.3.4.1.3 Elution Buffer

NaH_2PO_4 [0.05M] (Saarchem)	6.9 g
NaCl [0.3 M] (Saarchem)	17.54 g
Imidazole [0.5 M] (Sigma)	34.0 g
Water to	1000 ml

Adjust pH to 8.0 using NaOH

Autoclave and de-gas by filtering solution through 0.22 µM pore-size filter under vacuum

A.3.4.2 Solutions for recombinant protein purification under denaturing conditions

A.3.4.2.1 Lysis / Equilibration / Wash Buffer 1

Tris base [0.02 M] (Roche)	2.4 g
NaCl [0.5 M] (Saarchem)	29.22 g
Imidazole [0.005 M] (Sigma)	0.34 g
Urea [8 M] (Sigma)	96.02 g
Water to	1000 ml

Adjust pH to 8.0 using NaOH
Autoclave
De-gas by filtering solution through 0.22 µM pore-size filter under vacuum

A.3.4.2.2 Wash Buffer 2

Tris base [0.02 M]	2.4 g
NaCl [0.5 M]	29.22 g
Imidazole [0.02 M]	1.36 g
Urea [8 M]	96.02 g
Water to	1000 ml

Adjust pH to 8.0 using NaOH
Autoclave and de-gas by filtering solution through 0.22 µM pore-size filter under vacuum

A.3.4.2.3 Elution Buffer

Tris base [0.02 M]	2.4 g
NaCl [0.5 M] (Saarchem)	29.22 g
Imidazole [0.5 M] (Sigma)	34.0 g
Urea [8 M]	96.02 g
Water to	1000 ml

Adjust pH to 8.0 using NaOH
Autoclave and de-gas by filtering solution through 0.22 µM pore-size filter under vacuum

A.3.4.3 5X Sample Application Buffer

Tris-Cl (pH 6.8)	250 mM
DTT (Promega)	500 mM
SDS (Sigma)	10%
Bromophenol blue (Saarchem)	0.5%
Glycerol	10 ml
Water	5 ml

A.3.5 Solutions for Western Hybridization analyses

A.3.5.1 Blocking buffer (5 % skim milk)

Skim-milk (Elite)	5 g
10 X PBS [1 X] (Section A.2.4)	10 ml
Water to	100 ml

A.3.5.2 10X TBS

Tris base [0.05 M]	50.5 g
NaCl [0.15 M]	73.14 g
Water to	500 ml

Adjust pH to 7.4 with HCl
Autoclave

A.3.5.3 TBS-T (0.1 % Tween in 1X TBS)

Tween 20 (Saarchem)	0.1 ml
10X TBS [1 X]	100 ml
Water to	1000 ml

Make up fresh
Do not autoclave

A.3.5.4 Towbin Buffer

Tris base	3.03 g
Glycine (Merck)	14.42 g
Methanol [20%] (Merck)	200 ml
Water to	1000 ml

A.3.5.5 4X Tris-Cl / SDS solution (pH 6.8)

Tris base [0.5 M]	6.05 g
SDS (Merck)	0.4 g

Adjust pH to 6.8 with 1 M HCl

Water to	100 ml

Filter-sterilize and store at room temperature

A.3.5.6 4X Tris-Cl / SDS solution (pH 8.8)

Tris base [1.25 M]	91.0 g
SDS	2.0 g

Adjust pH to 8.8 with 1 M HCl

Water to	500 ml

Filter-sterilize and store at room temperature

A.3.6 Solutions for Protein Isolation from *G. gracilis*

A.3.6.1 Urea Lysis Buffer (ULB)

Urea [8 M]	24.0 g

Dissolve in 25 ml water by stirring and then add:

CHAPS [2 %] (Calbiochem)	0.25 g
Thiourea (Sigma-Aldrich)	7.6 g

APPENDIX B

DNA PRIMER SEQUENCES AND PCR CYCLE PROFILES

B.1 DNA PRIMER SEQUENCES

All synthetic oligonucleotide sequences were supplied by the Oligonucleotide synthesizing service of the Department of Molecular and Cell Biology, University of Cape Town, Cape Town, South Africa.

B.1.1 Primers used in colony PCR

M13-F (Universal)	5’	GTTGTAAAACGACGGCCAGT	3’
M13-R (Universal)	5’	CACAGGAAACAGCTATGACC	3’

B.1.2 Primers used for sequencing

LIB-F	5’	AACGACCGAGCGCAGCGAGTC	3’
M13-R (Universal)	5’	CACAGGAAACAGCTATGACC	3’

B.1.3 Primers used to assess residual genomic DNA contamination

E18F	5’	CCTGGTTGATCCTGCCAGTGG	3’
E18R	5’	GGAGAAGTCGTAACAAGGTTTCCGTAG	3’

B.2 PCR CYCLE PROFILES

B.2.1 Cycling parameters for colony PCR

1 cycle	Denaturation	96^{o}C	5 min
25 cycles	Denaturation	94^{o}C	30 s
	Annealing	61^{o}C	30 s
	Extension	72^{o}C	3 min
1 cycle	Extension	72^{o}C	7 min
1 cycle	Cooling	4^{o}C	10 min

B.2.2 PCR cycling parameters for DNA sequencing

1 cycle	denaturation	96^{o}C	1 min
25 cycles	denaturation	94^{o}C	30 s
	annealing	50^{o}C	5 s
	extension	60^{o}C	4 min
1 cycle	cool down	4^{o}C	10 min

APPENDIX C

SCRIPTS USED FOR MICROARRAY DATA ANALYSIS
MULTIPLE SEQUENCE ALIGNMENTS
TRANSIT PEPTIDE PREDICTION

C.1 Predefined filter generated in GenePix 6.0.27 Pro software used in image pre-processing

[% > B635+1SD]	> 55	And
[% > B532+1SD]	> 55	And
[Rgn R2 (635/532)]	> 0.5	And
[Flags]	<> [Bad]	And
[Flags]	<> [Absent]	And
[Flags]	<> [Not Found]	And
LCase([ID])	<> "empty"	And
[F635 % Sat.]	< 3	And
[F532 % Sat.]	< 3	And
[Sum of Medians (635/532)]	> 500	

C.2 R script written for use in Bioconductor package LIMMA for normalization of microarray data and identification of differentially expressed genes (compiled by Wiesner Vos)

```
## DATA ANALYSIS using LIMMA

setwd(Choose the working directory which contains all the files to be analysed)

library(limma)
## Load the script that modifies the NormailzeWithinArrays to a maxit = 150

targets <- readTargets()
targets

## Read Image Files and incorporate weights assigned to spots flagged "BAD"
f <- function(x) as.numeric(x$Flags >-49)
RG <- read.maimages(targets$FileName, source = "genepix", wt.fun=f)
RG

## Read gene information (GAL file)
RG$genes <- readGAL()
RG$genes[1:20,]

## Set print layout
RG$printer <- getLayout(RG$genes)
```

```
spottypes<-readSpotTypes()
RG$genes$Status<-controlStatus(spottypes,RG)

## Exploring the raw data
par(mfrow=c(2,2))
plotMA(RG,array=1,ylim=c(-5,5))
plotMA(RG,array=2,ylim=c(-5,5))
plotMA(RG,array=3,ylim=c(-5,5))


## Incorporating weights of spots - Exploring the raw data BEFORE normalization
par(mfrow=c(2,2))
plotMA(RG,array=1,ylim=c(-5,5),zero.weights=FALSE)
plotMA(RG,array=2,ylim=c(-5,5),zero.weights=FALSE)
plotMA(RG,array=3,ylim=c(-5,5),zero.weights=FALSE)

## If more arrays, just change the "array=" in the command according to array number
## zero.weights = FALSE excludes spots with a zero/-ve weight - Spots are not printed by default.  If
you want them
## to be printed, add zero.weigthts=TRUE


## Normalization within arrays

## I used the robustspline method, df = 4 (n-2 and there are 6 treatments)
MA.p<-
normalizeWithinArrays(RG,method="robustspline",bc.method="normexp",controlspots=200,df=4)
plotDensities(MA.p)


## Normalize between Arrays
MA.pAq<-normalizeBetweenArrays(MA.p,method="Aquantile")
plotDensities(MA.pAq)

## MA plots AFTER normalization
par(mfrow=c(2,2))
plotMA(MA.pAq,array=1,ylim=c(-5,5))
plotMA(MA.pAq,array=2,ylim=c(-5,5))
plotMA(MA.pAq,array=3,ylim=c(-5,5))


## Differential expression
## sort by gene ID to get spacing regular

i <- order(MA.pAq$genes$ID)
MA.pAq<- MA.pAq[i,]

## remove controls: Bactin,rc,cc,uc,nc
MA.pAq$M<-MA.pAq$M[MA.pAq$genes$Status=="gene",]
MA.pAq$A<-MA.pAq$A[MA.pAq$genes$Status=="gene",]
MA.pAq$genes<-MA.pAq$genes[MA.pAq$genes$Status=="gene",]

#average over 4 replicate spots on each array (M and A values)
fit1 <- lmFit(MA.pAq,design=diag(ncol(MA.pAq)),ndups=4,spacing=1,correlation=0)
fit2 <- lmFit(MA.pAq$A, design=diag(1, dim(MA.pAq)[2]), ndups=4, spacing=1,correlation=0)
```

```
MA.pAq$M<-fit1$coef
MA.pAq$A<-fit2$coef
MA.pAq$weights<-NULL
MA.pAq$genes<-fit1$genes


design<-modelMatrix(targets,ref="Healthy")
fit<-lmFit(MA.pAq,design)
fit<-eBayes(fit)
toptable<-topTable(fit,adjust="none",number=1000)
toptable[1:100,]

write.table(toptable,file="C:\\RTablelimma.txt",row.names=FALSE,sep="\t")
```

C.3 Multiple sequence alignment for *Gracilaria gracilis* Phosphoserine phosphatase (PSP) DNA sequence

The *G. gracilis* PSP sequence was aligned to three other PSP sequences (human NP_004568, plant NP_973858 and bacterial NC_014002) using DNAMAN (Version 4.13).

```
HumanPSP    ....................................MV      2
AthalianaPSPMEALTTSRVVPVQVPCRKLSSLFANFSCLELRRYPCRGLV     40
MethanoPSP  ....................................       0
GgracilisPSP....................................       0
Consensus
HumanPSP    SHSELRKLFYSAD...........................      15
AthalianaPSPSIMNHPKLLRPVTASVQPHELSTLGHEGNIVPSKEILDLW     80
MethanoPSP  .......................MIRFDEVTLASNSNR     15
GgracilisPSP....................................       0
Consensus
HumanPSP    ....AVCFDVDSTVIREEGIDELAKICGVEDAVSEMTRRA     51
AthalianaPSPRSVEAVCFDVDSTVCVDEGIDELAEFCGAGKAVAEWTARA    120
MethanoPSP  KFSKLIIFDMDSTLIDAECIDELALAAGVADKVSQITDRT     55
GgracilisPSP....MVVFDLSWTLVQCDSTNVLLHAADVDVPKEEEQKFR     36
Consensus           fd   t     i  l
HumanPSP    MGGAVPFKAALTERLALIQPSREQVQRLIAEQPPHLTPGI     91
AthalianaPSPMGGSVPFEEALAARLSLFKPSLSKVEEYLDKRPPRLSPGI    160
MethanoPSP  MRGELDYNLALLERVKLLKGLEITKATEAVEKIELM.PGA     94
GgracilisPSPSGQISGSEWVRFRVQLLKGFDAHTVNSKAVQSLVYTNGAV     76
Consensus                  l
HumanPSP    RELVSRLQERNVQVFLISGGFRSIVEHVASKLNIPATNVF    131
AthalianaPSPEELVKKLRANNIDVYLISGGFRQMINPVASILGIPRENIF    200
MethanoPSP  KELLEHVQSLGYKTAMLSGGFTLSSDRVAKLLNIDYVY..    132
GgracilisPSPQLCKGLKRLGCKLAIVSSGSKDIAMAAKEALHLDFVFXNV    116
Consensus                    sg
HumanPSP    ANRLKFYFNGEYAGFDETQPTAESGGKGKVIKLLKEKFHF    171
AthalianaPSPANNLLFGNSGEFLGFDENEPTSRSGGKAKAVQQIRKGRLY    240
MethanoPSP  ANTLEVKNGYLTGVVSGPMTQNLSKEFAFEEIARKNGFLP    172
GgracilisPSPIEIDSAGRFTGTLKEPVVDAQRKADLIAMLAMQERIDTEQ    156
Consensus
HumanPSP    KKIIMIGDGATDMEACPP..ADAFIGFGGNVIRQQVKDNA    209
AthalianaPSPKTMAMIGDGATDLEARKPGGADLFICYAGVQLREAVAANA    280
MethanoPSP  EDCIVVGDGANDVCIFKRAGYSIAFNPKPILHQHADVIIS    212
GgracilisPSPVIAVGDGPVSAKMLESVGMSIAFDQPDALDSVHSGRIGSK    196
Consensus         g
HumanPSP    KWYITDFVELLGELEE.....                    225
AthalianaPSPDWLIFKFESLINSLD......                    295
MethanoPSP  KKDLRALISVIDSLQ......                    227
GgracilisPSPSLASVLYLLGVSGHDFRTVTA                    217
Consensus
```

C.4 Alignment of *Gracilaria gracilis* and *Chondrus crispus* peroxiredoxin-like genes

DNAMAN (Version 4.13) was used to perform a Fast alignment of the two DNA sequences.

Chondrus crispus peroxiredoxin-like gene: Accession number CO650899.

Ktuple=2 Gap_penalty=7

Upper line: *Gracilaria gracilis* peroxiredoxin-like gene, from 10 to 510
Lower line: *Chondrus crispus* peroxiredoxin-like gene, from 28 to 528

Upper line:Lower line identity= 59.88%(300/501) gap=0.00%(0/501)

```
10    CATACTTTCAACTCTCACTGAATCCCAGAACCAGCCCATCGACATGACTGTCTCTACCGC
      ||      |  | | | | |  ||||   |||| ||  || ||||| |   |    ||||
28    CACCTCACCCTCACCCTCCGCCTCCCCCGACCAACCTCTCAACATGCCCACCGACGCCGC

70    ACCGTCGCTCGCCCCGCAGCTCCAGGCCTTCAAGGAAAAGTTCCTGGGTTCAGCGTCACC
      |    | || ||| |  | |||   |||||| || ||  |||||       |  | ||
88    AGGCCCCCTGGCCACCGAACTCGCTGCCTTCGAGCAATCGTTCCGCTCCAAGACCCCGCC

130   AGAGATGGTGGCCACCGTTTCCAGGACGTCGGCCGAGCTGGAGGAAGCGAACCCGATCAA
       |||||| |   | || |   |    |||| |||| ||| ||||       | |    |
148   CGAGATGATCAGCGCCATCGACGCCTCGTCCGCCGCGCTCGCGGCCCGCTTCACCCCGAC

190   CCCGCTCAAGGTCGGCGATCGCGCGCCGGACTTCACGCTGCCGAACGCCACCGGCAAGCA
      |||||    | | |||      |    |   ||| ||||||| |||||||  ||| ||
208   CCCGCCGTCGCTGGGCACCAAGGTCGAGTCGTTCTCGCTGCCCAACGCCAAGGGCGAGAC

250   GGTGACGCTCTCCTCGCTGCTCGAGAAGCATTCCGCCGTAGTGATGACGTGGTATCGTGG
       ||| |||||  ||||||||||   |     ||||||| ||  |||| |  || || ||
268   CGTGTCGCTCGACTCGCTGCTCAGCAGCAGCACCGCCGTCGTCCTGACCTACTACCGCGG

310   AGGGTGGTGCCCGTACTGCAATCTGGCGCTGCGCGCCTTCACGCAGCTCAACGGCGAGAT
        | ||||||||||||||||| || || ||| | ||| || |  ||  |  ||  |||||
328   TAGCTGGTGCCCGTACTGCAACCTTGCCCTGAGGGCCCTCGCCAAGTCCGTCGATGAGAT

370   TGAAGCCAGCGGCGCCAAGTTGGTGGCGTTGACGCCGGAGACGGCTGACGAGAGCCTTTC
        | |||  |||||| |   |||| ||  | |  ||||||||| | |||||||| ||  |
388   CAAGGCCCTCGGCGCTACCCTGGTCGCTCTCAGCCCGGAGACGCCCGACGAGAGTCTCAC

430   CACTTCGGAGAAGAACGAGCTGAATTTCGAGGTGCTGTCAGATGACGGCTCGACTGTTGC
      |||   |||||||||||||| |    | |||||| || || ||||| ||       || ||
448   CACCNAGGAGAAGAACGAGATCCCGTACGAGGTCCTCTCCGATGAGGGGCTTGTCGTCGC

490   CGAGAAGTTCGGCGTTGTGTT
       |  ||||| ||||| | |||
508   AGGCAAGTTGGGCGTCGCGTT
```

C.5 ExPASy tools (http://expasy.org/tools/) used to predict transit peptides for *G. gracilis* thioredoxin DNA sequence

```
### chlorop v1.1 prediction results ############################
Number of query sequences: 1

Name                  Length       Score  cTP       CS-    cTP-
                                                    score  length
-------------------------------------------------------------------
Sequence                 462     0.464   -     4.057     83
-------------------------------------------------------------------
```

Name	Name of the submitted sequence, truncated if longer than 11 characters
Length	Length (bp) of the submitted sequence
Score	The output score from the second step network. The prediction cTP/no cTP is based solely on this score.
cTP	Tells whether or not this is predicted as a cTP-containing sequence; "Y" means that the sequence *is* predicted to contain a cTP; "-" means that is predicted *not* to contain a cTP.
CS-score	The MEME scoring matrix score for the suggested cleavage site
cTP-length	Predicted length of the presequence (**Please note** that the prediction of the transit peptide length is carried out and presented *even if* its presence is not predicted).

```
### targetp v1.1 prediction results ###################################
Number of query sequences:  1
Cleavage site predictions not included.
Using PLANT networks.

Name                  Len     cTP    mTP     SP  other  Loc  RC
-------------------------------------------------------------------
Sequence              462   0.103  0.102  0.029  0.843   _    2
-------------------------------------------------------------------
cutoff                      0.000  0.000  0.000  0.000
```

Name	Sequence name truncated to 20 characters
Len	Sequence length
cTP, mTP, SP, other	Final NN scores on which the final prediction is based (Loc, see below). Note that the scores are not really probabilities, and they do not necessarily add to one. However, the location with the highest score is the most likely according to TargetP, and the relationship between the scores (the reliability class, see below) may be an indication of how certain the prediction is.
Loc	Prediction of localization, based on the scores above; the possible values are: **C** Chloroplast, i.e. the sequence contains **cTP**, a chloroplast transit peptide; **M** Mitochondrion, i.e. the sequence contains **mTP**, a mitochondrial targeting peptide; **S** Secretory pathway, i.e. the sequence contains **SP**, a signal peptide; **_** Any other location; ***** "no prediction"; indicates that cutoff restrictions were set and the winning network output score was below the requested cutoff for that category.

RC Reliability class, from 1 to 5, where 1 indicates the strongest prediction. RC is a measure of the size of the difference ('diff') between the highest (winning) and the second highest output scores. There are 5 reliability classes, defined as follows:
1 : diff > 0.800
2 : 0.800 > diff > 0.600
3 : 0.600 > diff > 0.400
4 : 0.400 > diff > 0.200
5 : 0.200 > diff
Thus, the lower the value of RC the safer the prediction.

```
###### Predotar v. 1.03 using plastid prediction results #################
```

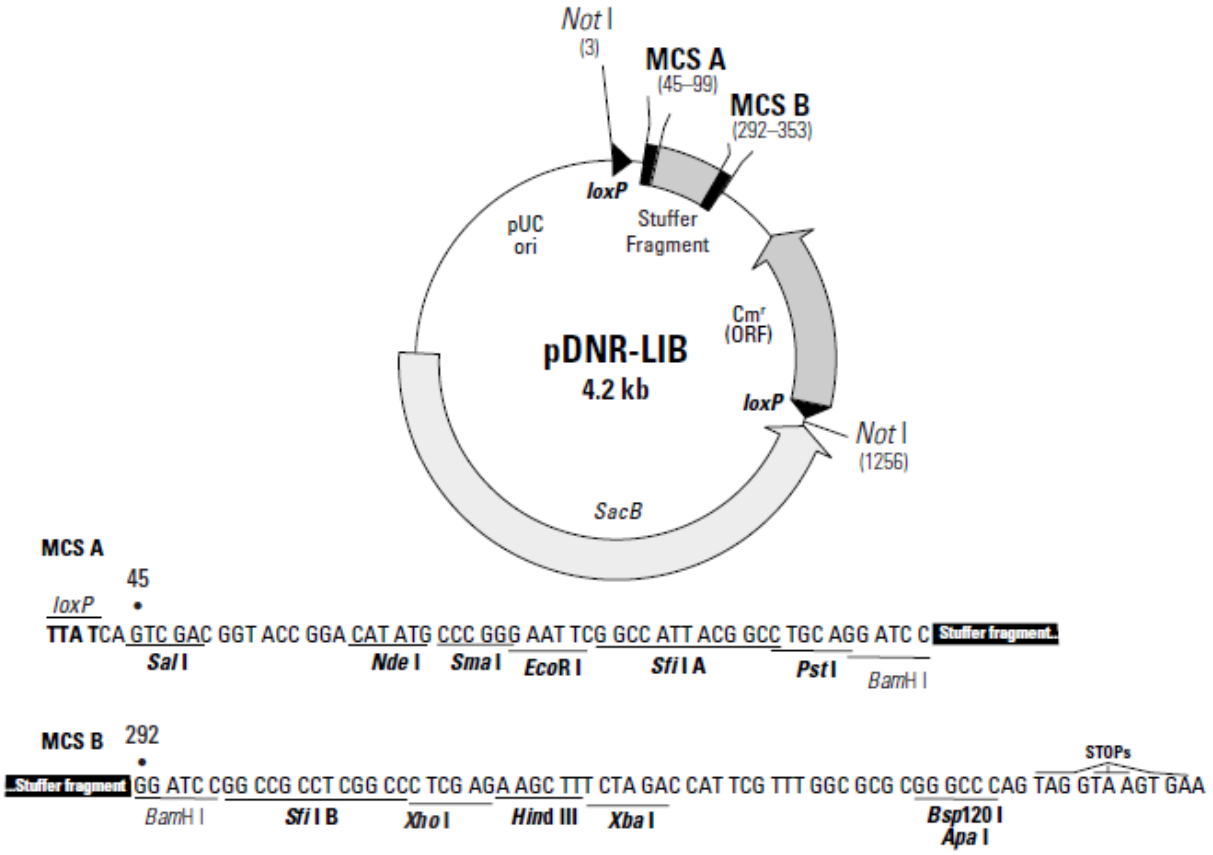
Sequence	Mitochondrial	Plastid	ER	Elsewhere	Prediction

```
Thioredoxin  Discarding thioredoxin: no terminal Met
```

APPENDIX D

CLONING VECTOR PLASMID MAPS

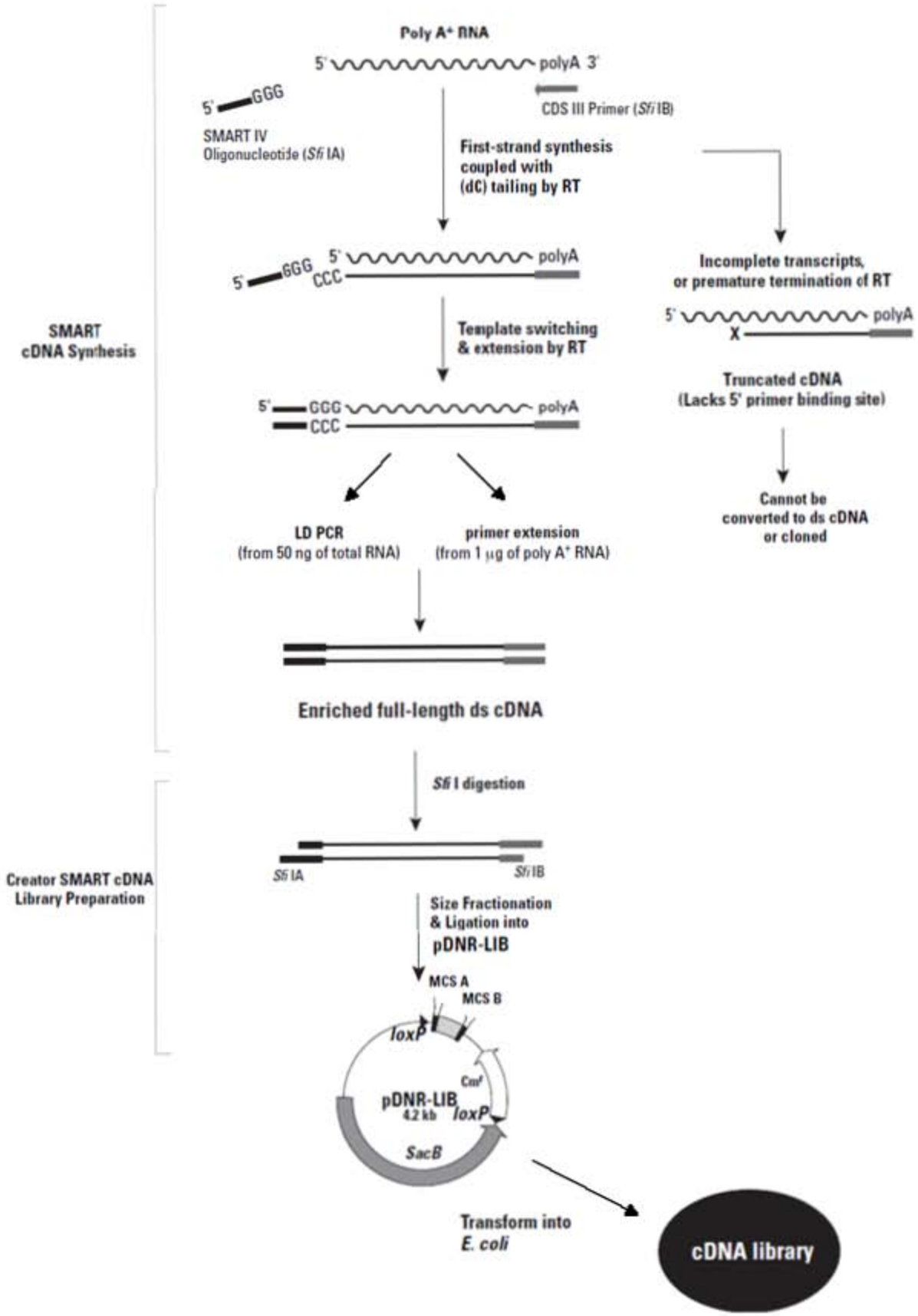
D.1 pDNR-LIB plasmid map



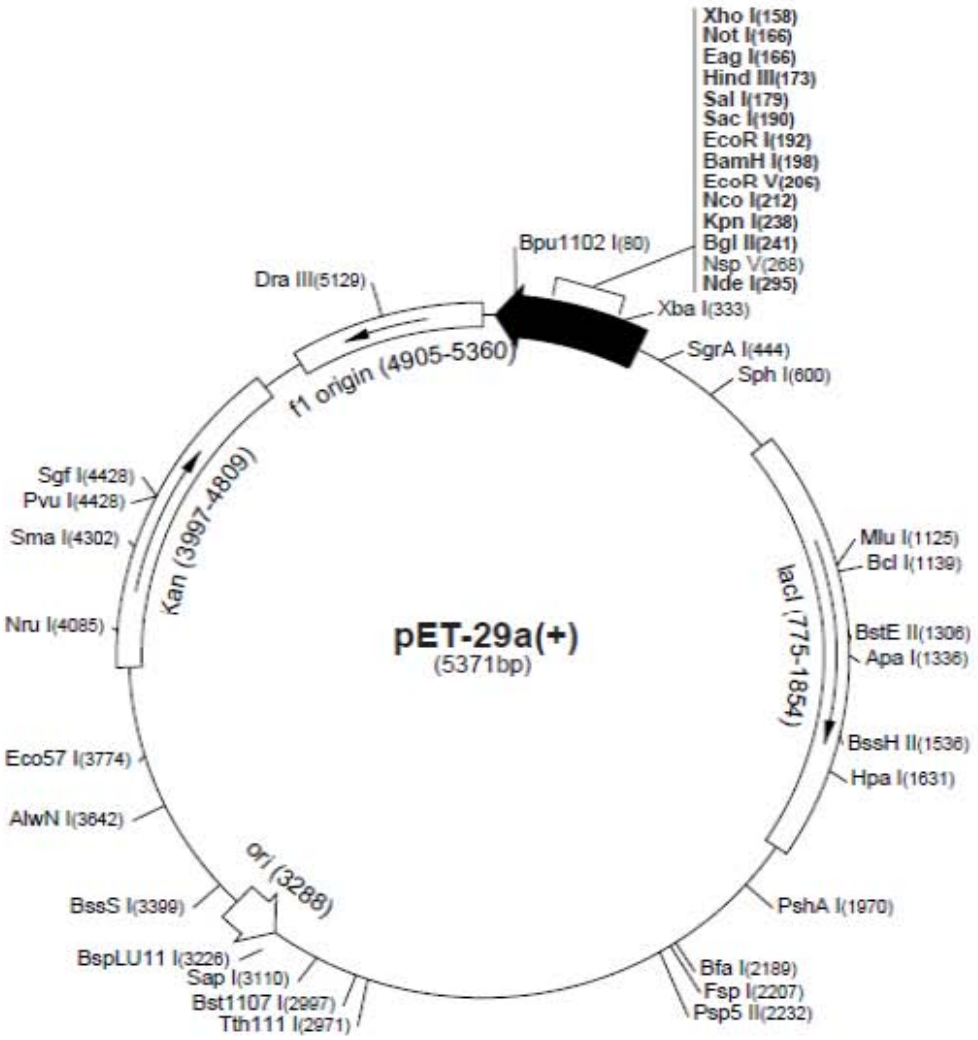
Restriction Map and MCS of pDNR-LIB Vector. Unique restriction sites are shown in bold. MCS A is shown in frame with the *loxP* site. The last four nucleotide bases of the *loxP* site can be seen at the left hand side of MCS A in bold. **Note:** The stuffer fragment is replaced by a cDNA insert when the library is constructed.

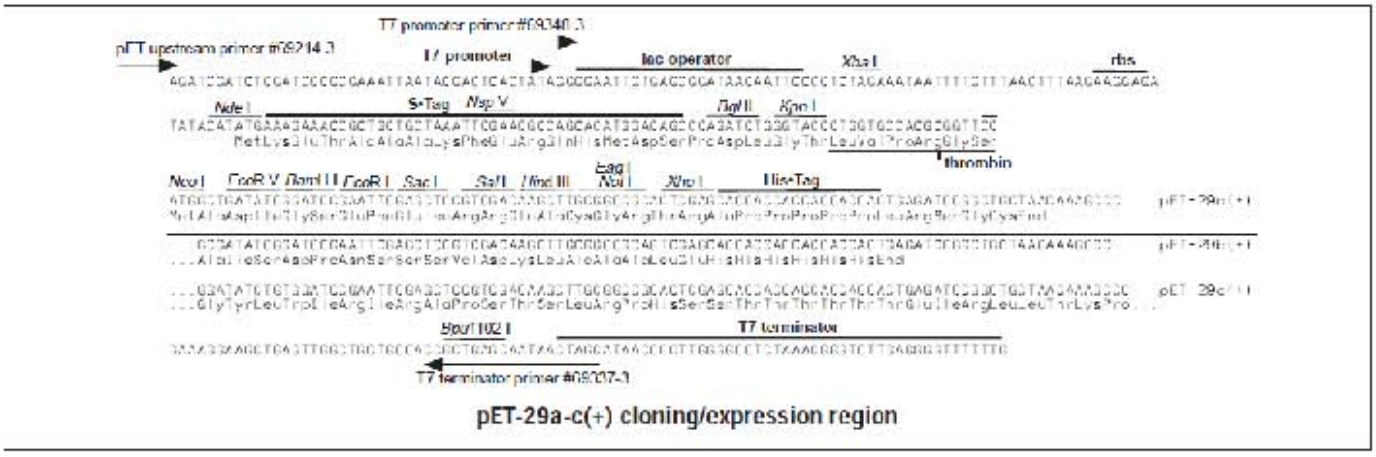
Cloning vector supplied with "Creator™ SMART™ cDNA Library Construction Kit" (CLONETECH Laboratories, Inc.)

D.2 Creator™ SMART™ cDNA library construction kit protocol



D.3 pET29a (+) plasmid map





APPENDIX E

GRACILARIA GRACILIS DNA SEQUENCES

E.1 DNA SEQUENCES OF PUTATIVE DEFENCE GENES

These following sequences correspond to all the cDNAs that represented differentially expressed *G. gracilis* genes in the cDNA microarray experiment (Chapter 2; Table 2.6). Sequences annotated to unknown functions or hypothetical proteins were not presented while redundant sequences were merged and represented by one only sequence. The DNA sequences are represented as follows: > Clone ID – putative functional annotation; DNA sequence.

> 4 – Thioredoxin

1 CGGCCGGGGG AGGCAGCGTG CGGTCAGCGC GGACAAGTTG CAGCACATAT ACACAATATG
61 GTGAACGAAA TCAAGCGAAG GTCAGAGTTT GCGCGCGCGG TCAACTCCAA GCCGGTGGTG
121 GTGGTGGATT TCTTCGCAAC ATGGTGCGCG CCATGCAAGA AGATTGCGCC GGTGCTGGAA
181 GACTGGGAGG TGGGCATGGC GGACAACAAG GAGGACTGGC TGGTGAACTT CTTCAAGGTG
241 AACGTGGACG GGCTACCAAA GCTGGCGCGC GACTTTGACA TTTCGGCGAT GCCCACGTTC
301 ATCATCTTCG TGAACGGCGA GCGCACGGCG ACGGTGACGG GCGCGGACCG CAAGGCGCTG
361 CAGCGCAAGA TCGAAGACGC CATGCAGCAG GTGACGCAGC AGCGCCAGCG CAAGTAAATG
421 CAACTTTTTG CTTTTACGAA AAAAAAAAAA AAAAAAAAAA AA

> 828 – Hypothetical trans-membrane trafficking protein

1 GGGGAATGCA AATACCTTGC TAACCAGAGC CAATATCATA TATGCATGTG TTCTCGAGAT
61 CATTCTCATC ACTTGTGGAT CTGATGTGAT CTGAGAACTC TTCTCGAGTT CTGTGAGGTA
121 TGCGTACTTT TTGTTGAGAA GATATACGAA TTGGTATTAT TCAAGCGAGC TGTGAACACC
181 GATGTCTTTC AACAAGCTTT TATGCAGTGT TGTACTTGTC TCCCTCTGTT TTCGATCTAT
241 AAGCGCTCTC CAGTTCGTCT TACCAGCGAA AACGCGAAAA TGCTTCCGGG AGGACATTCC
301 TCTCAGTACA AACACATTGT TTACTTATAC TGTTGCCCAG GGTAATGGAG AAATGCCCGT
361 ATCTGTGCGA ATAACGGACA TGACAGGGAA GGTGATTCTC GTTCGTCAAG CGGTTGACCA
421 TGGAGTGTTT ACGTTTCAAT CCCCGGAAAA AATCCCCAAT GTCCCTCAAA AGAACGATTG
481 GTCCTTACGC GACGAAGATA CAGACGACGA TGCCTATTTC CGGGCGATAC CTGACGGGGC
541 CGGAGATAAT CGCATTCCCT ATATGTTTTG TTTCGAGCAC G

> 116 – pG1 protein

1 GGGGTTTTTT AAAAGAGTTT GATCCTGGCT CAAAATGAAC GTTAATGATT AGCTTTACAC
61 ATGCAAATTA AATAAATTTA TTTTGATAAT TATAAATTAA TAGTGAACGG GTGAGTAAAA
121 TATAGAAATC GACCTCAAAT AATTGTTTGA TACATGTAAA GATTTTATCA ATTTGAGAAA
181 AGTCTATAAA AGATTAAGTT GTTGGTATAA TTAAAAGTTT ACCAAGCTAA TGATCTTTAG
241 TTGGTCTATG AAGATGTACA ACCACMTTGG AATTGAGATA CGGTCCAAAC TTTATATAAA
301 GGCAGCAGTG AGGAATATTG GGCAATGAGC GCAAGCTTGA CCCARCTAAA TCATGTATGT
361 GGGTAAAACT TTTTGTTTTA AAACATTAAA ATTAAAAAAA TAATGATAAC TTAAAAAAAA
421 GTCCTGGCTA ATTTCGTGCC AGCAGCCGCG GTAAAACGAG AAGGGCGAAC GTTATTTGGA
481 ATTATTGGGC GTAAAGCATA TGTAGGTGGA AACTTAACTT TTAGTTTAAA GCTTAAACTT
541 TATTTTTAAG TAGGCTGAAA A

> 216 – Putative 23S ribosomal RNA

1 GGGGGGACAG TATTTCAAAA TTTTGTTATG CTAGATGAAT CAATTGAAAA ATTGTACCGC
61 AGAAGGTTAT AGTCCTGTAG TCGAAAGCAA AACAATATTT TTATTGAATC CCAAGTAGCA
121 TGGGACACGT GAAACCCCGT GTGAATCAAC GAGGACCACC TCGTAAGGCC G

> 685 – Novel asparaginase

1 GGGGGCTAAA CAGCTGCCTT CTTCCACCGC TCACGTGGAG GCAGCAGTCA TCTCACCGCC
61 CGACGTGCTT CCTTCTATTC ATAAACGCTG CCTAGTCTCC CCAGCACCCT GCCTCGTCCC
121 AACGCTAGAG CTCTAGCTTA CTTTGCAAAT CCTTCAACAA CCACGCAAAC AGCATGACCA
181 ACCTTCCTGC TCATTCCCGG TCGCACCATC GCACCCCAGT CATCATTGTC CATGGTGGCG
241 CATGGGCTAT TCCGGACGCA ACAGCCGACG CCACAGAAGC ATCAGTCCGT CGTGCCGCCT
301 CCATCGGCCA TGCCTTACTG ACGCAATCGC GTTCCGCGTT AGATGCCGTC RAGGCTGCRG
361 TGCGCCATCT CGAGGAAGAT CCGCTGTTCG ATGCGGGTAT CGGCTCCTGT CTCACTGAGA
421 CCGGTACCGT CGAGATGGAT GCCGCCGTCA TGTACGATGC ACCGTCGGGT CTCCGCTCCG
481 GCGCTGTAGC ATGTGTGTCA AATTCTGTAC ACCCTATTAG CGTGGCGAGG GGAGTGATGG
541 AGCAGACTCA GCATTGTCTG CTGGTCGGCC GTGGAGCCGA TGCGTTCTCA AATCATATGG
601 GACTCGGAGG TGCCTCGAGC GATCAGTTGG

> 623 – Phosphoglycolate phosphatase

1 CGGCCGGGGG TCATTGAACC TGTTGCAACT CTTTGCCCGA CATCATTGTA TTCCATAATG
61 GCTTTTGTTC CTGCATTTTC TAAGAGCACC ACAACGTTTG TTCCTGCCGA GAAGTTCACT
121 ACCACCAAAC GCCATATTTT CCAATCCGCA AAGCAAGCCG AGCGAAGCGT GGTTCGTGGA
181 CAACCTCGTA TGACTGTGGC TCCAGATATG GCTTCGATTA CACAGGTGGC GAAGCTCACA
241 AGCCCATCCC AACTCCTGTC GAAAACAAGT ATTTTCATTT TCGATTGCGA TGGTGTTGTC
301 TGGCGTGGCG ATTCTGTCAT TGACCGAGTT CCCGAAGTTC TGGAGTATGT GCGCTCACTA

361 GGGAAACAAG TATTCTTCGT CACAAATAAC TCCACGAAAA GCCGGGCCGG ATATCTAAAG
421 AAGTTTACAA AGCTTGGGTT GGATGCCCAA GCAGAGGAGA TATTCTCCTC TAGCTTCGCA
481 GCAGCTGCCT ATCTCGAGCA GACTAATTTC AAAGCCACTG GAAAAAAGGT GTACGTCATC
541 GGCGAGGTGG GTATTCCGGA GGAACTAGAC CTGCTTGGTA TTCCGTACAT TGGTGGCCCG
601 GCTGATGCTG AAAAACAACC AAACATGGCT TTTGGCGGTC GACTGGAACA CGATCATGAT
661 GTAGGAGCAG TCATTGTAGG ATTTGATCGG CACATCAACT ACTACAAGAT TCAATATGCA
721 CAGCTATGTA TTAACGAGAA CCCGGGTTGC AAGTTCA

> 233 – Peroxiredoxin-like

1 GGGGGTGGAC ATACTTTCAA CTCTCACTGA ATCCCAGAAC CAGCCCATCG ACATGACTGT
61 CTCTACCGCA CCGTCGCTCG CCCCGCAGCT CCAGGCCTTC AAGGAAAAGT TCCTGGGTTC
121 AGCGTCACCA GAGATGGTGG CCACCGTTTC CAGGACGTCG GCCGAGCTGG AGGAAGCGAA
181 CCCGATCAAC CCGCTCAAGG TCGGCGATCG CGCGCCGGAC TTCACGCTGC CGAACGCCAC
241 CGGCAAGCAG GTGACGCTCT CCTCGCTGCT CGAGAAGCAT TCCGCCGTAG TGATGACGTG
301 GTATCGTGGA GGGTGGTGCC CGTACTGCAA TCTGGCGCTG CGCGCCTTCA CGCAGCTCAA
361 CGGCGAGATT GAAGCCAGCG GCGCCAAGTT GGTGGCGTTG ACGCCGGAGA CGGCTGACGA
421 GAGCCTTTCC ACTTCGGAGA AGAACGAGCT GAATTTCGAG GTGCTGTCAG ATGACGGCTC
481 GACTGTTGCC GAGAAGTTCG GCGTTGTGTT CACAGTTCCC GACGATATCA AGACCTTGTA
541 CAAGGGATTC CTCGTTGACC TCGATAACAA GAACGGTAAC GGGGGCAAGA GGATGGCGCG
601 TCTGCCAGCA CCGGCTACGT ATGTGATTGA CAAGGAGCGT ACCATCAAGT ATGTCTTCGC
661 TGATCTGGAC TACACGAAGA GAGCTGAGCC GTCGGAGGTA ATGGAGGTGC TCAAGACTCT
721 TGCTCAGTAA CGGAGCATAG AATGCTTCTG GTAGTTTAAG AACCACGGCG TTGTAGATAT
781 CGTAGCGTCG GCGTAAAAGT TAAAGAGTCT ATCCATCTCT CAAAAAAAAA AAAAAAAAAA
841 AAAAAAAAAA

> 422 – Putative ATP synthase CF0 subunit I

1 GGGGGGACAG TATTTCAAAA TTTTGTTATG CTAGATGAAT CAATTGAAAA ATTGTACCGC
61 AGAAGGTTAT AGTCCTGTAG TCGAAAGCAA AACAATATTT TTATTGAATC CCAAGTAGCA
121 TGGGACACGT GAAACCCCGT GTGAATCAAC GAGGACCACC TCGTAAGGCT TAATATTCCT
181 AAATGACCGA TAGTGAACTA GTACCGTGAC C

> 1034 – Cytochrome c oxidase subunit I

1 GGGGACCAAC GGTAACGCTA CCACGCTACC ACGGGAGAAA GTTGTCTTTC TTCATGGCCA
61 TTTCCCAATC TTGTATCTAA ACGATCGACC ACGTGGTTCC TCCGGTAACT GCTCACGTCC
121 GGAGGCATGT GTGGACGGGC AGCGCAATAA AACGTGCATC AAAGAAATAG GCAGTATGGS
181 ACTCTATCAG CAAGCTCTTT GAGTCACAGA CTGTARCAGA GCTTCTCTCT ACTAAGTACA
241 CGGGAACTGG GTGTATCATA GAGTCGGCAC CTGACCGGTG AAAGACTGTG GCAAAGTGAG

301 CTTGGSGTGT ATTGCAGCGT TTTTGCGTCT GATAAATTCT GGCGGAGTCG CCAGCATTCT
361 ACAAGATTTC ACTTCAGTTG GAATCATGTC TTGACTATCA GGAATGTCMA TTGAGATTTG
421 AAGTTGGTCT GTAATCGAGA TTGCTAGTCC TCGATTGTCT ACGCTTTCGT TTTCACGGTG
481 GCCTGGAATT TACGAGATAT GCACCGTATT TTTGGACCAG ATATATAACG TTTATATACA
541 TACAGCTTGT ACATAATAGT AAACATGTTT ACTACTATAA GAGATACTT

> 924 – Polyubiquitin

1 GGGGACAAGT CGACCTTCAG TTCGCCTTAT TACACACACA ACCGCCTTCT TTCGCTCGTT
61 CCCAGCTTTC TCTCTTGTTT TCCACCATGC AGATTTTCGT CAAGACTCTC ACCGGAAAGA
121 CCATCACCCT CGAGGTCGAG TCTTCCGACA CCATTGAGAA CGTCAAGACC AAGATTCAAG
181 ATAAGGAAGG CATCCCGCCG GACCAGCAGC GTCTTATCTT CGCCGGCAAG CAGCTCGAGG
241 ACGGTCGCAC TCTTTCTGAC TACAACATCC AGAAGGAGTC TACCCTTCAT CTCGTGCTCC
301 GTCTTCGTGG TGGAATGCGG ATTTTTGTCA AGACCCTCAC CGGTAAGACC ATTACCCTCG
361 AGGTCGAGTC TTCTGACACC ATCGAGAATG TGAAGACCAA GATTCAGGAC AAGGAGGGTA
421 TCCCTCCGGA CCAGCAGCGT CTTATCTTCG CTGGAAAGCA GCTCGAAGAT GGTCGCACCC
481 TTTCTGACTA CAACA

> 134 – tRNA-dihydrouridine synthase

1 GGGGGTTGCA CCAAGGGTAC TAACTTGGGC TACAGCTGCG CGCGCCCGTC GCCATGTCGG
61 GGGCAAGGCG TGGGAGCAAT GGCGCCGGCT GGGCAGTCCG CGCTACGTCG TGGCGCCCAT
121 GGTGGACGCG TCCGAGCTGG CGTTTCGCGA CCTGTGTCGC AACTACGGCG CCAAACTGGC
181 GTATACGCCC ATGCTGCATT CGGCCTTCTT CGCGCGCGAC GCCAAGTATC GCGCCGAGCA
241 TTTCACAACC CATCGCCATG ATCGCCCTCT CGTCGCGCAG TTCTGCGCCA ATGACCCCGA
301 TACGTTCGTC GCGGCTGCCA AGCTGGTGCA GGCGCATGTG GACGCCGTGG ATTTGAACCT
361 GGGCTGCCCG CAGGGCATTG CCAGACGTGG CCGCTACGGC GCCTTTCTCA TGGACCATCT
421 CAACGTCGTC TCGCGCATTG TTTCAAACGC GGCCGCGGAG CTAGATGTGC CGGTCTGGGT
481 GAAGATTCGC GTCTTTGATG ACTTGCAACA AACGCTGCAT TATGCCACTA TGTTGGAGCG
541 TGCCGGCGC

> 282 – Putative cysteine desulphurase / Glycerol-3-phosphate dehydrogenase

1 GGGGGGGGGA AAAGTTGGAA GACATCCTAA GTACATCGAG ACAGATTGCG GAAGGAGTAG
61 CAACGGCAAA AGCGCTGGTA CGATGGTTGG GAGGTGGGGA AGAAAACTCT CCGATGGCGG
121 TGAGGGCCAG CCTGAAGTAT CCAATTCTGT ATGGGGTGGC CGCAATTTTG GACGGAAAGA
181 TAACGCCAAG AGAGGGTCTC CATATGTTAC TAGCTATGCC GCCTCGTCAG GAAGACTGAA
241 ATCGAGCTTA AAAGGTAAGA AGTGAACTAA TTGACATGGT CGACAGGAAT TCCTTTTGCG
301 AGCGTCGGCA GAAATGCCTT CGACTTGTGA CCGTGTTCAT GACGTTACCT TGTTCACAAT
361 GCTTCTCAAG CATTGGCAAA AGCGGTCCAT GTCCTCGTTT GTGTGAGTTT TGAGATTCAC

421 ATCTACTCTA ATGGAGCCGC TTACACCCAG CTCGGCATGG GCTCGGTGAA CACCATGGAG
481 CCCAGTTCTC ACATACACCC CATTCTCCGC GAGTTCGATT GCCACTCGGG AAGAACTAAC
541 CCCTTCCACG TTGAAGCAAA CAAACGGAGC CCGACGCTCA GCAGACCCGT ACACATCCAC
601 TTTGTCAAAG GAGCGAAGCT GTTCATATAA ATATATGGCC TGACTGCCCA AAAAAAAAAA
661 AAAAAAAAAA AAAAAAA

> 1003 – High light inducible protein

1 GGGGACTCAT TCGAACTTCT TCACATCTTA TCCTTCAGAC CATTCCAACT CTCTCATCTT
61 GTACTTCAAA GCAACAATGG TCTCCACAGC TTTCGTCGCC GCTACCCCGG TCCTCTCCCG
121 CACAACAACC TCTGTTGCCT CTACATCTTT CACATCCCAC CGCACTCGCG TTAACCCGCC
181 TCGTCGTTCG GCCACGCTCA TGATGGCCGA TAAAGAGAAG ATCCCCCAAG GCCTTACAGC
241 CTTCTCTGAA GTTCTCAACG GTCGTGCCGC TATGCTCGGA CTCGTACTTG CAATCACCAC
301 CGAGGCCATC ACCGGAAAGG GCATCATTGG TCAACTTGCT GCCCTTGGGG ATATCTCCGC
361 CATTACGCAC GCATTGGGAC TTTGAGCCTG ACTGCCACGC TGTGTATGCT ACGTAGTGGA
421 TCTAAATTGT ATGAGTAAAT AAGCTTTCAT

> 252 – SNF2 family chromodomain-helicase DNA-binding protein

1 GGGGTATCGA AGAAGATCTT GCGAAATCGC TTCTTGACGA CTGCCTGTCT CAAGCTCAAG
61 AGGCTGTTGA GAGCTCAAAA AGGAAGGGAA GGAACTTGAA CAAGGACCCC GAATACGGTG
121 ATCATAGAAT GAATGGAAAG GATTCAAAGT CCAAGGCTCT ACGTGTGCAA ATCGACATTC
181 TTGGCGAGAG GAGGGTAGAT GCTCACGATC TCTTGAAGAG ATGCCGAGAC CTGAAGATGC
241 TACGAGATGC TATCGGATCA TTTAATTCAG ATCTGCAGTT CAGATTGCCC GGAGTCATTA
301 GGCCACCGTC GTTTGGTATT CGATGGAAGC GGTATCATGA CGCAATGCTG CTGGTTGGAA
361 CTTGTAGGCA TGGATTCGGG AACTGGACAC AGATCGCCAA GGACGACCAG CTGGATTTAG
421 GTGATAAGAT GAATGTAGCT GGTAATTCGG CTCAGGCAGG AGCGCCAGAT ACGACGAAAC
481 TGGCACGGCG AATCACGGCG CTGCTCCGAG AGCTTGAGCG CGAATCCCGC TTGCGAGCTG
541 CTGCCGATCG CAG

> 476 - GDP-fucose transporter 1

1 GGGGGCGAAG AAGTTGTTCG TCGTTGAGGC AGACGGCGTT TGCACGGCGA GCGAGTTGGG
61 TGTTGCTGCT GTTGCTCTAT TCACTATGTC CGTCGATAAT GCCAGTTCGC AGCCGCTGTC
121 GCGCCAGATG AGCAACATTG CCATGGCCGT CGCGTTCTAT TGGGTCATTT CTATCTCCAT
181 GGTGTTCGCG AATAAGACCT TGCTGGGTGG ACAGGATCGT TCGTCGGCGC CTTTCTTCAT
241 CACTTGGTCG CAGTGTGTTG TCACTGTGCT GTTTTGTTAT ATTGCTGGCA AGCTGCGCCT
301 CGCCAATGTG CCGCCTTTTG AGGTGCGCCC CGACGTTCTC AAGCAGATGC TTTCCTTGTC
361 CTTTGTGTTC ACGTCCATGA TTGTCTTCAA CAACTTGTGT CTTAAGTATG TCGAGGTGTC
421 TTTCTACCAG GTTGCGCGAT CGCTCACTAT TGTCTTCAAC GTTGTTTTTG ACTACGTCGT
481 TCTGGGCCAG GTTACTTCCT TCAAGGCCAT GGTGTGCTGT GCTTTCGTCG TGTCCGGTTT

541 TGCTCTTGGA AACGTGGAGG AGATGCGCTG GTCTCTTCTT GGAGTTCTTT

> 1041 – CHCH domain containing protein

1 GGGGAACGCT TTTCATTTTG GACTAGATGC CTCGCCGCAG CTCACGCCCT AGTGGTGGTC
61 GCCGTCCGGC ACCAGCCAGA ACTCCTGCGC GCCCACCCGC GCGATCTGCC GCAACTCGTA
121 CACGTCCTGC GCCATCATCC ACCCAGACGC GTGCGAGATC TACCGCTCCA GCCACACAAA
181 AGTCAGCAAC AGCCCCGCCG CCTGTTCCAG CTCAGCCCGC CCAGCAGGGG TCTGCCATGG
241 GCGGTTTCAT GGCCAACGTT GCTTCAACCG GTCTGGGTGT TATGGCTGGC CGTGCTATGG
301 ATCGTACCTT TTTTGGAGGC GGTGGATCTG CGCCTGAGCC AGTGGCAGAT GCTGAAACCG
361 CTGCTGCTCC CGCACCATAC GAATATTCGG CGGAAGAACC AGTACCGGAC ACTGTTTGCG
421 CTCGAGAGGT TTCGGAATTC AGGAAATGTA TGGAGAGGAG CGGTTCGGAT GTTGTGGCAT
481 GCCAGTGGAA TATTGACATT CTCGCAGCGT GTCAGAGGCA GCAGCAGGAG TTAGGCTTTG
541 CTGGCGCAGG TCGTCTATGA CCATTTTTAA ACTGCTTGCC GAGCAATTCT GAACCTCTAT
601 CTAGACT

> 1010 – Cytidine and Deoxycytidylate deaminase family protein

1 GGGGGACGAC ATATTCGTAC GTGGCCGAGT TCTGGTGCGG TCCGCCGCTG GTGGCGCGGA
61 GAAATCCGCT GTGCTTTGTT TTTCCTACCG TTGGTAACGT ACTCGCCGTC GCAGACGCTC
121 TATGGCTCTA TGAGAGCATG CAGCGCGCCA TCTGCGAGGC AAGGAACGCT CTGCAACGCG
181 GCGAGGTACC CGTGGGTTGC GTGCTTTGTA CAGCGGACGG CACCGTGGTT GCAACCGGCT
241 CCAACCGCAC CAACGAGCGA GGAAACGCCA CCTCGCACGC CGAAATGGTG GCGTTTGAAC
301 AGCTCTCTCG ACTCAACAGG TCGTTCGAGC AAGGCGAGTT GAGCCTGTTC GTCACGTGCG
361 AGCCGTGCAT CATGTGCGCG TCGGCCATTA CGCAGATGCA AGTGGTGGGG CGTGTTGTAT
421 ACGGCTGTTC CAACCCTCGC TTTGGCGGCT GCGGCTCTGT GCGATCGCTG GCTTTGTACG
481 AGCGCTGTTC GCAAGCGCAC GTCCCCAAGG TGCAGGGCGG CGTCGAGGCT GCTACTGCTG
541 TGCAACTGCT CGCCGAG

> 466 – Glycolipid transfer protein domain-containing protein 1

1 GGGGGAATCT CTGCGTACAT ATCCCGACCG ACGTTGAGTC AAGCTCAACC TCGCGCGCAT
61 TGGACTCTTT GTGCGATTCT GCTCCGCCCC GCTGCGTGTG TTCCGCGCCT AACTTGGTTC
121 GTGCTCGTGA GAAGGCTCGT GTTGCGTATA CGTTCCCGCC TTCCGTCGTC GTAATCCGAT
181 TCGCCTCACG GTCGCTCGCA GCATTGCGTA CGCATGGATT ATCTGCGTTC GCGCGCGTTG
241 CCCGGCGTGC GCGCAGGCAG CAACCCGTCC GCCGGCGACG GCGTCGCGGT GGAGCAGCCG
301 GTCGGCGCCA TGTTCGACTT TGAGCAAATG ACGTGTGCGT TCGAAGCGTG CCACACCCAT
361 TTCCTCACCT CCAACAGCAC GCTGCCGACC ACGCAGCTGT TCATCGCCGC CATGCAGCAG
421 CTGTGCCACC TGTTCGACAC GCTGGGCGCG GCGTTCACGT TCGTCAAGTC GGACATTGAG
481 TCCAAGCTCA TCGTGCTCAA GCGCTTCCAG AGCACGCAGA GCGCGCACTT TTCCAGCCTG
541 CAAGCCGGCG TGCAGCACGA GCTGCGCAAC GACGCCACGT GCGCCACGGG CTTTGTGTAC

601 GAGGAGTCGC CCA

> 428 – Putative serine protease-like protein

1 ACGCGGGGCG GTTCGATAAC AACGATCGAG GACAGTTCAC GCGTTGTCGT ACGACGAAGG
61 TAACTCTTAC TCCCGGACAT AACTTCATTG AAACCGATAT GACTCGCTAT GCCATTGCGG
121 ACGCGCTGGT AAGATCGTCG CAATCTTTAT CAGAGGCCAT CTCGTGCTTG TGGGAACTTG
181 ACACCAATAG GTTGGTTCCC GGCAAACATT ATACCATCTG CTTGCAGGGC GAAACCGAGC
241 ACCGTTCTGA GGATGCTGCG CGTGTCCCGC TATTCGAATA CTTCTCGGAC GATGTCTGGC
301 AAATAGAGTC TTACAGCCTC TTCAAGCGAT TGCTAGATAA CTACACAGTG AATCAGGGTG
361 TCCCGGAACG TGTTTCAGAG GAGGAGAAGG AGGAAGAGGA GCGATTCTTG GAATGCATCT
421 GTGACACTGA TTGCATTACG TTTGTATATG AGTGGCTACG AGAGAACGGA TACAAAGCCG
481 CCTCCTCTAT GGACGAGTTC AAACAAGTTC TAAGTGATTT GTGGTTTGGC ATGTATGGTA
541 GGGGCCGCTA CAGGGATTCT TCCGCGTTTG AGCACGTGTT CTGTGGTGAG ATCAGTGGCG
601 ATGATGTCAA AGGTCTCCAC AACTACATTC AAGTGTACAT TGAAGAGCAA CGTGGAAACT
661 TCGATTACAT GGGATATGTC GACTTCSAGG GAGATCTTTG TGGCGCACCG TTGTCGAACC
721 AGCAGGCTTT GATGATTCGC TTCAAGTGGT TTAGTTGTCT CAAGAATACT TCTTCTATGT
781 TTGTTGGCAC TTCTCCGGAA TTTGAGATTG CCTTGTTTTC TTTGCTTTGG TTCACTCTGG
841 GAGTCACCGA AGATGTTGAA CAAGAATTTC AGCTTGGGCC TTACTTTGTT GAATTAAGGC
901 TCTATTCTCA TCGCTCTAAC ATGTGCACTG CATTTCCCTG CTTGAAAGGG GTCGATACAG
961 AGACCTTAGA GGATAGTCAA AATGAAACAA AGAGAAGGAA TGAGGAGCAA TATGCGGAGC
1021 ACCAACAGGA GGGCGATCAG CTTGAGCAGA TTGACGCCGC AGTTCTGCAG GATTCGGAAG
1081 AGTTCCCACC ATTGGGATAG AGTTGATAGT GTAGTTGCTG AAATAGCTCT GAGATAAAAG
1141 CTTCTTGCCA TAAAAAAAAA AAAAAAAAAA AAAAA

> 991 – Phosphoserine phosphatase

1 GGGGAACTGA GTTTCAATGT GACCTTGCCT TGCATAAGGG AAACATATCT CGCAAGGCCA
61 AGAGAATGGT TGTCTTTGAT CTCAGCTGGA CGCTTGTTCA ATGCGATTCC ATTAACGTCT
121 TGCTGCACGC AGCCGATGTA GATGTACCGA AAGAGGAGGA GCAAAAATTC CGTAGCGGAC
181 AAATCAGCGG TTCGGAGTGG GTACGCTTCA GAGTGCAGTT GCTTAAAGGG TTCGATGCGC
241 ACACCGTGAA CAGCAAAGCC GTTCAAAGTT TGGTTTACAC AAACGGCGCT GTTCAACTGT
301 GCAAGGGTTT GAAACGCCTG GGATGCAAGC TTGCAATTGT GTCCTCCGGT TCCAAAGACA
361 TTGCTATGGC TGCAAAGGAA GCTCTTCACC TGGACTTTGT STTTDGAAAT GTTATTGAAA
421 TTGATTCCGC GGGAAGGTTC ACTGGCACTC TCAAGGAACC TGTTGTTGAT GCTCAGCGGA
481 AAGCTGATCT GATTGCAATG TTAGCCATGC AAGAACGTAT TGACACCGAA CAGGTCATTG
541 CTGTGGGCGA CGGTCCGGTG AGTGCGAAGA TGCTCGAGAG CGTCGGTATG AGCATCGCCT
601 TTGATCAACC TGATGCGCTG GATTCCGTCC ATAGTGGCAG AATTGGTAGC AAGAGTCTTG
661 CAAGCGTTCT GTACTTGCTC GGTGTTTCAG GACATGATTT CCGCACAGTT ACGGCAAGCT
721 AGCTAGGAAT ACTATTTGTC CCTTTCGTGA TTTAGTCGGC AGGGCAGCCT CGTTATCTTC

781 ACAGCAACAC TACTCTTGTA ACAAAGCCAC GATGCAGGTT CGTTGAACGT TGTCATCGCA
841 ACGCTCAGAC CATTTGGACC TAGTTTCCTG TGTATGGAAG GAACGCATTG GTTGTGAGAC
901 TCTAAGGGTG ACTGCTGCAT TTCCGACCCC TGGAACCCTC CCCGTGGTGA CATACACTGC
961 CACCGGTAAC TATAAGACAC CCCCAGCGTT TATAAAGAGA GCACTAACGC GTGCTCAAAT
1021 TTAAAAACCA TTTATTCTAT AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

> 897 – Phosphotidylserine decarboxylase

1 GGGTTCGCAA TTGAAGAAGA GCGGAAAACA CACCGAATTG AAGGCAAAAC ATCCAGCATG
61 GCCGACCACT TCGATGCATA TGAAGAGGAG TTTGCTTACT CTATGGGTCT CCAAGCCTAT
121 ATCTTTGGCA TGTCTCTCAC TATCTTTGAG CGTGAACGCA AGGTTCGCCT CTCTCCTGAT
181 TTCGATGGAC AACAGGGCGT CGCACCAGTT GCGAGGCTCA ACGATATCGG ACACATGCCC
241 GCACTGGCCA CGAGCGATGG TATACTTCCC TATACTCCCA ACAACGACAC GGTTTACAGC
301 GGTGCCTTGC TTGAGCTCAT CGACGAGCCA ATTATTCTCA CTGCTCCAGA CATTTTTGAC
361 CGCTACTGGA GCGTCGAAGT TGCTGACTGC TTTACCGAGA ACGTTTTTTA CATCGGTTCG
421 CGAGCAACGG AGGGCAAAGG CGGAAATCAC GCTTTTGTTG GGCCGGACTG GGACGGAACT
481 CTTCCCGAGG GGGTCGTAAA ACACCAAATG GACTACAACA GCTGTATGTT TGCGCTTCGT
541 ATCGCCGTGA CAAAAAGCAC TCCCCATGCA GAGAAAGTAG ATACCCAAAG GGTTCGTGTA
601 CTACAGAACA GGTTCATATT AACTTCTTTG AGCAACTGGG GCACCAATCC GGGTAAAGCA
661 GCTGTGCCCG ATAGCATTGA ACCGCGACCA GACTACACAG GACCCTTGTC CTTCTTCCAA
721 ACTCTTGCTG ATCTGATGAA GGAAAACCCT CCCATCGAGA AGCATATCGC CCAAACGGAT
781 CCATTCAAGT ACATTGGTCT GGTTGTCGGG AAGAAATTTA ATCCCGACGC ACTGTCCAAG
841 CCGACTCAAC TTGGTTTGGC GCGCGCGGCC GCAATGGGAC CAAAGGTTCT CGACTGGAAA
901 GTGAAGTATC GCGGCACCGC CTATACGACG CGTTGGAATC GCCTCCATCC TGGCACCTAC
961 AACTATGACT ATCTCGATCG CGCGGCAGGA GCCACAGAAG GTCTTTTCGT CCATGATTAC
1021 GTGGAGAGCA CTTACTACAG CACTTACGAG AGCTGTGATC TTGACAGCGA CGGAAATCCA
1081 GCTGGTGGTA CTTTCTTTGA TAGTTCCAAG AAGTACACCA TGCATATTGG CAAACATCAA
1141 ATTCCAGAGA TCAACAAGAA GCTTAATGGC TTCTGGTCCT TGACGATGTA TGGACCGGAC
1201 TTCCAGCTCG TCAGCAATGA GATAAATCGC TTTTCAGTCA GCGACCAAAC CCTCCGGGCG
1261 AACGAAGACG GCTCGATCGA CATATACTTT CAGAGTGAGG ATCCTCTCAC CGAACTTGGG
1321 GAGTTAGCCA ACAACAACTG GCTTCCGTGC CCAGCGCCAC CGGGCCAGTT GTTCCGCTTG
1381 AACTACCGTA TATACCTTCC GTTAAAGACT GTTCTCCACC CGGTACCCAC CAACCGCTTC
1441 ATCCCTCCGA TTGTCGAACG TATTGTGTAG GAGGGCTGAG CATGCCTTCG AGAAGAGATG
1501 GACATACGAG TTTCTGACTC ATTATATCTA GCTCAGGTGG CTCCGCCGTT GTGTTTGGCA
1561 TGACAAGGGG GATCTCCAGA AGTTGATATG AGTACAGGAG GAGGACTTCC ACGATAAATA
1621 GGAAAGTTTA CCAATCCTTT TTAGAAGTTA AAGACTACTC TTGCAAAAAA AAAAAAAAAA
1681 AAAAAAAAAA AAA

E.2 DNA SEQUENCES OF qPCR REFERENCE GENES

> RG 329 – Sugar epimerase

1 GGGAAACAAT ATGACTTGTT GAAGTCGAAC AGTGGAGGAA GTTTTTTGGT GAAGTGCAAG
61 AGGGTCGAAG TCGAGCCAGT TTGTGAAGAT CCTCGGTACT CTTCTGCAAT CTTTCCATTC
121 AATTTATGTA CACTTGTTTA TGATACACGC GAAGTTATAC AAGGCACTCA GCTTAACTAC
181 GTTACTTGCA CAGATATGCA CTCAGTTCTC TTTCTCAGGC TATGAGACCG AGCACATCTT
241 ATTTTTTTCC ATCGTATAAA GATCACAAGT AAGAGCTTCG TCTACCAGTA GCGAGATATA
301 AAGGATGTTC GTGCCCGTCA TAGTCTCAAT TTCCAAAAAA AAAAAAAAAA AAAAAAAAAA
361 AAAA

> RG 801 – Conserved hypothetical protien

1 GGGGAATGAG CGCAAGCTTT CAACACGCGC GCGCAGTCAC ACGAGCGTGT TGCGCGCCGC
61 ACGCGAAGTC AGGCAGGCAC GTGCTCAAGC GTGAGGGCTG GTTGAAGCGA AGAGCGCACG
121 CAGTGGAAGC AAGTGAGAGA AAGAGCGCGC GCCGCAACAG TGGACTGCAC GCGGAGAAGG
181 CGAAGCAGCA GGCGGCGACA AAAGCAGTCG GAGAAGAAGG CGAGGCAGCA GCAGCAGCAG
241 CAGCAGCAGC AGCAGCAGCG CGCTGTATGT AAAGTTACTG TATCTACAAG TGGAACATGC
301 AGCAGCGCGT GCCAATTTCT ACGTCGCAAC TGGCACGAGG TTTGATGTGG AGCGAGCGAG
361 ATAGACAAGA TAGCACACGC CGCGCGTGTG TGTGACAGCC GAGAGAGTTT ACGAAGCGAG
421 AGGCGCGCGT AAACGGGGGA GACGCGGCGG GCGGCTTGTA CTAGACAAGG CCCGCCTTCT
481 TCTGGCCGAA CCACACGGCA AACAGCGGGA TGCCAATACC GGCCGAGACG CCGCCGCCTG
541 TCAAGTTGCG TGCGGTGCGG TGCGGTGCGG TGCGGTGCAG GGCACGGGCA GCGCGCGGGT
601 GCGAAAAGGG GGTCAGCGTC CACGAGCGGT GTGAGCAGCG CCGAGAAAGA GAGAAAAAAA
661 AAAAAAAAAA AAAAAAAAAA AAA

> RG 245 – Hypothetical protein

1 GGGAGAACAT GAAACCATAA GCTTACAAGC AGTAGGAGGA TGACTAATCG TCTGACTATG
61 TGCATGTTGA AGAATGAGCC GGCGACTTGT AGGTAGTGGC AGGTTAAGAT AGAGAATATC
121 GGAGCCATAG TGAAAGCGAG CTTTAATTAG AGCATTAGTC ACTATTTACA GACCCGAACC
181 CGGATGATCT AGCCATGGCC AGGGTGAAGC TTAGGTAACA CTAAGTGGAG GTCCGGACCG
241 ACTGATGTTG AAAAATCAGC GGATGAGCTG TGGCTAGGGG TGAAATGCCA ATCGAATCCG
301 GAGCTAGCTG GTTCTCCCCG AAATGCGTTT TGGCGCAGCG ATTGATACTA TAACTTAGGG
361 GTAAAGCACT GTTCCGGTGC GGGCTGCGAG AGCGGTACCA AATCGATGCA AACTCTGAAT
421 ATTAAGCGAG AAGTCAATCA GTGAGACAGT GGGGGATAAG CTTCATTGTC AAAAGGGAAA
481 CAGCCCAGAC CACCAGCTAA GGCCCCCAAA TAATTACTAA GTGAAAAAAA AAAAAAAAAA
541 AAAAAAAAAA

E.3 DNA SEQUENCES OF THE *GRACILARIA GRACILIS* 18S RIBOSOMAL RNA GENE

> 18S rRNA (Bird *et al.*, 1990)

1 CAACCTGGTT GATCCTGCCA GTGGTATATG CTTGTTTAAA GGACTAAGCC ATGCAAGTGC
61 AAGTATGAGT GAATTGTACA ACGAAACTGC GAATGGCTCG GTAAAACAGC TATAATTTCT
121 TCGGTGCTAA ATACTACTCG GATACCCGTA GTAATTCTAG AGCTAATACG TGCCTCCAAA
181 ACGACGCAAG TCGTGGTACA AATTAGAGAT ACAAGCCAAC TTGTTGGTGA TTCTAGATTT
241 TTTTTCTGAT CGCAATTATT GCGACGCACC GTTCAAATTT CTGACCTATC AACTTTGGAT
301 GGTAAGGTAT TGGCTTACCA TGGTTGTGAC GGGTAACGGA CCGTGGGTGC GGGATTCCGG
361 AGAGGGAGCC TGAGAGACGG CTACCACATC CAAGGAAGGC AGCAGGCGCG CAACTTACCC
421 AATCCGGACA CCGGGAGGTA GTGACAAGAA ATATCAATAG AGGGCCCGAT GGGTTTTCTA
481 ATTGGAATGA GAACAAGGTA AACAGCTTAT CGAGGAGCCA GCAGAGGGCA AGTCTGGTGC
541 CAGCAGCCGC GGTAATTCCA GCTCTGTAAG CGTATACCAA AGTTGTTGCA GTTAAAACGC
601 TCGTAGTCGG ATTTTGGTGT CTGACTTGGG TCGTCCTCGC GGACGCTCTC AGGTTGGGCG
661 CCTTTGTGGA TGGGAGCTAG GTGGTGCTTA ATTGGATCAC CTAGCTGCCG CCACCGTTTA
721 CTGTGAAAAA AATAGAGTGT TCAAAGCAGG CGATTGCCCT GAATACATTA GCATGGAATA
781 ATAGAATAGG ACCCGGTCCT ATTTTGTTGG TTTGCTTGAA TCGGGTAATG ATTAAGAGGG
841 ACGGTTGGGG GCATTCGTAT TCCGACGTCA GAGGTGAAAT TCTTGGATTG TCGGAAGACG
901 AACAGCTGCG AAAGCGTCTG CCAAGGACGT TTTCATTGAT CAAGAACGAA AGTAAGGGGA
961 TCGAAGACGA TCAGATACCG TCGTAGTCTT TACTATAAAC GATGAGGACT GGAGATCGGA
1021 TAAGACTGAT ATATGGCTTA TCCGGCATCC TTCGAGAAAT CAAAGTGTTT GCTTTCTGGG
1081 GGGAGTATGG TCGCAAGGCT GAAACTTAAA GGAATTGACG GAAGGGCATC ACCGGGTGTG
1141 GAGCCTGCGG CTTAATTTGA CTCAACACGG GAAAACTTAC CAGGTCAGGA CATAGTAAGG
1201 ATTGACAGAT TGAGAGCTCT TTCTTGATTC TATGGTTGGT GGTGCATGGC CGTTCTTAGT
1261 TGGTGGAGTG ATCTGTCTGG TTAATTCCGT TAACGAGCGA GACCTGGGCG TGCTAGCTAG
1321 GCGCCGTTAC TATTTTTGGT AGCGAGGCTT GCCTTCCTAG ACGGACTGTG GGCGTCTAGC
1381 CCACGGAAGC TCCAGGCAAT AACAGGTCTG AGATGCCCTT AGATGTCCTG GGCCGCACGC
1441 GTGCTACACT GAACGGGTCA ACGAGTTAGG ATATGCGAAA GCATTTCCCA ATCTCTAAAT
1501 CCGTTCGTGA TGGGGATCGA CGGTTGCAAT TTTCCGTCGT CAACGAGGAA TACCTTGTAA
1561 GCGCGGGTCA TCATCCCGCG CTGAATACGT CCCTGCCCTT TGTACACACC GCCCGTCGCT
1621 CCTACCGATT GAGTGGTCCG GTGAGGCCTT GGGAGAGCTA GATGAACTGA TTATTCAGAT
1681 CTTTTGGCTT GAACTTGGTC AAACCTTATC ACTTAGAGGA AGGAGAAGTC GTAACAAGGT
1741 TTCCGTAGGT GAACCTGCAG AAGGATCAGA A